

INDEPENDENT OFFICE OF
EVALUATION OF IFAD

Evaluation Manual

SECOND EDITION



IFAD
Investing in rural people

© 2015 by the International Fund for Agricultural Development (IFAD)

Printed on FSC paper
by Palombi & Lanci Tipografia srl, Rome, Italy
February 2016

INDEPENDENT OFFICE OF
EVALUATION OF IFAD

Evaluation Manual

SECOND EDITION
DECEMBER 2015



Contents

Acknowledgements	4
Foreword	5
Abbreviations and acronyms	6
1.	
Background	8
The audience and purpose of the Manual	8
The context – the challenge of rural poverty reduction	8
Aid and agriculture in the global development agenda	9
The role of IFAD	10
IFAD's theory of change	10
Evaluation in IFAD	11
The Evaluation Policy	12
Operationalizing the Policy	12
The Evaluation Manual.....	13
2.	
Methodological fundamentals	18
Approach, design and methods	19
Attribution and contribution	20
Techniques and instruments for data collection	21
Analysing and interpreting data	24
Benchmarking	25
Evaluation questions and inter-evaluator variability	26
Theory of change.....	27
Evaluation and hindsight.....	28
Transparency and the evidence trail	30
Formulating recommendations	30
Learning accountability and follow-up	30
Joint evaluations	31
Participation.....	32
Evaluation management	32
3.	
Evaluation criteria, key questions and ratings	36
Background	36
Evaluation criteria and their definitions.....	36
Evaluation questions by criterion	39
Rural poverty impact.....	39
Project performance	39
Sustainability of benefits	39
Gender equality and women's empowerment.....	40
Innovation and scaling up	41
Environment and natural resource management.....	42
Adaptation to climate change	42
Performance of partners.....	43
Corporate-level evaluations	43
Country strategy and programme evaluations.....	43
Project evaluations and validations: PCRVs, PPEs and impact evaluations	44
Evaluation synthesis reports	44
The rating system	44
4.	
Project completion report validation	48
Background	48
Objectives and methodology	49
Process.....	52
Reports.....	53

5.		9.	
Project performance evaluation	56	Evaluation synthesis report	
Background.....	56	110
Objectives and methodology.....	57	Background.....	110
Process.....	59	Objectives and methodology.....	111
Reports.....	61	Process.....	113
		Reports.....	115
6.		10.	
Country strategy and programme		Communication, dissemination	
evaluation	64	and learning	118
Background.....	64	Learning.....	118
Objectives, scope and approach.....	65	Communication tools and dissemination.....	119
Methodology.....	66		
Process.....	78	Annexes	125
Reports.....	83	I Independent evaluation architecture for	
		promoting accountability and learning	
7.		at IFAD.....	126
Corporate-level evaluation	86	II Indicative set of additional questions.....	127
Background.....	86	III Template and process for agreement at	
Objectives and methodology.....	87	completion point.....	131
Process.....	90	IV Conflict of Interest Policy for IOE Staff	
Reports.....	93	and Consultants.....	133
		V IOE selectivity framework.....	137
8.		VI Bibliography and references.....	138
Impact evaluation	96		
Background.....	96		
Objectives and methodological considerations.....	98		
Process.....	105		
Reports.....	107		

Acknowledgements

This second edition of the Manual has been developed following broad-based consultation with key stakeholders within and outside IFAD.

Within the Independent Office of Evaluation of IFAD (IOE), this Manual was prepared by a core team under the leadership of Ashwani K. Muthoo, IOE Deputy Director, who was ably supported by Catrina Perch, Fabrizio Felloni, Simona Somma and Pradeep Itty, Evaluation Officers, and Melba E. Alvarez, Evaluation Communication and Knowledge Management Officer. All other IOE staff also provided valuable feedback on selected chapters throughout the process. Linda Danielsson and Laure Vidaud, Evaluation Assistants, provided administrative support and conducted research on selected topics. A number of external consultants also contributed to the preparation of the Manual including Dorte Kabell and Roger Slade. In-house consultants, Xiaozhe Zhang and Maurizio Navarra also contributed to the production of the Manual. Brett Shapiro conducted a preliminary editorial review and Jill Baskins, from the Office of the Secretary, performed the final editorial review of the Manual.

IOE benefitted from the insightful comments and guidance of a high-level panel of experts comprising: Osvaldo Feinstein, Professor at the Universidad Complutense de Madrid, who also contributed to the preparation of the Manual as

an external consultant; Susanne Frueh, Director of Internal Oversight at the United Nations Educational, Scientific and Cultural Organization; Ted Kliest, Independent Evaluator and Expert; Uma Lele, Independent Scholar; and Kevin Watkins, Executive Director at the Overseas Development Institute.

The Manual was also peer reviewed by colleagues in the evaluation offices of a number of multilateral development organizations, including the African Development Bank, the Asian Development Bank, the Food and Agriculture Organization of the United Nations and the World Bank. The Centre for Development Impact at the University of Sussex in the United Kingdom and the Swiss Agency for Development and Cooperation also provided important comments on selected chapters.

Appreciation is extended to IFAD Management and staff who reviewed and provided valuable insights. Last but not least, before its finalization, the Manual was discussed at an informal seminar of the Evaluation Committee of the IFAD Executive Board, and comments from Committee members were duly taken into account.

Foreword

Over the past decades, there have been substantive changes in the way countries accomplish economic development, with a rising number of new actors – investors, the private sector, civil society organizations – playing an increasingly prominent role in the development arena. Social protection and inclusion, environmental protection and equal access to resources – in addition to economic growth – are becoming central to ending extreme poverty and fostering prosperity.

This year, 2015, the Millennium Development Goals (MDGs) have arrived at their target date, and the international community has developed a new set of Sustainable Development Goals, which will carry on the momentum sparked by the MDGs and fit into a global development framework beyond 2015. They contain far-reaching commitments to eradicating poverty, promoting better nutrition and fighting hunger, protecting the environment and encouraging social and economic inclusion. Their universality is blurring the traditional North-South divide, planting the seeds for a completely new definition of partnership between developed and developing countries.

Such intense changes require a profound reflection on the role played by bilateral and multilateral development agencies and international cooperation in all its forms, including finance, technology, capacity-building and scientific research and development cooperation.

IFAD, as one of the leading agencies driving change, has launched its new strategic vision for the post-2015 agenda: to invest in rural people for economic, social and cultural impact, leading to a sustainable and inclusive rural transformation. This commitment has translated into a coherent programme of work for the Fund's Tenth Replenishment process, which includes key entry points such as: investing in smallholder agriculture for global food security and nutrition; promoting empowerment for rural livelihoods; promoting the resilience of poor rural women and men; and leveraging the rural-urban nexus for development.

Against this backdrop, the role of evaluation cannot be overstated. Evaluation has become a potent mechanism through which development agencies can better deliver, reach their target populations, understand the activities that can make a real and lasting difference, and most effectively disseminate successes and lessons learned. If the world cannot evaluate the effectiveness of the policies and actions that are put in place to achieve development outcomes, then there is no point in creating and implementing them.

This second edition of the Evaluation Manual contains the core methodology that the Independent Office of Evaluation of IFAD (IOE) uses to conduct its evaluations. The Manual has been developed based on the principles set out in the IFAD Evaluation Policy, building on international good evaluation standards and practice. It incorporates new international evaluative trends and draws from IOE's experience in implementing the first edition. The Manual also takes into account IFAD's new strategic priorities and operating model – which have clear implications for evaluation methods and processes – and adopts more rigorous methodological approaches, for example by promoting better impact assessment techniques and by designing and using theories of change. The Manual also contains the core methodology for evaluations that were not contemplated in the first edition, such as corporate-level evaluations, impact evaluations and evaluation synthesis reports.

The Manual is available in Arabic, English, French and Spanish to facilitate its use in all regions where IFAD has operations. It is also available online on the IFAD Evaluation website (www.ifad.org/evaluation).



OSCAR A. GARCIA

Director

Independent Office of Evaluation of IFAD

Abbreviations and acronyms

ACP	agreement at completion point
ARRI	Annual Report on Results and Impact of IFAD Operations
CLE	corporate-level evaluation
CLP	core learning partnership
CPM	country programme manager
COSOP	country strategic opportunities programme
CSPE	country strategy and programme evaluation
DAC	Development Assistance Committee
ECG	Evaluation Cooperation Group (of multilateral development banks)
ECU	Evaluation Communication Unit
FAO	Food and Agriculture Organization of the United Nations
GPS	good practice standards
IFAD9	Ninth Replenishment of IFAD's Resources
IFAD10	Tenth Replenishment of IFAD's Resources
IFI	international financial institution
IOE	Independent Office of Evaluation of IFAD
KEQ	key evaluation questions
KM	knowledge management
M&E	monitoring and evaluation
MICs	middle-income countries
MTR	mid-term review
NGO	non-governmental organization
OECD	Organisation for Economic Co-operation and Development
PB	partnership-building
PCR	project completion report
PCRV	project completion report validation
PD	policy dialogue
PMD	Programme Management Department
PPA	project performance assessment
PPE	project performance evaluation
PRISMA	President's Report on the Implementation Status of Evaluation Recommendations and Management Actions
RIDE	Report on IFAD's Development Effectiveness
RIMS	Results and Impact Management System
RMF	Result Measurement Framework
SIA	senior independent adviser
TOC	theory of change
UNEG	United Nations Evaluation Group

1. Background

The audience and purpose of the Manual

The Evaluation Manual's primary function is to guide staff and consultants engaged in evaluation work at the Independent Office of Evaluation of IFAD (IOE). It also serves as a reference document for other IFAD staff and development partners (such as project management staff and executing agencies of IFAD-supported operations), especially in recipient countries, on how evaluation of development programmes in the agriculture and rural development sector is conducted in IFAD.

The revision of this Manual was undertaken in recognition of the dynamic environment in which IFAD operates, and in response to the evolution in the approaches and methodologies of international development evaluation. It will help ensure that IFAD's methodological practice remains state of the art. It may be read in conjunction with the Evaluation Process Guidelines.¹

Global development trends and corporate reflection on IFAD's role, business model and theory of change are all issues the Manual seeks to align with and support. It is a Manual on how evaluation is performed at IFAD, not a sourcebook on evaluation in general. However, to enhance its usefulness, references for further reading on selected evaluation topics are included in annex VI.

The Manual is a living document that will be revised, updated and possibly expanded as needs arise and the context demands. Therefore, readers are encouraged to always refer to the web-based version, which will include the latest updates. To ensure its wider use, IOE has translated the Manual into the other IFAD official languages (Arabic, French and

Spanish). The Manual will be implemented starting January 2016 and IOE will organize training sessions for all concerned to ensure its proper implementation.

The context – the challenge of rural poverty reduction

Major changes are taking place in the global economic system and the international development architecture, and consensus is gradually emerging around the future design of a global agenda for post-2015 development. In particular, the establishment of the United Nations Sustainable Development Goals is an important landmark. These goals demonstrate the scale and ambition of this new universal agenda and seek to attain human rights for all and achieve gender equality and the empowerment of all women and girls. They are integrated and indivisible and balance the economic, social and environmental dimensions of sustainable development.

IFAD has been actively engaged in the main processes of developing this agenda in order to make sure that it will contribute to lifting poor rural people out of poverty. No poverty eradication and inclusive growth agenda can succeed without serious attention to rural areas, rural sectors, and smallholder and family agriculture. Indeed, poverty has multiple dimensions that go beyond low levels of income, consumption and material assets; this is why IFAD targets its investments towards rural transformation – a sustainable and comprehensive level of change that is social as well as economic, and which takes different shapes in different country contexts, such as middle-income countries (MICs), small island states, fragile states, low-income countries, etc.

¹ Evaluation Process Guidelines are intended for internal use by IOE staff: these will emphasize the procedures and administrative guidelines for each of the evaluation products undertaken by IOE, whereas the manual is focused on the methodological and technical aspects of evaluation.

Rural women and men – farmers, small entrepreneurs, workers – today face a very dynamic environment with new opportunities, vulnerabilities and risks. Through the right investments, tools and partnerships, poor rural women and men can make a key contribution to sustainable development and rural transformation. Through its targeting policy, IFAD proactively strives to reach poor rural people to pursue an inclusive and sustainable development agenda and provide them with social and economic development opportunities to improve their incomes, nutrition and overall livelihoods. More productive, sustainable and resilient smallholder and family agriculture can and will play many roles in creating a better future – feeding more populated urban areas, providing decent jobs and incomes, delivering environmental services, and fostering social cohesion.

While it is estimated that nearly 70 per cent of the world's population will be living in urban areas by 2050, it is important to recognize that urban and rural areas cannot succeed without each other. Interdependencies between rural and urban areas are increasing – rural growth (especially within food systems) underpins the sustainability of urban areas, while growth in urban sectors provides opportunities to raise rural incomes. And such interdependencies will be further shaped and sharpened by the country context and level of development.

Agriculture today indeed needs to deliver on a complex agenda of contributing to global food security and improved nutritional outcomes, offering expanded employment opportunities, and sustainably managing the natural resource base, while remaining resilient to the effects of climate change. The role of smallholders in this agenda is crucial, and it is therefore critical to ensure that poorer rural households are able to access and benefit from these opportunities and are not further marginalized, on the contrary, are able to better manage the growing risks they face.

A transformation of rural spaces, productive sectors and societies is needed – one that is inclusive, dynamic and sustainable. To promote this requires a new development paradigm that

empowers rural people to play their economic, social and environmental roles to the fullest. The post-2015 development agenda can encourage this through goals, targets, indicators and implementation modalities that give explicit attention to rural women and men and their role in building a better future.

The latest estimates by the Food and Agriculture Organization of the United Nations (FAO) put the number of chronically undernourished people at about 795 million for the period 2014-2016, down more than 216 million since 1990-1992. However, further declines will be highly dependent on global food prices, since the world's poor spend over half their income on food. Rates of extreme poverty have been halved between 1990 and 2010 – with declines recorded in all regions; however, there are still 1.2 billion people living on less than US\$1.25 a day. For IFAD, there are two critical associated points. The first is the fact that most of the world's extremely poor people no longer live in the world's poorest countries: three quarters of them live in countries that have reached middle-income status. The second is that poverty rates in rural areas are substantially higher than in urban areas, and 70 per cent of those living on less than US\$1.25 a day – 840 million – live in rural areas, and most are directly or indirectly dependent on agriculture.

Aid and agriculture in the global development agenda

Flows of official development assistance from international development agencies comprise a declining and relatively minor share of agricultural investment. Other forms of development initiatives are increasingly gaining importance, such as South-South and Triangular Cooperation, inter alia, with the aim of sharing experiences and knowledge of what works and what does not, and facilitating partnerships, fostering innovations, and promoting the scaling up of promising ideas. However, official development assistance can be important for some countries, enabling them to plug major funding gaps. For many more, including many MICs, official development assistance can serve to catalyse private

investment in the agricultural and rural sector and help governments strengthen their policies and add value to the quality of public investment in the sector. Increasingly, this is the role that IFAD is being called on to play, and it will become ever more important in the future.

The role of IFAD

IFAD's mission is indeed more relevant than ever. Its mandate is to contribute towards rural transformation by supporting inclusive and sustainable agriculture and rural development activities in developing countries. For the Ninth Replenishment of IFAD's Resources (IFAD9) (2013-2015), IFAD established a target of reaching 90 million people with its operations, 80 million of whom to be moved out of poverty by 2015. A similar commitment was made for the Tenth Replenishment of IFAD's Resources (IFAD10) (2016-2018). By the end of 2018, IFAD made a commitment to move 80 million people out of poverty, and reach between 110-130 million with services from IFAD-supported projects.

IFAD functions as a specialized agency of the United Nations system and an international financial institution (IFI). The Fund has a clear vision of its role: in different contexts, to facilitate both public and private investment, support national and global policy processes, generate and share knowledge, and develop partnerships, all in pursuit of a goal of sustainable, equitable and inclusive rural transformation, in which marginalized groups play a full and active role and from which they are able to benefit.

The main instruments for delivery are loan-funded projects and programmes, which IFAD helps governments, smallholders and other stakeholders to develop and implement. The Fund also has a small grant-financing window, and a new grant policy has just been developed. This will help ensure a more strategic use of resources, including in scaling up its innovative approaches and initiatives whenever feasible.

Indeed, IFAD views scaling up its results and impact as mission-critical, and in this respect

also sees knowledge management, partnerships and policy dialogue as essential for achieving its goals. Through its partnerships and cofinancing arrangements, IFAD seeks to leverage its own resources, experience and knowledge, all of which are essential for scaled-up impact.

IFAD's country programmes are made up of the projects IFAD supports, the related policy processes in which it engages, and the policy changes at national or state level to which it contributes; these facets are supported by the knowledge it gathers and shares and the partnerships it develops to carry them out. It is the country programmes – this mix of projects and policy engagement – that helps deliver development impact at scale, assessed in terms of the numbers of rural people moved out of poverty, but also in terms of the quality of development results, paying attention to inequalities, environmental aspects, climate change, etc. and the various challenges faced by the rural poor.

One important dimension for IFAD to effectively and efficiently deliver on its mandate is to promote and nurture strategic partnership with governments, the private sector, civil society and the non-governmental organization (NGO) community, think tanks, and other development organizations such as bilateral aid agencies, international financial institutions and the United Nations specialized agencies, programmes and funds. With regard to the latter, the Rome-based agencies (FAO and the World Food Programme) offer special opportunities for collective action, not only because of their physical proximity to one another, but particularly due to the complimentary nature of their mandates in the food and agriculture domain.

IFAD's theory of change²

IFAD's unique role in contributing to inclusive and sustainable rural transformation is built upon a theory of change (TOC) that was explicitly articulated in the context of IFAD10 and included in IFAD's Result Measurement Framework (RMF).

The TOC is composed of five levels. Level 1 of the RMF – Global poverty, food security and agricultural investment outcomes – embodies

² Based on a recommendation included in the Corporate-level evaluation (CLE) on IFAD Replenishments, the Fund developed its first explicit theory of change in the context of IFAD10; see the final replenishment report at <https://webapps.ifad.org/members/gc/38/docs/GC-38-L-4-Rev-1.pdf>.

the broader development goals to which IFAD contributes, tracked by measures of extreme poverty, food insecurity and investment in agriculture. Level 2 – Country-level development outcomes and impact – assumes that success in achieving global poverty and food security targets is predicated on development projects – such as those supported by IFAD and is measured by the number of people moving out of poverty. Level 3 – Country-level development outputs delivered by IFAD-supported projects – underpin the achievement of impacts and outcomes reflected by the number of people lifted out of poverty (level 2). The supply of programme outputs enables smallholders to overcome key constraints, giving them access to a range of productive resources and inputs, knowledge, infrastructure, market, financial and business services, and responsive institutions. The number of people using the services of IFAD-supported projects is considered the most important indicator at this level. RMF level 4 – Operational effectiveness of IFAD-supported country programmes and projects – assesses the performance of IFAD-managed processes aimed at positively influencing the design, delivery and achievement of programme outputs. The assumption is that enhancing the quality and efficiency of IFAD’s business model and associated processes will lead to improved programme outputs. Achievement of agreed programme results is supported by: better project design responding to the needs and priorities of poor rural people; an effective portfolio management process identifying underperformance in time for a course correction; offering proactive assistance to entities for project implementation; and deepening knowledge on how policies for smallholder agriculture and rural development work in practice. RMF level 5 – IFAD’s institutional effectiveness and efficiency – reflects a commitment to ensuring institutional effectiveness and efficiency in the way the organization mobilizes, allocates and utilizes its financial and human resources. The assumption is that strong performance in this area will focus

resources on what adds most development value, with overall impact and quality delivered for less, in terms of expenditures.

Looking forward, IFAD will continue to focus on the areas it knows and on which it has built its reputation, while at the same time working in partnership with others to address development challenges that go beyond its area of specialization, or to leverage its experience and knowledge for greater development impact. An important element of this vision is better integration between viable, profitable and sustainable small-scale agriculture and the range of non-farm sectors that contribute more or less directly to agricultural value chains. This agenda reinforces the role of evaluation, to assess the poverty impact and the role of partners, to understand the pathways to change, and to examine and assess what works, for whom, where, when and why.

Evaluation in IFAD

Evaluation plays an important role in the way IFAD’s business model, operational focus, corporate processes, and accountability systems evolve. They are refined, adjusted and sharpened through feedback from evaluation to ensure that the Fund is in the best possible position to fulfil its mandate and meet its corporate goals. When new issues come to the fore, and the world is developing new goals, targets and indicators (for example, on inequality and equity considerations), IFAD’s evaluation function needs to evolve alongside these changes. New demands are put on IFAD’s ability to understand rural transformation, to report on results in new areas, on its capacity to measure new dimensions of poverty, and on its data, evaluation tools, approaches and methodologies. To meet these demands, IFAD’s evaluation architecture has been refined and is comprehensive.

This architecture has been developed to meet the twin purposes of accountability and learning. Annex I illustrates the overall architecture of IFAD’s independent evaluation function, and how its products³ mutually reinforce each other

3 Including the Annual Report on Results and Impact of IFAD Operations (ARRI), CLEs, country strategy and programme evaluations, evaluation synthesis reports, impact evaluations, project completion report validations and project performance evaluations.

and contribute to strengthening organizational accountability and learning. Each product is essential and has a specific place in the architecture, which also comprises IFAD's self-evaluation system.

The Fund also has a comprehensive self-evaluation function,⁴ maintained by IFAD Management. For example, for each project, IFAD conducts periodic supervision missions and a mid-term review (MTR), and prepares a project completion report.⁵ Rigorous impact evaluations are also undertaken selectively in a number of operations. There are other instruments in the self-evaluation system, including the Results and Impact Management System (RIMS), client surveys, quarterly corporate performance reviews, and annual portfolio reviews. These and other self-evaluation products serve as essential inputs for IFAD's independent evaluation activities.

The Evaluation Policy

The Fund's independent evaluation processes are guided by the principles outlined in the Evaluation Policy, first approved by the Executive Board in April 2003 and revised in 2011. This policy is built on the four principles of accountability, learning, independence and partnership. These principles govern the undertaking of each evaluation by IOE, and the methods and processes that IOE adopts must therefore support them. This requires careful management, as the promotion of one principle may have implications for the others. For example, IOE is committed to ensuring participation throughout the evaluation process as an intrinsic aspect of partnership and to promoting inclusiveness and ownership in evaluation findings and recommendations; yet, in observance of the principle of independence, participation and ownership should not be allowed to influence the evaluative process.

Impartiality and the independence of IFAD's evaluation function are of special importance. To quote the policy: "The IFAD evaluation function shall operate in line with internationally accepted principles for the evaluation of development

assistance. Foremost among these is the principle that the evaluation process should be impartial and independent from both the policymaking process and the delivery and management of development assistance." However, this does not mean isolation; to ensure the equally important dimensions of relevance and usefulness of the evaluation products and processes, dialogue with Management is required.

The independence principle is reflected in a number of provisions of the Evaluation Policy, which stipulate that: (i) the Director of IOE reports to the Executive Board rather than to the President of IFAD; (ii) the IOE work programme and budget are prepared independently of IFAD Management and presented directly to the Board and Governing Council for approval; (iii) the President has delegated his authority to make all human resource decisions related to IOE to its Director; and (iv) the Director of IOE is authorized to issue evaluation reports to IFAD Management, the Fund's governing bodies and the public at large without seeking the clearance of any official outside IOE.

Operationalizing the Policy

IOE undertakes various types of evaluations, including country strategy and programme, corporate-level, project and more recently impact evaluations. It also prepares syntheses of evaluation findings as a key knowledge product. Efforts are made to ensure that appropriate feedback loops are established and learning opportunities developed so that findings and conclusions from evaluations feed into ongoing work, and can inform and renew corporate reflection.

To ensure maximum usefulness of all products, three core issues are considered: (i) evaluations are completed in a timely manner to ensure their usefulness; (ii) the timing for conducting a specific evaluation is carefully orchestrated, for example, by prioritizing a country strategy and programme evaluation (CSPE) so that the findings feed into the preparation of the new country

4 Self-evaluation is an evaluation by those who are entrusted with the design and delivery of a development intervention.

5 The preparation of project completion reports is normally the responsibility of the recipient governments. However, IFAD has developed guidelines for such reports for governments to follow. Also, on a case-by-case basis, IFAD will support governments, for example, by fielding a consultant to support local authorities in the preparation of these reports.

strategic opportunities programme (COSOP) in that country; and (iii) the selection of topics for evaluation is transparent and is made according to the evaluation's usefulness. Therefore, IOE has developed a selectivity framework to ensure full transparency on the selection of topics and undertakes a consultation process with Management on a range of possible themes and issues. The selectivity framework and its application are presented each year as a part of IOE's annual Work Programme and Budget to the Board and the Governing Council (see annex V).

The Annual Report on Results and Impact of IFAD Operations (ARRI) deserves special reference as it is grounded in the Evaluation Policy's statement that "the accountability of an international financing institution as a whole is facilitated if the results of individual evaluations can be aggregated and consolidated at the organizational level. This allows a better analysis of the effectiveness of a given development organization and of the cross-cutting issues and challenges that impinge on its overall performance."

The ARRI is IOE's flagship annual report, first produced in 2003, and addresses precisely this as the main purpose of the ARRI is to provide an aggregated insight into the performance of IFAD operations and identify systemic issues and lessons, as well as generate recommendations to enhance IFAD's development effectiveness. The ARRI draws on the different evaluations (e.g. project evaluations and validations, CSPEs, etc.) done by IOE based on a common methodology, as enshrined in its evaluation manual. In fact, IOE is one of the few evaluation offices across multilateral or bilateral development organizations that produce such an annual report, which is a reflection of the Fund's commitment to promoting accountability, learning and transparency in reporting on results.

The field of evaluation is constantly evolving and IOE aims to be both efficient and effective in terms of its evaluation processes, and to be "ahead of the curve" in terms of methodological development and development of new evaluation products. In

this respect, the active engagement with peers is essential and CSPEs, for example, are carried out taking into account the good practice standards (GPS) developed by the Evaluation Cooperation Group (ECG). Approaches to project evaluations too have been informed by consultation with peers and benefitted in particular from the 2010 Peer Review of IFAD's Office of Evaluation and Evaluation Function by the ECG, which recommended that IOE transform its approach to project evaluations. As a result, IOE discontinued the way project evaluations were done in the past,⁶ and instead introduced project completion report validations (PCRVs) and project performance evaluations (PPEs) – (see chapters 4 and 5).

The Evaluation Manual

This second edition of the Manual seeks to renew, update and consolidate current guidelines in order to achieve high quality standards in IFAD evaluations in support of IFAD's mandate and within the framework of the agreed Evaluation Policy. It represents an evolution of the first edition of the Evaluation Manual (2009) in a number of ways, including guidance for undertaking all types of evaluations carried out by IOE.

Table 1 shows some of the salient differences between the first and the second editions of the Evaluation Manual. In addition to the changes presented in table 1, inter alia, the second edition of the Manual includes a set of enhanced questions to assess performance across each evaluation criteria, introduction of dedicated criteria for systematically assessing adaptation to climate change and environmental and natural resource management in line with IFAD10 priorities, wider discussion on key methodological fundamentals for independent evaluations (e.g. use of theory of change approach, new techniques for data collection and analysis, and participation of beneficiaries and key stakeholders in evaluation processes), the transformation of project performance assessments into project performance evaluations and country

⁶ In the past, project evaluations cost around US\$100,000 each (excluding staff costs) and took nearly one year to complete, whereas PPEs cost around US\$30,000-US\$35,000 each. PCRVs are normally done internally by IOE staff and thus only incur staff costs. PCRVs and PPEs are both done in less time than the old generation project evaluations.

TABLE 1 Main modifications between the first and second edition of the Evaluation Manual*

Features	First edition of the Evaluation Manual (2009)	Second edition of the Evaluation Manual
Evaluation products	<ul style="list-style-type: none"> • Project evaluations (completion evaluations and interim evaluations) • Country programme evaluations (CPEs) • Corporate-level evaluations • Annual Report on Results and Impact of IFAD Operations 	<ul style="list-style-type: none"> • Project completion report validations • Project performance evaluations • Impact evaluations • Country strategy and programme evaluations (CSPEs) • Corporate-level evaluations • Annual Report on Results and Impact of IFAD Operations
Chapters	<ol style="list-style-type: none"> 1. Background 2. Methodological fundamentals 3. The project evaluation methodology 4. The country programme evaluation methodology 	<ol style="list-style-type: none"> 1. Background 2. Methodological fundamentals 3. Evaluation criteria, key questions and ratings 4. Project completion report validation 5. Project performance evaluation 6. Country strategy and programme evaluation 7. Corporate-level evaluation 8. Impact evaluation 9. Evaluation synthesis report 10. Communication, dissemination and learning
Evaluation criteria	<p>Project evaluations and portfolio performance in CPEs</p> <ul style="list-style-type: none"> • Project performance (relevance, effectiveness and efficiency) • Rural poverty impact (household income and assets, human and social capital and empowerment, food security and agricultural productivity, natural resources and the environment, institutions and policies) • Other performance criteria (sustainability, and promotion of pro-poor innovation, replication and scaling up) • Overall project achievement (relevance, effectiveness, efficiency, rural poverty impact, sustainability, innovation, replication and scaling up) • Partner performance (IFAD, government, cooperating institution, non-governmental organization/community-based organization) <p>Country programme evaluations</p> <ul style="list-style-type: none"> • Non-lending activities (policy dialogue partnership-building and knowledge management), including grants • COSOP performance (relevance, effectiveness) • Government and IFAD partnership (portfolio performance, non-lending activities, and COSOP performance) 	<p>Project evaluations and portfolio performance in CSPEs</p> <ul style="list-style-type: none"> • Rural poverty impact (household income and assets, human and social capital and empowerment, food security and agricultural productivity, and institutions and policies) • Project performance (relevance, effectiveness, efficiency and sustainability of benefits) • Other performance criteria (gender equality and women's empowerment, innovation and scaling up, environment and natural resources management, and adaptation to climate change) • Overall project achievement (rural poverty impact, relevance, effectiveness, efficiency, sustainability of benefits, gender equality and women's empowerment, innovation and scaling up, environment and natural resources management, and adaptation to climate change) • Partner performance (IFAD and government) <p>Country strategy and programme evaluations</p> <ul style="list-style-type: none"> • Non-lending activities <ul style="list-style-type: none"> – policy dialogue – partnership-building – knowledge management – South-South and Triangular Cooperation, grants and reimbursable technical assistance • COSOP performance (relevance, effectiveness) • Government and IFAD partnership (portfolio performance, non-lending activities, and COSOP performance)
Rating	Six-point scale	Six-point scale

* Climate change and gender equality and women's empowerment were included following the 2010 peer review of IOE by the ECG.

programme evaluations into country strategy and programme evaluations, and more thorough attention to communication and outreach of results and lessons learned. There are also numerous new chapters (e.g. on corporate-level evaluations [CLEs], impact evaluations, evaluation synthesis reports and communication and dissemination) and features that contribute to the completeness of the document.

The Manual has been prepared through a process of engagement with multiple internal and external feedback opportunities from various stakeholders, including peer institutions (African Development Bank, Asian Development Bank, FAO, Institute of Development Studies (University of Sussex), Swiss Agency for Development and Cooperation and the World Bank).⁷ The Manual was also reviewed by a high-level panel of experts.⁸ Their main role was to provide guidance and inputs to IOE and to confirm to IFAD governing bodies that the Manual is in line with good practice standards in international development evaluation. Consultations were held with IFAD Management and the Evaluation Committee of IFAD's Executive Board prior to the finalization of the document.

The Manual comprises guidance on all evaluation products: PCRVs, PPEs, CSPEs, CLEs, impact evaluations and evaluation synthesis reports. It also includes essential guidance on evaluation fundamentals and criteria that are applicable to all evaluations, and it contains an important section on learning and dissemination, recognizing that reports are of little value if the knowledge therein is not appropriately used by as many people as possible.

While the Manual seeks to instill a degree of consistency across IOE evaluations, it is important to stress that it should be used intelligently and

flexibly. The field of evaluation is continually evolving, and there is no one method or approach that is the right one in all circumstances; neither is it possible to assess the full range of evaluation criteria in every evaluation. Therefore, users of the Manual are encouraged to apply critical thinking; demonstrate creativity, innovation and flexibility; and when carrying out an evaluation, consider the costs and benefits of the approach proposed and keep updated on current international thinking and good practice on that particular theme or type of evaluation.

The development and implementation of this Manual should facilitate IOE's participation in joint evaluations with other development organizations and with recipient governments. Joint evaluations have the potential to lower transaction costs, although this is not necessarily always the case. However, benefits also include wider exchanges of knowledge, experience and evaluation capacity development as they expand the scope to include all major development partners in the programmes being evaluated; these elements deserve consideration when determining whether an evaluation is best carried out as a joint initiative.

This Manual is organized as follows: chapter 2 provides an overview of methodological fundamentals and chapter 3 contains guidance on evaluation criteria. Chapters 4-9 include details to guide the respective evaluation products and chapter 10 is about communication, dissemination and learning. A selection of support material is included in the annexes to facilitate implementing the Manual and conducting the different types of evaluation.

⁷ Specifically, the Swiss Agency for Development and Cooperation commented on chapters 2 and 3 (methodological fundamentals and evaluation criteria; key questions and ratings); the Asian Development Bank commented on chapter 4 (project completion report validations); FAO and the African Development Bank on chapter 6 (country strategy and programme evaluations); the Institute of Development Studies on chapter 8 (impact evaluations), and the World Bank on chapter 9 (evaluation synthesis reports).

⁸ Osvaldo Feinstein, Susanne Frueh, Ted Kliet, Uma Lele and Kevin Watkins.



2

2. Methodological fundamentals

This chapter presents methodological fundamentals that are the basis for all IOE evaluations. Specific issues and considerations for different types of evaluations are covered in individual chapters. Awareness of these methodological fundamentals by evaluators helps to set standards and ensure consistency in methodology and in reporting formats across evaluators and evaluations; it enhances the robustness and rigour of IOE evaluation products and facilitates comparison of results and aggregation.

IFAD's Evaluation Policy constitutes the overall framework for the independent evaluations conducted by IOE and contains the evaluation principles and operational policies. The key principles are independence, accountability, partnership and learning. According to the Policy, the Manual is a key element for promoting learning and accountability, and "sets out the methodology and the process that the division will apply in

the assessment of the results of IFAD policies, strategies and operations."⁹

IFAD's Evaluation Policy also states that the IFAD evaluation function shall operate in line with internationally accepted principles for the evaluation of development assistance. IFAD is a specialized agency of the United Nations with a governance structure and operating model similar to that of an IFI. As such, IOE subscribes to the overarching norms and standards adopted by the United Nations Evaluation Group (UNEG),¹⁰ ECG¹¹ and Organisation for Economic Co-operation and Development (OECD).¹² Consistency and compliance of IOE's work and approaches with these broader principles was at the core of the Peer Review of IFAD's Office of Evaluation and Evaluation Function,¹³ conducted in 2009/2010.

These core evaluation principles, which also govern IFAD's evaluative work, are presented in box 1.

BOX 1 Principles for Evaluation of Development Assistance

- Aid agencies should have an evaluation policy with clearly established guidelines, methods and a clear definition of its role and responsibilities and its place in institutional aid structure.
- The evaluation process should be impartial and independent from the process concerned with policy-making, and the delivery and management of development assistance.
- The evaluation process must be as open as possible with the results made widely available.
- For evaluations to be useful, they must be used.
- Feedback to both policy-makers and operational staff is essential.
- Partnerships with recipients and donor co-operation in aid evaluation are both essential; they are an important aspect of recipient institution-building and of aid co-ordination and may reduce administrative burdens on recipients.
- Aid evaluation and its requirements must be an integral part of aid planning from the start. Clear identification of the objectives which an aid activity is to achieve is an essential prerequisite for objective evaluation.

Source: Development Assistance Committee, Paris, 1991.

⁹ <http://www.ifad.org/gbdocs/eb/102/e/EB-2011-102-R-7-Rev-1.pdf>, para. 26 (v).

¹⁰ <http://www.uneval.org/document/foundation-documents>.

¹¹ <https://www.ecgnet.org/document/ecg-big-book-good-practice-standards>.

¹² <http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf>.

¹³ <http://www.ifad.org/gbdocs/eb/ec/e/62/e/EC-2010-62-W-P-2.pdf>.

The main purpose of evaluation, according to IFAD's Evaluation Policy, is to promote accountability and learning for better development effectiveness. The range of IOE evaluation products ensures that both these fundamental dimensions of evaluation are adequately covered. Some IOE products (e.g. PCRVs) are strongly oriented towards accountability, whereas others have a greater focus on learning (e.g. evaluation synthesis reports). But the majority of IOE products (e.g. CLEs, CSPEs, impact evaluations and PPEs) cover both accountability and learning dimensions of independent evaluation.

Approach, design and methods

IOE evaluations are designed to enable judgement about what has been achieved while also supporting forward-looking operational improvement and strategic guidance at the corporate level. Irrespective of the type of evaluation, the primary purpose of the evaluation drives the design of the evaluation and choice of various evaluation methods. As is pertinently stated in the Evaluation Policy: "Accountability through evaluation analysis requires a rigorous methodology for the assessment of developmental results and impacts and the performance of the partners concerned."

In line with the overarching purpose of evaluation in international development cooperation, IOE evaluations follow an objectives-based¹⁴ approach. Such evaluations focus on understanding the pathways to change, the measuring of achievements along the results chain, and generating lessons and recommendations for better performance. The starting point is the broader strategic intent, objectives and design of a policy, strategy or operation, as approved by the Fund's Executive Board, and will examine both the achievement of objectives but also unintended results.

IOE also closely follows the international discourse related to theory-based evaluation approaches, as these have advantages over the traditional objectives-based approach. As stated in a World Bank document:¹⁵ "Theory-based evaluation

has similarities to the LogFrame approach but allows a much more in-depth understanding of the workings of a program or activity – the "program theory" or "program logic." Recommended reading on theory-based approaches can be found in annex VI.

Theory-based approaches will be used in the design of IOE evaluations as and when appropriate. Theory-based approaches will be particularly appropriate, but not limited, to the undertaking of CLEs, especially in those cases when IOE is required to evaluate a policy, strategy or business process for which a specific policy document is not available as a main reference for evaluation. It is also highly relevant for impact evaluations, CSPEs, and selected PPEs.

Different methodologies and methods have comparative advantages in addressing particular concerns and needs; and in the end, the question is not whether an evaluation design is ideal, but whether it is adequate given the available budget and capacities and hence "fit for purpose". Given constraints and challenges faced, is it able to provide the evidence that allows passing judgement on what was achieved and exploring alternatives? This is also a fundamental principle of the Evaluation Policy: "IOE shall require the evaluators to ensure that the methods and processes in the manual are adequately customized – as needed – to reflect the specific circumstances of the policy, strategy or operation being evaluated." Indeed, a given set of evaluation questions could be answered by a range of different evaluation designs, and different options should be considered, keeping in mind the purpose, context, available resources and budget, and known or anticipated availability of data.

Using a mix of methods – combining the breadth of quantitative methods with the depth of qualitative – and "triangulating" information from different approaches – can be useful to assess different facets of complex outcomes or impacts. This will yield greater validity than using one method alone and will usually be the preferred approach in IOE. It is also in line with UNEG good practice: "A 'mixed method' approach utilizing quantitative, qualitative, participatory and blended (e.g. quantifying qualitative

¹⁴ <https://www.ecgnet.org/document/good-practice-standards-evaluation-public-sector-operations>.

¹⁵ World Bank: *Monitoring and Evaluation, Some Tools, Methods and Approaches*, 2004.

data) approaches is now widely accepted as advisable to address the types of interventions that are now predominant in international development.”¹⁶

Developing the approach paper is the first step in any evaluation, and a critical one, in which the different methodological options available are being considered. The approach paper covers, consistent with IFAD’s Evaluation Policy, the following elements: background and rationale, objectives, key questions, methodology and process, time frames, human resources, and communication and dissemination activities. An evaluation framework (i.e. in the form of a matrix) should be developed as the core of the approach paper for each evaluation; the purpose of the evaluation framework is to ensure that evaluators clearly establish the link between the key questions to be addressed, and the data collection instruments and methods, so that a clear evidence trail is established and evidence can be consolidated to support findings. In this respect, evaluators should also observe the implementation modalities of the interventions being evaluated because they may offer opportunities for innovative evaluation designs.

Furthermore, an evaluability assessment¹⁷ should customarily be carried out before the evaluation framework is developed. Evaluability is assessed against the quality and availability of: (i) data deriving from IFAD programme documents available in house; and (ii) data and information available at the country/programme levels. Special attention will be devoted to the availability of gender-related data and information in all evaluability assessments. The results of the evaluability assessment will allow evaluators to choose the different methods and instruments for collecting data and information to answer the key questions specific in the evaluation framework.

Attribution and contribution

IOE conducts evaluations with the dual purpose of learning and accountability, but all evaluations also have an element of exploration to detect to what extent a particular outcome or impact can be attributed to a particular intervention.

This is no easy assessment. First, IFAD-supported activities involve many partners. Second, they are exposed to external factors that influence results; in particular, donor countries’ policies, beneficiary countries’ domestic policies, other development programmes, socio-economic fluctuations, structural changes and climatic phenomena can affect the results. And third, baselines that may provide strong support for dealing with attribution issues are often not available or not of the required quality. IFAD-supported projects rely on national statistical systems which are generally weak. The database in IFAD’s RIMS is continually improving, but lack of solid data often remains a constraint to good evaluation.

Therefore, making a credible assessment attributing the results achieved on the ground to a particular project or programme, let alone to IFAD’s own performance, is challenging. While this recognition has led to a widespread use of “contribution analysis” rather than “attribution” in evaluations, addressing the issue clearly is critical.

Where data allow, two scenarios can be used: the “before and after” and the “with – without”. The “before and after” technique is often hindered by the lack of baseline data and inadequate monitoring and evaluation (M&E) systems, although specific techniques (e.g. memory recall, wealth ranking, community mapping) can shed light on the situation before the project/programme/policy was introduced, thus facilitating the “before and after” assessment.

The “with – without” is also referred to as the counterfactual and provides an assessment of impact attribution on an even more comprehensive basis, but is not applicable in all cases. It calls for the definition of a plausible counterfactual, which is the situation or scenario that would hypothetically prevail were there is no development intervention. While actual measurement of a counterfactual poses significant challenge and most IFAD projects, even those with a well-functioning M&E system, do not incorporate counterfactual observations (e.g. survey data of households that have not been assisted

¹⁶ <http://www.unevaluation.org/document/detail/1433>.

¹⁷ For guidance on conducting evaluability assessments see: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf.

by a project), evaluators should always apply counterfactual thinking, i.e. asking the question: what would have happened without the intervention?

In cases where data are available, the use of control groups may help to acquire an understanding of the impact of IFAD-supported operations on livelihoods (e.g. in terms of income, nutritional status, access to resources) by assessing the results of interventions on target groups and by comparing them with the situation of populations outside the target group in similarly situated regions. In choosing the beneficiaries and a control group to represent the project population and ascertain project impacts, care should be taken to avoid selection biases. Moreover, evaluators should also reflect at the time of design if there is likelihood of any systematic errors in measuring impact, and how they will be addressed to ensure greater accuracy in the assessment (see also chapter 8 on impact evaluations).

In the absence of a control group, alternative methods include general elimination method¹⁸ – asking what else may have led to changes – or scenario analysis to estimate the “without project” situation. While scenario analysis is usually considered a predictive technique, it may also be applied to simulate a counterfactual in retrospective analysis.¹⁹ Also, benchmarking can be used to make a comparative analysis, and the “causal contribution analysis” approach suggested by Mayne²⁰ is an approach that has been used in a number of IFAD evaluations.

Notwithstanding the increasing attention to impact, at this stage no consensus has emerged regarding agreed methodologies for rigorously attributing the impact of development interventions to specific factors or causes. On the one hand, some researchers call for a rigorous assessment

of causality through quantitative measures of impact and advocate the use of randomized control trials and other experimental and non-experimental approaches (e.g. using propensity score matching methods)²¹ as the “gold standard” of impact evaluation. On the other hand, these approaches have severe limitations in complex and volatile development environments and are skills-intensive and costly, as well as requiring a vast amount of data. The issue of attribution and related methodological fundamentals will be further analysed in chapter 8 on impact evaluations.

Techniques and instruments for data collection

Data collection tools will vary according to the type of evaluation, availability of data, local context, resources and time available, and other variables. Using a mix of different methods enhances the robustness and hence credibility of an evaluation. Evaluators are encouraged to carefully consider the relevant options to decide on the most cost-effective technique that will provide the most robust evidence for the evaluation. In order to enhance capacity and widen the experience with different techniques in IOE, testing new techniques alongside well-known methods is also encouraged (see annex VI for suggested reading on data collection techniques). Approach papers should clarify the rationale for using innovative/experimental methods in the design of a particular evaluation so there is clarity among key stakeholders at the outset of the process.

Specific techniques are required for both data collection and analysis. IFAD’s practical guide for monitoring and evaluation at the project level contains a list of 34 different methods, many of

18 Introduced first by Michael John Scriven, this requires the listing of plausible alternative causes.

19 See Asian Development Bank. 2008. *Strengthening the Quality of Project Economic Analysis in ADB Operations*, Manila: ADB, pp. 7-8.

20 Mayne, J. 1999. *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*, Mimeo, Office of the Auditor General of Canada, Canada, and <http://www.innonet.org/resources/node/437>.

21 Propensity score matching methods are often used to control for bias when randomization is not possible. These methods were developed to ensure comparability between the treatment and the comparison group in terms of propensity to participate in the development programme. The first step involves estimating the likelihood (the propensity score) that a person/household would have received the treatment or intervention given certain characteristics. After estimating the propensity scores, the scores are used to group observations that are close to each other. Comparisons of development results can then be applied to different groups of observations which have the same propensity to participate, hence ensuring comparability (see Ravallion M. 1999. “*The Mystery of the Vanishing Benefits: Ms. Speedy Analyst’s Introduction to Evaluation*”. Working Paper No. 2153. Washington, D.C., World Bank).

TABLE 2 The most frequently used data collection methods

	Typical methods	Short description	Remarks
Qualitative – semi-structured	Direct observations	Observations of sites, practices, living conditions, physical constructions (e.g. grain warehouses) according to a pre-agreed checklist (can be combined with rating).	Necessary in most evaluations; can be a source of unexpected findings. Generalization of findings can be an issue.
	Key informant individual interviews with semi-open questionnaires or checklists	Individual interviews on a number of selected topics according to a pre-agreed checklist. The majority of questions are open-ended and meant to stimulate discussion rather than elicit one-word or one-sentence responses.	Useful for discussing sensitive issues that would not normally be discussed in public. Generalization of findings can be an issue. Requires careful preparation of instruments.
	Focus groups	Interaction of a relatively small group of people (normally 6-12) on a limited set of topics, facilitated by a moderator. Beneficiaries agree on a number of preferences, conclusions, beliefs, attitudes, etc.	Valuable for understanding interactions and areas of agreement and disagreement. Generalization of findings can be an issue.
	Memory recall	Entails interviews with beneficiaries and other stakeholders, individually or in groups, who reconstruct their situation before the project.	Necessary generally in all evaluations, especially when baseline surveys are not available, but has limitations especially with long recall periods.
	Participatory techniques: wealth ranking, problem ranking, community mapping, historical transects	Participants are requested to come up with their own criteria and indicators to assess a situation, a process or a distribution of resources and how it has changed over time.	Indicators and parameters are elicited from people rather than pre-selected by researchers. Generalization of findings can be an issue.
Quantitative – structured	Case studies	In-depth assessment of a very limited number of observations (e.g. some microfinance organizations, community development projects or farms).	The criteria for the selection of cases matters. Options include selecting best cases, worst cases or a mix of good-, medium- and low-performing cases.
	Large quantitative surveys	These include the collection of primary data from large numbers of beneficiary households and control and comparison groups. Among other issues, such data provide the basis for rigorous assessment – using statistical methods – of the changes in the lives of the rural poor.	Normally done in experimental or non-experimental (project) impact evaluations.
	Mini-surveys (typically samples of 100-200 respondents, including project and control observations)	A sample of the programme population (and possibly of a control group) is extracted. Interviews are conducted by enumerators on the basis of a pre-written and pre-coded questionnaire. Entries are recorded on electronic support media and analysed using computer software on the basis of standard descriptive, inferential and econometric techniques.	The sampling procedure should try to capture the “true averages” in the population. This technique is feasible in the context of a project or country strategy and programme evaluation. Trained specialists are required for survey design planning and data analysis.

Note: This table does not provide an exhaustive repertoire of available methods. Evaluation teams are invited to examine the specialized literature.

which are highly relevant to IOE's evaluations. Table 2 highlights some frequently used techniques.²²

Data can be primary or secondary, qualitative or quantitative, and the evaluability assessment carried out as part of the preparatory process will normally have examined the availability and quality of different types of data. Relevant data for evaluation may be drawn from existing reports compiled by project authorities or IFAD operations staff and from a variety of other sources, including government statistical or administrative offices, national censuses, and world development indicators from the World Bank and United Nations Development Programme. Therefore, during the evaluability assessment, it is important to assess the availability and quality of secondary data. This enables IOE to target efforts towards the collection of additional data. In particular, it is important to ascertain whether baseline data are available and, if so, to determine their quality. Where baseline data do not exist and surveys have not been undertaken or are not of the required quality, the approach paper should identify how data collection ought to proceed in order to secure a plausible proxy for the assessment of initial conditions. In this respect, Internet searches may be useful to identify studies and databases that are relevant for establishing the "initial conditions".

To establish baselines, evaluators may conduct in-depth interviews with project beneficiaries and have them reconstruct – using memory recall, structured interviews and/or focus groups discussions – the logical chain of behavioural, productive or organizational changes generated or supported by the project, and the underlying assumptions. Evaluators should exercise caution and triangulate the information secured from diverse sources. This is done before deciding on a set of variables deemed to represent initial conditions and those resulting from project interventions; this is

particularly important in the case of income and cost indicators, which may be subject to measurement errors in recall methods.²³

Case studies are often used in IOE evaluation and present a number of methodological challenges²⁴ – in particular in terms of validity and generalization. In terms of internal validity, the concern is how to ensure quality, reliability and robustness of methods and design. In terms of generalizations, the concern is on external validity, i.e. the extent to which it is possible to generalize, and in which circumstances. And lastly, there are issues relating to aggregation and synthesizing for learning purposes. There is an ongoing international debate on this²⁵ and different ways to analyse case study material can be considered, including Qualitative Comparative Analysis, which emphasizes configurations within cases and uses Boolean algebra to support stronger conclusions from a modest number of cases. Systems modelling that relies on qualitative and quantitative methods and allows existing knowledge to be incorporated into models and then "run" under different conditions is also an option.

In general there are two main ways to make it easier to draw conclusions and put forward explanations on the basis of relatively few cases: (i) looking at internal relationships within the case as well as comparisons across variables; and (ii) examining the role of theory, whether to test programmes against theory or to develop theories of change during an evaluation.

When primary data collection is necessary, a combination of (mixed) methods should normally be used to ensure data accuracy and facilitate its interpretation. Concerning quantitative data about agricultural production patterns, incomes and livelihoods can be secured through surveys and possibly national statistics, while insights about the performance of development partners may be dealt

22 See annex D in A Guide for Project M&E (2002) (www.ifad.org/evaluation/guide). For more information on these techniques see: Yin, R. K. 2013. *Case Study Research: Design and Methods* (Fifth Edition); Yin, R. K. 2011. *Applications of Case Study Research* (Third Edition); Hancock D. and B. Algozzine. 2011. *Doing Case Study Research: A Practical Guide for Beginning Researchers*, Second Edition (no electronic document) and Stake, R.E. 1995. *The Art of Case Study Research*.

23 Typical problems with recall methods include: (i) incorrect recollection; and (ii) telescoping, i.e. projecting an event backward or forward. For example, the purchase of a durable good which took place seven years ago (before the project started) might be projected to a point in time just four years ago, during project implementation.

24 See for example http://www.gao.gov/special.pubs/10_1_9.pdf.

25 For a thorough review, see papers prepared for the International Workshop: Case studies in Development Evaluation, Validity, Generalization and Learning, Copenhagen 2012. <http://um.dk/en/danida-en/results/eval/>.

with through focus group discussions. These may cover issues such as government responsiveness to community needs, availability of privately supplied inputs, or usefulness of IFAD's involvement, all of which help to interpret the data.

Security restraints on mobility, ethical considerations or efficiency concerns may constrain the systematic adoption of random techniques. Nevertheless, there are practical ways to minimize potential biases, for example: (i) selection of project sites so as to cover different agroecological zones; (ii) surveys of beneficiaries at varying distances from a main road to ensure that the direct and indirect impacts on communities are accurately captured; (iii) examination of results in sites where project activities are at different maturity stages; (iv) targeting of communities and organizations endowed with diverse capacities (e.g. a mix of rural credit cooperatives combining high-, moderate- and low-repayment records); (v) interviews of large- and small-scale landholders, sharecroppers and landless labourers; and (vi) surveys focused on older and younger women and men.

A judicious approach should be taken to the sequencing of data collection. For example, when there is more than one mission, interpretation of quantitative survey results secured before the first mission can usefully be checked or probed through participant interviews during subsequent missions. Or, when there is only a single mission, a "sequential sampling" approach of progressive validation of initial findings can be used.

Analysing and interpreting data

The Oxford dictionaries define "evidence" as the available body of facts or information indicating whether a belief or proposition is true or valid. Keeping quality of evidence at the centre of the evaluation is key to a rigorous and credible evaluation, and the analytical phase is where the strength and validity of evidence is assessed, and any weaknesses of data gaps are addressed. Quality assurance at this stage is essential, and evaluators

must ensure that findings are adequately supported by evidence. Reviewers should assess both the quality of the evidence presented and the clarity of the analysis, identifying any possible gaps and weaknesses in the evidence.

Credibly aggregating and synthesizing evaluative evidence poses methodological challenges. Different techniques are used to categorize, clean, systematize, code, and analyse data. Some qualitative analysis can be done using simple categorization in an excel format using filter templates for example, followed by synthesis. In other cases, quantitative data need robust statistical analysis, for example to enable a distinction to be made between correlation and causation. Consideration should be given to methods such as the Qualitative Comparative Analysis²⁶ and other techniques, and the approach paper should indicate the methods foreseen, with the caveat that the data ultimately available may not support these methods and alternatives will therefore have to be found.

The quality of evidence is the backbone of a credible evaluation.

While ensuring that independent evaluations serve as instruments for strengthening accountability, concerted efforts need to be made to understand the proximate causes of good performance or to identify areas of IFAD operations that need further improvement and attention. Hence, evaluation reports should devote adequate coverage and attention to the "why" question and ensure that the numeric rating attributed to each evaluation criterion is consistent with the evidence secured by the evaluation. In addition to reporting on "what" the performance was, evaluations should provide a deeper understanding of "why" the performance was as it was. This in turn facilitates the identification and consolidation of lessons to be considered in country strategy formulation, as well as project design and implementation.

26 See for example Rihoux, B. and Ragin, C. 2009. Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and related techniques; Schneider C. and Wagemann, C. 2012. Set-theoretic methods for the social sciences. A Guide to QCA.

TABLE 3 Different types of triangulation used in mixed method evaluations

Method	Examples
Different conceptual frameworks	Comparing feminist, human rights, social exclusion or economic (e.g. cost-benefit) analysis frameworks
Different methods of data collection	Comparing structured survey, direct observation, secondary data, artifacts
Different interviewees	Comparing interviewee sex, age, ethnicity, economic status, language, etc.
Different times	Comparing responses or observations at different times of day, days of week, times of year
Different locations and contexts	Comparing responses and observations when interviews are conducted in the home when other people are present, in locations where the respondent may be able to speak more freely, in the street and other public places, at work, in the classroom, etc.

Source: Introduction to mixed methods in impact evaluation (No. 3), InterAction, August 2012.

The collection of data and other types of information from different sources and methods²⁷ – a mixed methods approach – allows the evaluation team to formulate well-founded assessments regarding important dimensions of a project and to triangulate the information and data collected. According to OECD/Development Assistance Committee (DAC), triangulation entails the use of three or more sources or types of information, or types of analysis to verify and substantiate an assessment. This allows evaluators to overcome the bias that comes from single informants, single methods or single observations and thus helps to ensure the robustness and reliability of evaluation findings.

Triangulation in an IFAD project context for example could entail looking at the views and perceptions of: (i) project beneficiaries (using, for example, a combination of survey work and participatory techniques); (ii) the country programme manager (CPM) for the relevant country and/or staff in line departments responsible for project execution (captured using a semi-structured questionnaire); and (iii) secondary sources as documented in project-related self-evaluation reports, such as periodic progress reports, MTRs and project

completion reports (PCRs). Triangulation can also use other national statistical data, when they are available at a disaggregated level.

Benchmarking

Benchmarking allows the performance of IFAD to be compared with that of other development organizations. It involves the use of a reference point or standard against which performance or achievements can be assessed.²⁸ However, there are challenges in benchmarking performance, because no two organizations are exactly the same, and often the evaluation methods and processes used by development organizations may be different.

Notwithstanding the aforementioned, benchmarking is a key aspect in IOE evaluations. Benchmarking performance against IFI and regional development banks is plausible for IOE, given that IFAD and such organizations have similar operating models (i.e. they provide loans for investments, conduct their own supervision, etc.), governance structure, and have committed to adopting harmonized evaluation methods as members of the ECG.

²⁷ See A. Tashakkori and C. Teddlie. 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*.

²⁸ See Glossary of Key Terms in Evaluation and Results-Based Management, OECD/DAC.

Benchmarking is conducted internally with IFAD operations and externally with other relevant institutions. Internally, the results of project evaluations and CSPEs will be benchmarked against the data contained in the ARRI,²⁹ but also against the data for the specific IFAD geographic region in which the project evaluation or CSPE was undertaken. Benchmarking the performance of IFAD operations across the Fund's five geographic regions will also be a feature in the ARRI.

Evaluations will also – to the extent possible – benchmark results against the indicators and targets contained in the Fund's overarching results measurement framework. Externally, as far as CSPEs are concerned, efforts should be made to collect data and compare IFAD's performance with the results of other IFI (in particular the World Bank and the appropriate regional development banks) and other international institutions, including the United Nations organizations working in agriculture and rural development, preferably in the same country.

IOE will also use benchmarking as an instrument to learn from the processes or approaches of other development organizations including from United Nations specialized agencies, programmes and funds as well as international NGOs, and bilateral aid agencies. For example, many such organizations have established country presence for much longer than IFAD, and learning from their experiences and good practices of relevance to IFAD is another form of benchmarking that will be a regular feature of IOE evaluation work, as and when appropriate. Such form of benchmarking can be particularly useful when IFAD's own experience in a particular development approach or business process might be relatively limited at the time of evaluation.

Evaluation questions and inter-evaluator variability

The evaluation questions are the point of departure for a robust evaluation. And carefully crafting the

evaluation questions is essential to ensure the usefulness and strong ownership of the process.

Evaluation questions are developed at different levels and help focus the evaluation work on a limited number of key issues, thus allowing better reflection on the evaluation criteria, more targeted data collection, more in-depth analysis, and a more useful evaluation overall. Some questions are mandatory in all project and programme evaluations to ensure consistency and comparability. In the absence of questions, choices are likely to be implicit and at risk of being biased; they may be made in relation to the most easily available information, the importance of the various stakeholders, or the evaluation team's preconceptions. Evaluators should ask themselves the question: "What is it really that we should know?"

A limited set of higher-order questions should guide the evaluation to ensure it fulfills its purposes – these are often generic and linked to specific types of evaluations (impact evaluation, evaluation syntheses, etc.). At the next level, specific questions are tailored to address the issue of focus. Questions should be further unpacked in the form of research questions in the evaluation framework. When evaluations are theory-based, a series of more detailed questions should be developed to inform the theory of change analysis. A range of generic questions, including mandatory questions linked to the evaluation criteria should also inform the evaluation, as explained in chapter 3.

“The most serious mistakes are not being made as a result of wrong answers. The truly dangerous thing is asking the wrong question.”

Peter Drucker

When the evaluation framework builds on an underlying TOC, two types of evaluation

29 The ARRI is produced every year and contains the aggregate performance assessment of all evaluations conducted in a given year. It normally covers all five IFAD geographic regions.

questions are relevant: those aiming to uncover “implementation failure” and those addressing “theory failure”. The first set of questions addresses questions linked to what was achieved and how well the approach was implemented, understanding why or why not. The focus of the second set is on the underlying theory and its assumptions, and the degree to which these can be confirmed. It answers various questions in relation to how and why, under what conditions, it worked as it did; knowledge that is essential for scaling-up and replication purposes. In all circumstances, one of the key characteristics a theory-based evaluation has is a solid analysis of the external environment that can generate findings on the influence and implications of external factors to the achievement of results.

The evaluation questions must meet certain criteria. They must first of all be relevant and pertinent. If the evaluation fails to answer key questions for major stakeholders, it will lose relevance and credibility; if it fails to ask the right questions to explore the TOC, it may end up with wrong assumptions and hence irrelevant findings and recommendations; and if it fails to ask mandatory questions that are common to all evaluations of a certain type, it will jeopardize IOE’s ability to draw comparisons. Secondly, evaluation questions must be feasible; this relates to the scope of the exercise, the boundaries of the evaluation, and the availability of resources for the evaluation; questions must be commensurate with resources, including time. Getting the evaluation questions right is the sound footing for any evaluation, and evaluators cannot achieve this alone – thus the third criteria relates to the process; questions must be developed using a participatory process to ensure relevance and usefulness for stakeholders, and sufficient time and a strong dialogue process are at the core of the process.

Finally, it is essential, that IOE evaluations minimize inter-evaluator variability. The term “inter-evaluator variability” refers to differences in assessments and judgements related to

performance, which primarily stem from differences in the understanding and application of evaluation methodology by IOE staff and consultants. This is a legitimate cause for concern – not only within a given evaluation but also across evaluations – as inter-evaluator variability limits the degree of reliability obtained when aggregating results as well as in comparing results between evaluations.

In addition to providing guidance for the use of numerical ratings, special efforts are being deployed by IOE to minimize such variability, although it should be recognized that it is not possible to completely eliminate this due to the human factor. These efforts include holding a mandatory briefing session for all senior consultants and selected team members on the evaluation methodology and the process to follow; undertaking systematic internal peer reviews within IOE of major deliverables³⁰ produced during evaluations; requesting each evaluation team member to comment on the ratings and draft final evaluation report; and conducting periodic workshops in IOE to provide guidance to staff and evaluation consultants on the methodologies to be followed and their application. Furthermore, IOE is introducing the use of descriptors to support ratings, which will further strengthen comparability and robustness (table 6).

Theory of change

Particularly for the theory-based evaluations, understanding the TOC³¹ is critical. The TOC may be explicitly stated or, as is often the case, is implicit and therefore needs to be articulate and clarified. A TOC articulates how activities are linked to outcome and impact, clearly outlining the assumptions behind, including the implicit assumptions. Assumptions would pertain to external influence, for example, or the assumed behavioural effect of a particular intervention on the target group. It should also capture assumptions about the distributional effects of the interventions in terms of anticipated project benefits to different groups, based on gender, age, geographical location, etc.

30 Mainly the evaluation approach paper and the draft final report.

31 Mayne, J. 2015. *Useful Theory of Change Models*.

The purpose of a TOC is to describe how an intervention leads to results. A TOC illustrates, generally both in graphic and narrative form, the series of assumptions and links underpinning the presumed causal relationships from input to impact; it describes the pathways to change. A TOC also has to factor in the context, something that is often missing in classical logframes, but is essential for understanding when, how, why, and under what conditions certain actions lead to certain outcomes, or not, i.e. it is to a very large degree tailored for learning purposes. In this respect, a distinction can be made between programme-level and project-level change. The former would apply for example to initiatives aiming at reforms that might facilitate entry into higher value-added areas of production, enhance systemic climate resilience, etc., while the latter might aim, for example, at delivering a set of inputs designed for a beneficiary group of farmers who were not previously using that input. In the former, evaluation tools would need to be able to capture systemic and institutional change, whereas in the latter, the evaluation tools would need to identify uptake, whether beneficiaries understand and trust the technology/input, and whether they have sufficient skill and knowledge.

A TOC may be developed based on past evaluations or theoretical and empirical research on processes of change in similar interventions. It may also be derived from the intervention's existing logical framework; however it must be developed not merely based on a document describing the logframe, but based on stakeholder interviews that reveal insights and expectations on the presumed way target groups are affected. In all cases it is important

Getting the TOC right is one of the basic conditions for a good evaluation and therefore sufficient time, effort and resources must be allocated for this.

to critically look for, and articulate, plausible explanations for the anticipated changes that can be tested through the evaluation questions and for which evidence may be found.

This said, it should also be kept in mind that use of TOC is a fairly recent phenomenon in IFAD. Therefore, an explicit TOC may not be available in many policies, strategies or operations, and reconstructing the TOC is not always an easy task. In newer projects, it is to be expected that an explicit TOC has been developed, as TOC is not only an evaluation tool, but is also recognized as an excellent tool for stakeholder dialogue in project preparation.

Strong input from stakeholders and beneficiaries is required and therefore the process leading to the articulation of the TOC is important. A robust dialogue with stakeholders will help build ownership, clarify the assumptions, confirm the plausibility of the TOC, and give visibility to those assumptions that are implicit; it can shed light on the possible discrepancy between what is written in various documents and what is the reality.

This process is time-consuming, may be costly, and is not applicable in all cases. Sound judgement should therefore be exercised in determining the necessity and scope of the exercise. In some cases, it is advisable to keep the TOC simple, to use the logframe and make the tacit hypotheses clearer and highlight gaps. However, it should also be noted that as experience is gained, and with systematic sharing of experience and building up a knowledge base, it will be less time-consuming and the quality of the TOC in IFAD documents may improve.

Much has been written about TOC and how to develop a strong TOC. Annex VI provides suggested reading and sources on learning, and more details are given in chapter 8 on impact evaluation, where the construction of a TOC is critical.

Evaluation and hindsight

There are four evaluation dilemmas that evaluators often need to address:

- How to evaluate performance of a strategy or operation if the context has changed in terms of, for example, the country's policy framework or institutional arrangements?
- How to evaluate performance if development understandings have changed since the beginning of a strategy or operation?
- How to evaluate performance if IFAD policies, processes or features of its operating model have changed during implementation?
- How to evaluate operations whose original targets, approved by the Board as part of design, have changed in the course of implementation?

Common sense might appear to suggest that the evaluation of past performance should be measured by yesterday's metric and the advantages of hindsight disregarded. However, it also stands to reason that results should be judged based on up-to-date information regarding actual achievements and policy standards. Performance cannot be rated as if project and programme designs were immutable and immune to adjustment in the course of implementation. Given the learning dimension of IFAD operations, the adaptability of its instruments and practices should be an important aspect of its performance evaluations. At the same time, revisions of project and programme designs are not cost-free and require the concurrence of partners, in particular the borrowers. Hence, a sound assessment of adjustment feasibility should be carried out by evaluators to ensure fairness in performance ratings.

In other words, in this case learning should be distinguished from accountability. Learning is maximized when it is evaluated against today's standards. But to hold managers accountable for failing to achieve today's standards before they were known may be unfair. For example, to judge the relevance and quality of project designs without reference to the limits of the knowledge available at the time would not be fair. Equally, one cannot expect a rural finance project that is scheduled to close at the end 2015 to be retrofitted to meet the provisions of a new rural finance policy introduced

by the Fund in 2014. In cases where standards or policies have changed late in the life of a project – too late for retrofitting – managerial performance must be evaluated without the benefit of hindsight.

The costs of retrofitting must also be taken into account, for example, in the case of projects or components that cannot be readily changed without prohibitive consequences (e.g. irrigation systems or rural roads). On the other hand, where project and programme designs could have been adjusted economically and in a timely fashion so as to remain relevant as time passed and circumstances changed, performance evaluation with the benefit of hindsight is both legitimate and fair. To sum up, the judicious use of a current evaluative lens³² allows project performance and impact to be assessed against current standards, and these standards should be clear in the approach paper. This is particularly pertinent in relation to the assessment of IFAD's performance and in this case, evaluation with the benefit of hindsight is important for learning and also for accountability of IFAD's supervision, and should be used to facilitate and encourage adaptation of IFAD's interventions.

Experience shows that the overall development context (e.g. political, agroecological, policy, institutional) in a given country and project area has a significant effect on results. This means that project design teams must factor in context issues up front and that project strategy should be fine-tuned as required during implementation to respond to changes in the implementation environment. Evaluation teams should therefore ascertain the nature of the development context at the design stage, and track its evolution together with adjustments made to the project, including targets, and determine the adequacy, feasibility and rationale of adjustments in the course of implementation, including at the MTR stage. Targets may be changed up until the MTR, provided that the MTR takes place no later than midway through the project. The MTR constitutes the cut-off point for making major adjustments to project/programme targets. Teams are therefore encouraged to ascertain the nature of

32 Evaluation approach papers should clarify instances when a current evaluative lens will be applied to assess a specific policy or operation.

the development context at the design stage and at the time of the evaluation, as a basis to determine the adequacy of the adjustments made during the course of project implementation. Thus, the benefits of hindsight are used to learn from experience and answer the important question of what could have been done differently in order to maximize results.

Transparency and the evidence trail

The credibility and quality of each evaluation are based on the robustness of its evidence and analysis; one consideration in this regard is the importance of ensuring transparency and a clear evidence trail. In terms of transparency, the sources on which findings are based must be clear, and their reliability and validity transparently assessed. Records of interviews should be kept to allow for tracking back to the sources behind a particular finding. And the findings and conclusions of a particular evaluation should be coherently anchored in the analysis and documented in evaluation reports. Each recommendation should find its genesis in the conclusions contained in the evaluation. Moreover, in order to delineate the evidence trail, evaluation reports should contain cross-references to the pertinent sections and paragraphs in the document to help readers easily identify the findings that led to a particular recommendation and the analysis that led to a particular conclusion.

Formulating recommendations

The above sections focus on how to make evaluations credible and useful. In this regard, the Executive Board noted that IOE should not be expected to make cost-neutral recommendations: it is the role of the Board, together with IFAD Management, to decide on the final trade-off between budget implications and the impact of specific recommendations.³³

Care must be taken to ensure that recommendations are appropriate for achieving the objectives of the interventions, are positioned

strategically, and once implemented, will add value to the organization.

Recommendations must be grounded in solidly evidenced findings and follow logically from the conclusions. It must be possible to track back from a recommendation to the evidence that supports this recommendation and with the transparency described above, be assured that the recommendation is solidly anchored in coherent analysis.

The full utility of an evaluation hinges on dissemination, learning and follow-up and therefore recommendations should be presented in a form that allows different decision-makers to clearly identify their responsibility, and to track follow-up action in IFAD's system for the purpose of the President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA) (see also section below). Annex VI contains a reference to the UNEG good practice note on ensuring appropriate follow up to evaluations.

Learning accountability and follow-up

Given IFAD's learning mandate, IOE's performance assessment should include an evaluation of the Fund's quality assurance mechanisms, risk management systems, and the adaptability of its instruments and practices. In particular, IOE evaluations should analyse the extent to which recommendations from past evaluations, PCRs, MTRs and supervision and implementation support missions were reflected in the project/programme/policy under consideration. Evaluations should review whether the advice generated through IFAD's quality enhancement and quality assurance processes³⁴ was internalized in the subsequent phases of project/programme/strategy/policy development.

Related to this, IOE will undertake evaluations more than once in the same country (e.g. repeat CSPEs in the same countries, after a certain number of years from the previous CSPE) or on the same theme (for example, IOE did a CLE on gender in

33 See minutes of the December 2014 Executive Board: <https://webapps.ifad.org/members/eb/113/docs/EB-113.pdf>.

34 These are new processes introduced under IFAD's Action Plan to improve quality at entry. The quality enhancement process is the responsibility of the Programme Management Department (PMD), whereas the quality assurance process is undertaken at arm's length from PMD under the responsibility of the Office of the Vice-President.

2010 and might conduct a similar evaluation on the same topic in the near future). In this regard, such evaluations will specifically entail a careful assessment of how IFAD Management and/or governments (for CSPEs) have acted upon the lessons and recommendations from the previous evaluation(s). A more comprehensive treatment of how evaluation results strengthen the learning feedback loop is included in chapter 10 on communication, dissemination and learning.

As required by the IFAD Evaluation Policy, the PRISMA contains an account of the follow-up actions taken by Management and government (where applicable) on the recommendations made by independent evaluations. It is a key instrument to promote learning and improvement, and is presented each year to the Evaluation Committee and Executive Board together with IOE comments thereon. The PRISMA includes follow-up actions taken to recommendations made in the ARRI, CLEs, CSPEs, PPEs and evaluation synthesis reports. The PRISMA also includes an account of how IOE comments on the Report on IFAD's Development Effectiveness have been treated by Management.

For the ARRI, this includes recommendations agreed with the Board. For CSPEs, these are extracted from agreements at completion point finalized at the end of the evaluation. For CLEs and evaluation synthesis reports, these are outlined in Management's responses or action plans submitted to and agreed by the Executive Board following evaluations. As far as PPEs are concerned, the PRISMA only includes the recommendations addressed to IFAD Management.

In terms of process, after the evaluations are selected, recommendations are extracted, listed and classified by selected themes for follow-up. The list is prepared by the Programme Management Department (PMD) Front Office and reviewed by IOE. The classification criteria used include level (country, project etc.), nature (operational, strategic, policy) and theme.³⁵

Joint evaluations

Joint evaluations³⁶ can contribute to progress towards implementation of the provisions contained in the Paris Declaration which are aimed at promoting aid effectiveness and can help overcome attribution problems in assessing the effectiveness of programmes and strategies, the complementarity of partners' contributions, the quality of aid coordination, etc. Of course, there are various degrees of "jointness", depending on the extent to which individual partners cooperate in the evaluation process, merge their evaluation resources and combine their evaluation reporting.

Joint evaluations permit the pooling of resources to undertake more effective desk and country fieldwork that will almost certainly add to the coverage, quality and credibility of the evaluation. They also provide an opportunity to draw together the substantial volume of evaluative evidence, experience and knowledge accumulated by more than one organization. Properly managed, they may also help reduce transactions costs, which the country or countries covered by the evaluation must assume. On the other hand, joint evaluations are more demanding than single-owner studies because there is increased scope for disagreement on methods, priorities, findings or resource management and because of sensitivities about reputational risk. Similarly, the coordination of joint work has often proven complex and has increased the cost and duration of the evaluation exercise as a result of the coordination costs incurred in terms of staff time and travel, lengthy management processes and other factors.³⁷

In view of changes to the aid architecture and stronger evaluation capacity in some countries, and the increasing use of partnership approaches, evaluation on a broader scale than the traditional project, sector or thematic levels is likely to be required, not only to assess results at the country level but also to look more closely at the role of the different institutions in the process. Joint or multi-aid agency evaluations can provide this broader

35 Currently categorized in six broad categories: targeting and gender; technical areas; project management; non-lending activities; cross-cutting themes; and corporate issues, mainly information and communication technology and human resources issues.

36 IOE worked on a major joint evaluation with the African Development Bank on agriculture and rural development in Africa in 2007-2008. A section on the joint evaluation is found in the IOE website on the IFAD portal (www.ifad.org/evaluation).

37 For a more comprehensive discussion of these matters, see *Guidance for Managing Joint Evaluations*, DAC Evaluation Series, OECD, Paris, 2006.

perspective while fostering cross-agency learning and reducing evaluation transaction costs for in-country stakeholders; there may thus be a shifting of transaction costs from the partner country to IFAD, which is well in line with the Paris Declaration and IFAD's commitment to partnerships.

IOE will carefully consider joint evaluations and participate selectively in them, including with the Rome-based agencies, with which such evaluations have already been carried out. IOE will continue to contribute to the ongoing debate on joint evaluations taking place within the framework of UNEG, ECG and the DAC Network on Development Evaluation.

Participation

The DAC Quality Standards for Development Evaluation³⁸ contains a standard related to stakeholder engagement: "Relevant stakeholders are involved early on in the evaluation process and given the opportunity to contribute to evaluation design, including by identifying issues to be addressed and evaluation questions to be answered." Research in turn suggests that this seems to enhance the use of evaluation findings and that learning and the use of findings are stronger when the link between users and producers of knowledge is short. This is reinforced through one of the basic tenets of IFAD's Evaluation Policy, namely "partnership" as the policy states that IOE's procedures and instruments aim at the appropriate engagement of partners in the evaluation process while safeguarding the independent role of IOE.

Involving beneficiaries is particularly important, including at the design stage where they may bring a different perspective than policymakers. Involving beneficiaries in the early stages may broaden the focus of the evaluation, but is also perceived to enhance fairness and provide for a more balanced evaluation, and programme stakeholders are often aware of contextual considerations of which evaluators are not. Therefore, by including stakeholders in the evaluation process, the evaluation is more likely to identify important problems of

concern³⁹ and hence ultimately enhance credibility and utility.

IOE evaluations will pay due attention to promoting participation of key stakeholders in the evaluation process, but at the same time ensure that the independence of the evaluation's analysis and final judgements are not compromised. Key stakeholders who IOE will seek to involve – at different stages of an evaluation, which will vary depending on the type of evaluation being conducted – include IFAD Management and staff, recipient governments, cofinanciers, and ultimately, beneficiaries and their organizations.

Participation will be promoted in several ways. For example, key stakeholders will be given an opportunity to provide inputs at an early stage of an evaluation so that their concerns and priorities are captured (e.g. in the form of key questions) in the design of the evaluation. Evaluations will ensure that different stakeholders may interact with IOE at selected junctures, so that they may share their experiences, data and information to facilitate an informed evaluation. They will also be given a chance to comment on the draft final report, and IOE will carefully consider their feedback before finalizing it, with due consideration to the independence of IOE. Beneficiaries would also be involved at other stages of the evaluation, for example, at the time when final workshops are held at the country level by IOE to discuss the main conclusions and recommendations from CSPEs.

Evaluation management

Managing an evaluation involves agreeing on how decisions will be made for each cluster of the evaluation (from designing an evaluation to reporting and supporting use) and ensuring they are implemented well. Three aspects are particularly important: transparency, ownership and maximizing learning.

To demonstrate transparency, evaluation processes and methodology should be discussed with key evaluation partners,⁴⁰ while evaluation

³⁸ <http://www.oecd.org/development/evaluation/qualitystandards.pdf>.

³⁹ See A.E. Cullen, C.L.S. Coryn and J. Rugh: *The Politics and Consequences of Including Stakeholders in International Development Evaluation*. American Journal of Evaluation, 2011.

⁴⁰ Key partners may include, for example, IFAD Management, government authorities and project staff who are the main users of evaluation results.

reports should include data and information that adequately support the conclusions as well as relevant descriptions of the evaluation process, the literature consulted and the working papers generated to support report findings and recommendations. The entire evaluation report and supporting working papers are disclosed together with the approach paper.

In accordance with the IFAD Evaluation Policy, the designated IOE lead evaluator will be responsible for the quality and contents of the final evaluation report. In line with this Policy, IOE staff leads the evaluation with the required technical inputs from consultants. IOE will make efforts to mobilize national consultants, especially for project evaluations and CSPEs, and ensure they satisfy all necessary provisions in IOE's conflict of interest policy (see annex IV) for consultants. Efforts will also be made to ensure gender balance of evaluation teams.

With regard to learning, much informal learning actually takes place during the evaluation process as evaluators engage with stakeholders and beneficiaries and ask questions, often giving rise to new insights and ways of seeing and appreciating things. Formal learning is more embedded in "uptake chains" and good feedback loops; this requires that lessons learned are absorbed in a systematic, structured and effective way into policy and practice. This is further explored in chapter 10 on communication, dissemination and learning.

3. Evaluation criteria, key questions and ratings

Background

In line with good practice in international development evaluation, IOE uses a set of evaluation criteria to assess the performance of IFAD policies, strategies, operations and business processes. In the different types of evaluations it undertakes, IOE rates (i.e. scores) performance across each criterion using a six-point rating scale. It also has two specific evaluation criteria – project performance and overall project achievement – that provide an integrated assessment of performance of IFAD interventions.

Evaluation criteria applied in IOE evaluations ensure the harmonization of IFAD's evaluation methodology across international development organizations. The evaluation criteria adopted are in line with the practices set out in the OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management,⁴¹ the UNEG's Norms and Standards, as well as ECG's Good Practice Standards for the Evaluation of Public Sector Operations. Taken together, these criteria allow for a solid assessment of the key aspects of IFAD initiatives to be made, and also enable comparison with peers who apply similar criteria to judge performance.

IOE has introduced additional evaluation criteria⁴² which allows it to assess key dimensions that are specific to IFAD's work and operating model. The criteria used, their definitions and the rating system will be discussed later in the chapter.

To help conduct performance assessments and ensure consistency in the foundation for ratings, IOE uses a number of core questions for each evaluation criterion. These core questions have been carefully crafted, based on the internationally recognized

definition of each evaluation criterion. The core questions are therefore fundamental and evaluators should aim to provide answers to all of them and assign a rating to each evaluation criteria. The answers to each question should be clearly outlined in each evaluation report. Additional evidence for each question should be available, possibly in a synthesized form, in the documentation retained on the evaluation for transparency purposes, but does not have to be presented in the report. The use of core questions will help ensure consistency and comparability as well as allow for aggregation of ratings across IOE evaluations.

In addition to core questions, evaluators may also introduce other questions for each criterion – to ensure a comprehensive assessment of the intervention being evaluated. Additional questions are meant to deepen the inquiry and provide further evidence to support the finding and assessment. The approach papers, prepared at the outset of each evaluation, should specify the criteria, how they are applied, and the questions for assessing each criterion. Illustrative additional questions are presented in annex II.

Evaluation criteria and their definitions

The evaluation criteria used by IOE and their definitions are shown in table 4. However, it should be kept in mind that not all criteria are necessarily relevant to all types of evaluations.

⁴¹ <http://www.oecd.org/development/evaluation/2754804.pdf>.

⁴² Innovation and scaling up, gender equality and women's empowerment, environment and natural resources management and climate change.

TABLE 4 Definition and rating of the evaluation criteria used by IOE

Criteria	Definition*	Mandatory	To be rated
Rural poverty impact	Impact is defined as the changes that have occurred or are expected to occur in the lives of the rural poor (whether positive or negative, direct or indirect, intended or unintended) as a result of development interventions.	X	Yes
	<i>Four impact domains</i>		
	<ul style="list-style-type: none"> Household income and net assets: Household income provides a means of assessing the flow of economic benefits accruing to an individual or group, whereas assets relate to a stock of accumulated items of economic value. The analysis must include an assessment of trends in equality over time. 		No
	<ul style="list-style-type: none"> Human and social capital and empowerment: Human and social capital and empowerment include an assessment of the changes that have occurred in the empowerment of individuals, the quality of grass-roots organizations and institutions, the poor's individual and collective capacity, and in particular, the extent to which specific groups such as youth are included or excluded from the development process. 		No
	<ul style="list-style-type: none"> Food security and agricultural productivity: Changes in food security relate to availability, stability, affordability and access to food and stability of access, whereas changes in agricultural productivity are measured in terms of yields; nutrition relates to the nutritional value of food and child malnutrition. 		No
	<ul style="list-style-type: none"> Institutions and policies: The criterion relating to institutions and policies is designed to assess changes in the quality and performance of institutions, policies and the regulatory framework that influence the lives of the poor. 		No
Project performance	Project performance is an average of the ratings for relevance, effectiveness, efficiency and sustainability of benefits.	X	Yes
Relevance	The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, institutional priorities and partner and donor policies. It also entails an assessment of project design and coherence in achieving its objectives. An assessment should also be made of whether objectives and design address inequality, for example, by assessing the relevance of targeting strategies adopted.	X	Yes
Effectiveness	The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.	X	Yes
Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted into results.	X	Yes
Sustainability of benefits	The likely continuation of net benefits from a development intervention beyond the phase of external funding support. It also includes an assessment of the likelihood that actual and anticipated results will be resilient to risks beyond the project's life.	X	Yes

Criteria	Definition*	Mandatory	To be rated
Other performance criteria			
Gender equality and women's empowerment	The extent to which IFAD interventions have contributed to better gender equality and women's empowerment, for example, in terms of women's access to and ownership of assets, resources and services; participation in decision making; work load balance and impact on women's incomes, nutrition and livelihoods.	X	Yes
Innovation and scaling up	The extent to which IFAD development interventions: (i) have introduced innovative approaches to rural poverty reduction; and (ii) have been (or are likely to be) scaled up by government authorities, donor organizations, the private sector and others agencies.	X	Yes
Environment and natural resources management	The extent to which IFAD development interventions contribute to resilient livelihoods and ecosystems. The focus is on the use and management of the natural environment, including natural resources defined as raw materials used for socio-economic and cultural purposes, and ecosystems and biodiversity – with the goods and services they provide.	X	Yes
Adaptation to climate change	The contribution of the project to reducing the negative impacts of climate change through dedicated adaptation or risk reduction measures.	X	Yes
Overall project achievement	This provides an overarching assessment of the intervention, drawing upon the analysis and ratings for rural poverty impact, relevance, effectiveness, efficiency, sustainability of benefits, gender equality and women's empowerment, innovation and scaling up, as well as environment and natural resources management, and adaptation to climate change.	X	Yes
Performance of partners			
• IFAD	This criterion assesses the contribution of partners to project design, execution, monitoring and reporting, supervision and implementation support, and evaluation. The performance of each partner will be assessed on an individual basis with a view to the partner's expected role and responsibility in the project life cycle.	X	Yes
• Government		X	Yes

*These definitions build on the OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management; the Methodological Framework for Project Evaluation agreed with the Evaluation Committee in September 2003; the first edition of the Evaluation Manual discussed with the Evaluation Committee in December 2008; and further discussions with the Evaluation Committee in November 2010 on IOE's evaluation criteria and key questions.

Evaluation questions by criterion

This section provides an overview of the core questions to be addressed by each evaluation criterion listed in table 4. As mentioned before, an indicative set of additional questions by criterion is provided in annex II.

Rural poverty impact

Given the international emphasis on measuring impact of aid interventions and IFAD's commitment to a significant poverty impact, building up a strong database with evidence on impact is indispensable. IOE applies a specific criterion to gauge rural poverty impact, which should be treated with the utmost thoroughness. The criterion will be rated to enable comparisons and tracking of trends over time. The following are the core questions all evaluations should try to answer in assessing rural poverty impact:

- Has the initiative had the anticipated impact on the target group?⁴³
- To what extent have beneficiary incomes changed as a result of the project?
- In what way have household net assets changed due to the intervention?
- What changes have taken place in household food security and nutrition and what explains such changes?

In addition, and in light of IFAD's mandate, there are four impact domains in areas of priority for rural transformation promoted by IFAD: (i) household incomes and assets; (ii) human and social capital and empowerment; (iii) food security and agricultural productivity; and (iv) institutions and policies.

Covering the four domains is not mandatory in every evaluation, and the performance of operations in any of these areas is not to be rated. The evaluation approach papers should clarify which domains will be covered and why.

Project performance

This is a composite criterion made up of four core performance criteria: relevance, effectiveness, efficiency and sustainability of benefits. They are mandatory and are always applied and rated in CSPEs and project evaluations. Where an exception is made, the reason must be made clear in the approach paper and the final evaluation report. Each of these four criteria will be assessed and rated individually. Evaluations will determine overall project performance by taking an average of the four individual ratings for relevance, effectiveness, efficiency and sustainability of benefits.

The two criteria, effectiveness and impact, while distinct, are closely linked. The effectiveness criterion assesses whether the intervention has met (or is expected to meet) its objectives. By contrast, the rural poverty impact criterion takes on board all effects, intended or unintended, direct or indirect, positive or negative, and thus requires careful examination if it is to be used to shed light on IFAD's role as an agent of rural change. The risk of duplication when assessing these is mitigated by focusing the effectiveness criterion on the achievement of the immediate outcomes of the project and on the initial effects, whereas all side effects and longer-term effects are captured by the impact criterion. The use of these criteria thus leads to a deeper understanding of the forward and backward linkages of an IFAD-funded operation. As a result, they help guide efforts to scale up such operations and to orient the design of future IFAD projects.

Sustainability of benefits

This is a multi-dimensional concept as acknowledged in Occasional paper 8: *Sustainability of Rural Development Projects* (IFAD 2009), and merits careful assessment. In line with the OECD/DAC definition, sustainability entails two aspects: (i) measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn; and (ii) assessing if benefits are

⁴³ This entails trying to determine and report the total approximate number of people (gender disaggregated) who have been affected by the initiative being evaluated.

environmentally as well as financially sustainable, i.e. the likelihood that actual and anticipated results will be resilient to risks beyond the project's life. (Environmental aspects are explored separately under the criterion environment and natural resources management). This criterion should be assessed in all IOE evaluations. Country strategy and programme and project evaluations should include a rating for this criterion.

Below are the core questions for assessing and rating relevance, effectiveness, efficiency and sustainability of benefits.

- **Relevance:** (i) Are project objectives in line with key IFAD and government objectives for promoting sustainable agriculture development as well as the needs of the rural poor? (ii) Was project design appropriate (for example, in terms of components, financial allocations, institutional arrangements, etc.) to meet the intervention's objectives? (iii) Was the project adjusted during implementation to any changes in context to retain continued relevance?
- **Effectiveness:** (i) To what extent have the objectives of the project and its components been attained in quantitative and in qualitative terms? (ii) In particular, what changes in the overall context (e.g. policy framework, political situation, institutional set-up, economic shocks, civil unrest) have affected or are likely to affect project implementation and overall results? (iii) What factors in project design and implementation account for the estimated results in terms of effectiveness; are there valid alternatives?
- **Efficiency:** (i) How does the economic rate of return at evaluation compare with that at project design?⁴⁴ (ii) What are the loan costs per beneficiary (both at the time of appraisal and at the time of evaluation) and how do they compare to other IFAD-funded operations (or those of other donors) in the same country and/

or other countries? (iii) What are the total project management costs in relation to total project costs and how do they compare with similar projects?

- **Sustainability of benefits:** (i) Do project activities benefit from the engagement, participation and ownership of local communities, grass-roots organizations and the rural poor, and are adopted approaches technically viable? (ii) Is there a clear indication of government commitment after the loan closing date, for example, in terms of provision of funds for selected activities, human resources availability, continuity of pro-poor policies and participatory development approaches, and institutional support? (iii) What are the chances that benefits generated by the project will continue after project closure and what is the likely resilience of economic activities to post-project risks?

Gender equality and women's empowerment

This criterion is also assessed in all evaluations by IOE, and therefore country strategy and programme and project evaluations must include a rating for this criterion. A dedicated section in all evaluation reports provides an overall assessment of IFAD's efforts in promoting gender equality and women's empowerment. Each evaluation assesses performance based on the main objectives outlined in the relevant country strategies and project design, and the 2012 corporate gender policy⁴⁵ or any revisions to the policy in the future.⁴⁶ Evaluations will also assess IFAD's efforts to integrate the main provisions in the handbook issued by the United Nations Evaluation Group on human rights and gender equality.⁴⁷

The following are the core questions that evaluations must cover to assess efforts made to promote gender equality and women's empowerment:

⁴⁴ To be assessed only where adequate evidence exists, and using the following internal guidelines: https://xdesk.ifad.org/sites/pt/FAME/EFA/Economic%20%20Fiancial%20Analysis%20Documents/Guidelines/EFA_guidelines_Aug2013.pdf. (This is an internal link).

⁴⁵ http://www.ifad.org/gender/policy/gender_e.pdf.

⁴⁶ A gender marker system was developed by the Gender Desk of the Policy and Technical Advisory Division in 2014 to provide a uniform method to assess the gender sensitivity of IFAD projects at various stages of the project cycle. See <http://www.slideshare.net/GenderSecretariat/ifad-gender-markers>.

⁴⁷ See document *Integrating Human Rights and Gender Equality in Evaluation*. 2012. The document is found on the following website: <http://www.uneval.org/document/detail/980>.

- What were the project's achievements in terms of promoting gender equality and women's empowerment? This includes assessing whether there are changes: to women's access to resources, assets and services; to women's influence in decision-making; in workload distribution among household members; in women's health, skills, income and nutritional levels; in gender relations within households, groups and communities in the project area; etc.
- What percentage of total project resources was invested in activities to promote gender equality and women's empowerment and how does that compare with other projects funded by IFAD?
- To what extent did the project define and monitor sex-disaggregated results to ensure that gender equality and women's empowerment objectives were being met?
- Was the project implementation structure adequate to support effective implementation of gender equality and women's empowerment goals?

Innovation and scaling up

IFAD devotes priority attention to promoting pro-poor innovations and scaling up of successful innovative approaches introduced in the context of IFAD operations. The assessment of innovation is to be based on the IFAD Innovation Strategy⁴⁸ or further enhancements thereof over time, as well as findings and recommendations from the CLE on innovation and scaling up,⁴⁹ and the dedicated document presented to the Tenth Replenishment Consultation in 2014 on Scaling Up Results.⁵⁰

With regard to the latter, evaluations will also consider IFAD's "operational framework for scaling up results" (2015), in particular by paying attention to whether design and implementation paid due attention to the "pathways, drivers, and spaces" for scaling up results of IFAD operations. The operational framework notes that: (i) the definition adopted by IFAD for scaling up is: "expanding, adapting and supporting successful policies, programmes and knowledge so that they can

leverage resources and partners to deliver larger results for a greater number of rural poor in a sustainable way"; and (ii) scaling up results does not mean transforming small IFAD projects into larger projects. Instead, IFAD interventions will focus on how successful local initiatives will sustainably leverage policy changes, additional resources and learning to bring the results to scale.

In assessing this criterion, evaluators should give only one final consolidated rating. In determining the final rating, both innovation and scaling up should be considered equally in the analysis, narrative and judgement. It is important that there is a strong narrative for each component, even if it is a joint rating, so as to capture examples where innovative approaches were introduced but without scaling them up, for example, and/or innovative approaches were not introduced but scaling-up took place. The below core questions should be answered by evaluators in assessing innovation and scaling up.

- **Innovations:** (i) What are the characteristics of innovation(s) promoted by the intervention? (ii) Are the innovations consistent with the IFAD definition of this concept? (iii) Are the actions in question truly innovative or are they well-established elsewhere but new to the country or project area? (iv) Have grants been used to promote innovation?
- **Scaling up:** (i) What evidence was used to justify scaling up, and were successfully promoted innovations documented and shared to facilitate scaling up? (ii) Has IFAD proactively engaged in partnership-building and policy dialogue to facilitate the uptake of successful innovations? (iii) Based on the information available, have these innovations been scaled up and, if so, by whom? If not, what are the prospects at the time of evaluation that they can and will be scaled up by the government, other donors and/or the private sector? What were/are the pathways to scaling up?

48 <http://www.ifad.org/gbdocs/eb/91/e/EB-2007-91-R-3-Rev-1.pdf>.

49 http://www.ifad.org/evaluation/public_html/eksyst/doc/corporate/scalingup.pdf.

50 <https://webapps.ifad.org/members/repl/10/3/docs/IFAD10-3-R-2.pdf>.

Environment and natural resource management

This criterion should also be assessed in all evaluations by IOE.⁵¹ Country strategy and programme and project evaluations should include a rating for this criterion. A dedicated section in all evaluation reports would provide an overall assessment of IFAD's efforts in environment and natural resource management. Each evaluation would assess performance based on the main objectives outlined in the relevant country strategies and project design and IFAD's Environment and Natural Resource Management Policy.⁵²

The following are the core questions that evaluations would cover:

- To what extent did the project adopt approaches/ measures for restoration or sustainable management of natural resources (e.g. enhancement of ecosystem services, support to training and extension to foster efficient environment and natural resource management, uptake of appropriate/new technologies)?
- To what extent did the project develop the capacity of community groups and institutions to manage environmental risks (e.g. how governance-related factors are shaping the management of natural resources, influence of incentives and disincentives for sustainable natural resource use and natural resource-based livelihoods improvement)?
- To what extent did the project contribute to reducing the environmental vulnerability of the community and built resilience for sustainable natural resource management that contribute to poverty reduction (e.g. factors such as access to technologies, information/awareness creation)?
- To what extent did the project contribute to long-term environmental and social sustainability (e.g. through avoiding over exploitation of natural resources or loss of biodiversity or reduction of the community's livelihoods); and by empowering and strengthening the capacity of community-based natural resource management groups to ensure sustainable natural resources

management; and by ensuring strong stakeholder engagement, especially of vulnerable groups, in decision making affecting natural resources use?

- To what extent did the project follow required environmental and social risk assessment procedures (e.g. Social, Environmental and Climate Assessment Procedures), including meaningful consultation with affected and vulnerable communities, and have complied with applicable IFAD or national environmental and social standards or norms, to ensure any harmful impacts are avoided or managed/mitigated through, where needed, the implementation of effective environmental and social management plans, including robust monitoring and supervision?

Adaptation to climate change

The contribution of the project to increase climate resilience and increase beneficiaries' capacity to manage short- and long-term climate risks will be assessed in all evaluations by IOE contingent on the adequacy of evaluative evidence. This dimension has only recently become an IFAD priority and hence older projects may not have addressed this and little evidence will exist. Country programme and project evaluations would include a rating and a dedicated section in all evaluation reports will provide an overall assessment of IFAD's efforts. Each evaluation should assess performance based on the main objectives outlined in the relevant country strategies and project design and IFAD's Climate Change Strategy.⁵³

The following are the core questions that evaluations would cover:

- To what extent did the project demonstrate awareness and analysis of current and future climate risks?
- What are the amounts and nature of funds allocated to adaptation to climate change-related risks?
- What were the most important factors that helped the rural poor to restore the natural resource and environment base that (may) have been affected by climate change?

51 For projects and programmes pre-dating IFAD's Environment and Natural Resource Management Policy, this is optional and contingent on sufficient evaluative evidence.

52 IFAD (2012) Environment and Natural Resource Management Policy.

53 <http://www.ifad.org/climate/strategy/e.pdf>.

Performance of partners

The performance of each partner – IFAD and government – is examined and reported on separately, as each has specific functions and roles to discharge. All country programme and project evaluations should assess and attribute a rating to the performance of IFAD and the government.

With regard to the performance of the relevant government, it is important to note that IOE assesses the performance of government in the agriculture sector, with specific reference to IFAD operations. IOE evaluations do not aim to assess government performance in general terms.

Moreover, it is noted that different government ministries, departments and line agencies (at the national, provincial and local levels) may be involved in project execution. As such, evaluations will need to make an assessment of the individual institutions involved and, ultimately, come up with a comprehensive overall rating for the performance of the government concerned. In light of the sensitivities involved, special attention should be devoted to discussing the assessment and ratings of government performance with the main government agencies involved and to seeking, on a joint basis, ways and means to enhance government performance in the future, as may be required.

The following are the core questions to be covered in assessing and rating IFAD and government performance, respectively.

- **IFAD:** (i) How well were the comments and recommendations of quality enhancement and quality assurance processes, including from evaluations, included in the final project design? (ii) Did IFAD have a well-functioning self-evaluation system? In particular was adequate supervision and implementation support provided and a MTR undertaken in a timely manner, and portfolio performance monitored on a continuous basis? (iii) Did IFAD exercise its developmental and fiduciary responsibilities adequately, ensuring that projects had sound financial management systems, audit reports were submitted in a timely manner, the required

provisions in the financing agreements were fully met, etc.? (iv) What support did the IFAD country office (where applicable) provide to the country programme and individual operations? and (v) Did IFAD pay adequate attention to further cooperation and dialogue with the United Nations Rome-based agencies?

- **Government:** (i) Did the government ensure that a baseline survey was done in a timely manner and that M&E systems were properly established and functioning? (ii) How were periodic progress reports used and was the PCR provided in a timely manner and of the required quality? (iii) Were counterpart resources (funds and staffing) provided in line with the agreement at design stage? (iv) Were audit reports done and submitted as needed? (v) Were the flow of funds and procurement procedures suitable for ensuring timely implementation? (vi) Did the government have the required capacity at all levels to implement the project as per schedule?

The below section clarifies the use of evaluation criteria by different types of IOE evaluations.

Corporate-level evaluations

These are the most complex and challenging evaluations done by IOE, covering various types of IFAD policies, strategies, operations, and business processes. Therefore, depending on the nature and scope of the evaluation, the specific evaluation criteria used will be defined in the approach paper, as drawn from the criteria shown in table 4. There may be instances when additional criteria, not included in this chapter, are also used in certain CLEs.

Country strategy and programme evaluations

These evaluations will be based on three inter-related components to be assessed: (i) portfolio-level analysis of IFAD-funded projects included in the CSPE scope; (ii) non-lending activities – policy

dialogue, knowledge management, partnership-building; and (iii) the performance of partners (i.e. IFAD and the government) in managing the country programme, including process aspects. Based on the composite assessment of these three components, CSPEs will provide an aggregate rating for the IFAD-government partnership in promoting rural transformation. The evaluation criteria, their definitions and key questions are described in chapter 6, on the methodology and process for CSPEs.

With regard to the assessment of the project portfolio, each project included will be assessed and rated for the following criteria: relevance, effectiveness, efficiency and sustainability of benefits, project performance (as an average of the four aforementioned criteria), rural poverty impact, gender equality and women's empowerment, innovation and scaling up, environment and natural resources management, and adaptation of smallholders to climate change, as well as IFAD's performance, and government's performance as a partner. If a criterion is not applied in a particular CSPE, the reasons for this should be explicitly stated in the approach paper.

Based on these ratings – excluding IFAD's and government's performance – an overall project portfolio achievement rating will also be determined. However, it should be noted that – for ongoing projects assessed as part of CSPEs, it might not be possible to assess and rate all the criteria. This is because some ongoing projects might have been under implementation for only a short while at the time of the CSPE, and therefore it was not possible to assess their performance across all evaluation criteria adopted by IOE (e.g. in terms of effectiveness or impact). The criteria to be assessed in depth will be set out in the approach paper for CSPEs.

Each non-lending activity will be rated for relevance and effectiveness, but only one rating is to be provided for policy dialogue, knowledge management and partnership-building. Based on these three individual ratings, CSPEs will also generate a composite rating for non-lending activities.

Project evaluations and validations: PCRVs, PPEs and impact evaluations

These evaluations are done for completed operations, and therefore will cover all evaluation criteria to be rated as shown in table 4. There might be instances when not all evaluation criteria can be assessed or rated. For example, IOE might not be able to assess or rate “adaptation to climate change” in a project that mainly deals with microfinance for microenterprise development, or projects that were designed a decade ago that did not explicitly emphasize climate change. In such cases, the evaluation report will allocate a “not-applicable” rating to such criteria.

Evaluation synthesis reports

These IOE knowledge products primarily focus on summarizing lessons learned, and good or less-good practices. The analysis in a synthesis report will include a review of four evaluation criteria: relevance, effectiveness, efficiency and sustainability of benefits. However, these criteria will not be rated. The approach paper for an evaluation synthesis report will clarify if any other criteria is to be covered, together with the rationale for doing so.

The rating system

Starting in 2005, in line with the practice adopted in many other IFIs and United Nations organizations, IOE moved from a four-point to a six-point rating system⁵⁴ that allows for a more nuanced assessment of project results.

In addition to reporting on performance based on the six-point rating scale, in 2007 IOE introduced the broad categories of “satisfactory” and “unsatisfactory” for reporting on performance across the various evaluation criteria (see table 5).⁵⁵

As a general rule, evaluators will exercise their own judgement in assigning ratings, supporting their arguments with evidence. In this regard and based on the evidence available, evaluators are encouraged to make full use of all the points on the rating scale, rather than merely being satisfied with

54 Where 6 is the highest score and 1 the lowest score.

55 The Annual Review of Development Effectiveness produced by the Independent Evaluation Group of the World Bank uses a similar system of categorization.

assigning ratings in the middle of the range. It is essential that the narrative on findings is consistent with the rating assigned. All ratings (except project performance) will be a round number, with no decimal points. Consistent with most other evaluation offices (except the Asian Development Bank) and to keep the system simple, no weights

will be assigned to ratings when determining a final rating for project performance or overall project achievement. Evaluators are encouraged to consult the detailed score descriptors of the PCR guidelines. In addition, IOE has developed the following guidance to support evaluators to assign ratings on each evaluation criteria.

TABLE 5 **Rating system**

Score	Assessment	Category
6	Highly satisfactory	SATISFACTORY
5	Satisfactory	
4	Moderately satisfactory	
3	Moderately unsatisfactory	UNSATISFACTORY
2	Unsatisfactory	
1	Highly unsatisfactory	

TABLE 6 **Score descriptors**

Rating scale	Score descriptor
Highly satisfactory (6)	Under the concerned criterion, the activity (project, programme, non-lending, etc.) achieved or surpassed all main targets, objectives, expectations, results (or impacts) and could be considered as a model within its project typology.
Satisfactory (5)	Under the concerned criterion, the activity achieved almost all (indicatively, over 80-95 per cent) of the main targets, objectives, expectations, results (or impacts).
Moderately satisfactory (4)	Under the concerned criterion, the activity achieved the majority (indicatively, 60 to 80 per cent) of the targets, objectives, expectations, results or impacts. However, a significant part of these was not achieved.
Moderately unsatisfactory (3)	Under the concerned criterion, the activity did not achieve its main targets, (indicatively, less than 60 per cent) objectives, expectations, results or impacts.
Unsatisfactory (2)	Under the concerned criterion, the activity achieved only a minority of its targets, objectives, expectations, results or impacts.
Highly unsatisfactory (1)	Under the concerned criterion, the activity (project, programme, non-lending, etc.) achieved almost none of its targets, objectives, expectations, results or impacts.

4. Project completion report validation

Background

The 2010 Peer Review of IFAD's Office of Evaluation and Evaluation Function by the ECG recommended that IOE transform its approach to project evaluations. In particular, it recommended that IOE discontinue the way project evaluations were done in the past, and instead introduce project completion report validations (PCRVs) and project performance assessments (PPAs), subsequently renamed project performance evaluations (PPE).

The rationale for this recommendation was to enhance IOE's efficiency and enable the division to cover and assess the entire portfolio of closed projects funded by IFAD in any given year. Furthermore, undertaking PCRVs and PPEs would align IFAD's independent evaluation function even more closely with similar functions in other IFIs, which all also validate PCRs and conduct PPEs.

Both PCRVs and PPEs are enshrined in the IFAD Evaluation Policy (2011). In addition to PPEs, IOE also undertakes another type of project evaluation, namely impact evaluations. Chapter 5 contains guidelines on PPEs, while chapter 8 outlines the methodology and process for undertaking impact evaluations.

IOE developed guidelines on PCRVs in 2010, which were subsequently updated in 2013. This chapter builds on IOE's experience in conducting PCRVs, and supersedes past guidelines issued by IOE for the purpose.

Definition

A PCRV is a desk-based validation of a PCR, which contains an assessment of the results and impact of an IFAD-financed project or programme, including ratings for each evaluation criterion covered in

the validation. PCRs are prepared by recipient governments, following guidelines developed by IFAD Management within six months after project completion date.

Purpose and use

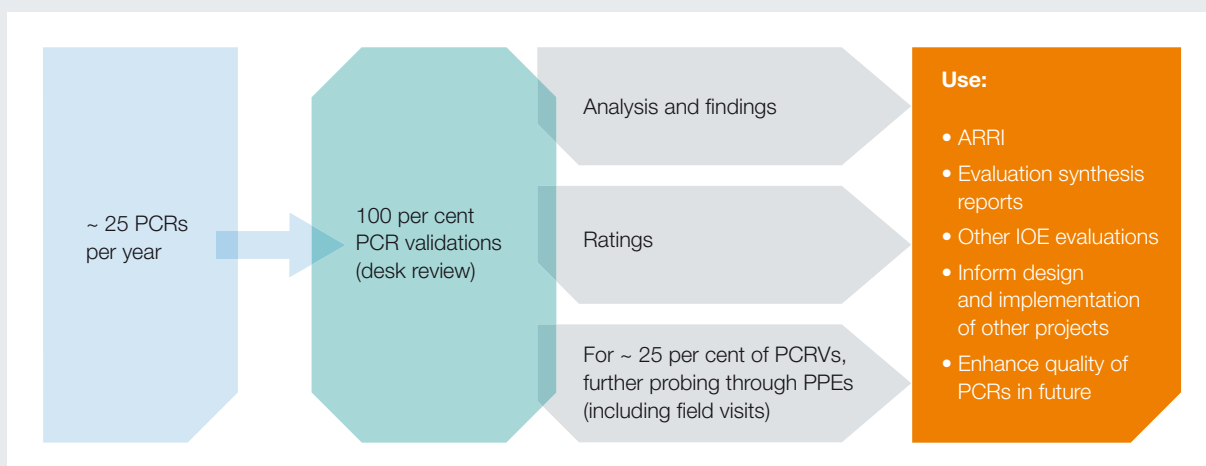
In line with the main principles of the IFAD Evaluation Policy, the overall purpose of PCRVs is to promote accountability and learning for better development effectiveness. PCRVs are indeed important building blocks for learning, and the balance between ratings and learning must be kept in mind to avoid that an emphasis on ratings results in missed opportunities for learning.

PCRVs include findings on performance and ratings across the evaluation criteria adopted by IOE. PCRVs contribute to strengthening a key component (i.e. the PCR) of a project's self-evaluation system, as PCRVs also include an overall assessment of the quality of PCRs. PCRVs do not include recommendations and therefore no IFAD Management's response is required.

PCRVs feed into and inform the design of new projects as well as the implementation of ongoing projects. As such, they contribute to enhancing the quality of IFAD's operations in general. They also provide essential evidence for the preparation of key IOE evaluation products, such as the ARRI, evaluation synthesis reports, and other evaluations such as CSPEs and CLEs.

PCRVs are normally done by IOE staff without the assistance of consultants and IOE is fully responsible for the underlying process of preparing PCRVs as well as for the contents contained therein. On average, eleven staff days are allocated for a PCRV.

FIGURE 1 PCRVs, PPEs and their use



Selection

PCRVs are prepared for all completed operations in any given year. This allows IFAD's independent evaluation function to gain a picture of the performance of all completed operations for which a PCR is available in a timely manner. Based on past records, IOE prepares around 25-30 PCRVs per year, which corresponds to the number of projects in the IFAD project portfolio. However, the number of PCRVs may vary from year to year, depending on the number of operations that close and the PCRs available in any given year.

Objectives and methodology

The objectives of PCRVs are to: (i) assess the results of IFAD operations, including the robustness of ratings contained in the PCR⁵⁶ for each evaluation criteria assessed; (ii) validate findings and lessons; (iii) highlight possible findings, lessons and systemic issues of importance to project design and implementation; and (iv) assess the quality of the PCR process and final report, with the aim of identifying areas for further improvement of the PCR as a self-evaluation instrument.

This section provides guidance on the evaluation design and building blocks of PCRVs, on the evaluation criteria, and how to validate and assign ratings.

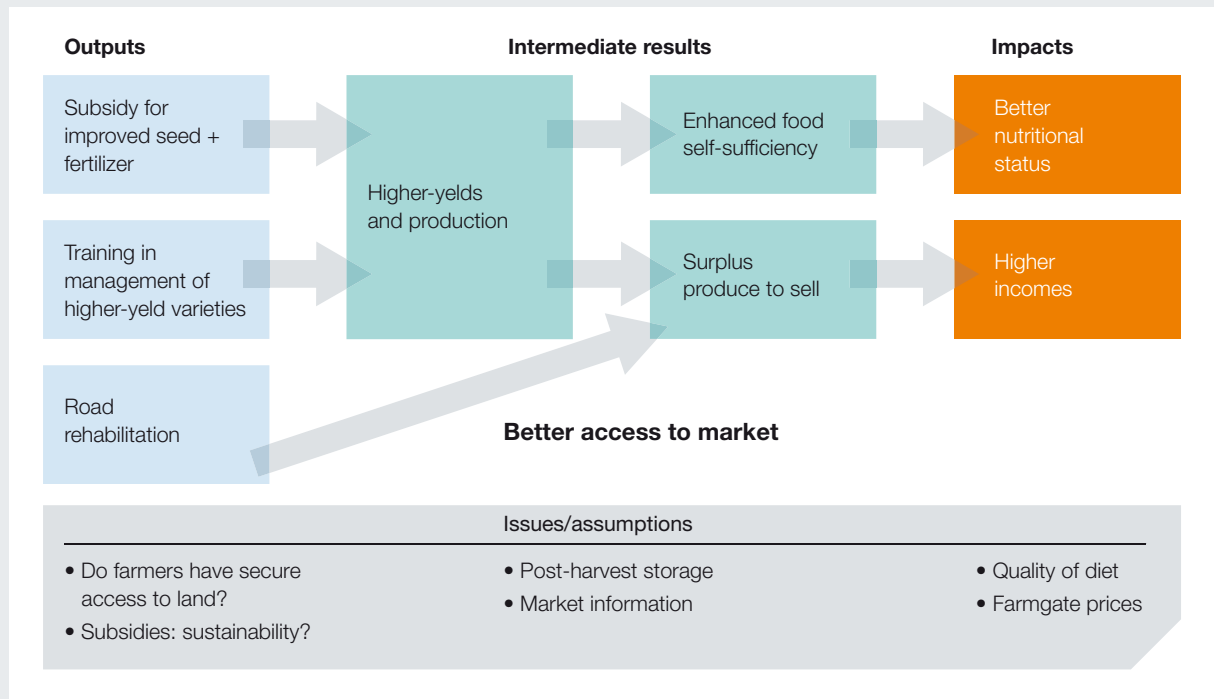
Design and building blocks of the PCRV

The basic element for PCRVs is an extensive documentary review, which will include a review of the relevant COSOP, all the relevant project documents and monitoring reports, as well as key strategic policy and strategy documents to assess the relevance criteria. The validation usually starts with clarifying and articulating the project conceptual framework (i.e. the theory of change, or the logical framework), if one was not explicitly included in design. To do so requires reviewing the final project design to understand the overall project objectives, the type of components the project seeks to support, the underlying assumptions, the cost structure, the institutional arrangements, and the overall strategy to achieve project objectives.

It may be helpful to draw up a simple project scheme along the lines of that shown in figure 2, which provides a simplified framework with a number of simple assumptions. However, a project framework

⁵⁶ This relates to the ratings assigned by PMD to each evaluation criteria in the PCR.

FIGURE 2 A simplified linear logical project scheme



may be more complex and the relationships between inputs, outputs, intermediate results and impact may not be linear. For the purpose of a PCRV, it is advisable to keep the scheme short and simple. If the project seems to lack a clear and coherent logical framework, this should be underlined in the PCRV.

It is essential to assess the lower levels of the results chain (i.e. outputs and activities), because this is the basis for explaining why the performance was as it was, including in terms of flow of funds and financial management. The latter is important for ensuring proper and timely project management and implementation, which is at the core of successful outcomes. Use of an intervention logic/theory of change that takes into account complexities arising from contextual factors and the interventions of other actors besides IFAD could help both in focusing evaluations better and in providing more of a basis for analysing factors related to observed outcomes. Assessing whether the initial project logical

framework was construed in a coherent manner and whether key assumptions were validated, or violated, is a key aspect of the assessment.

Based primarily on secondary information, PCRVs verify the logical consistency of the PCRs as well as the evidence base for the ratings assigned by Management to the various evaluation criteria.

PCRVs may face difficulties in verifying some information, especially relating to project impact. Although there is no “magic bullet” to solve the problem, evaluation literature recognizes a set of strategies for enhancing validity. These include:

- (i) Content analyses of PCRs; verifying logical consistency;
- (ii) Cross-checking and triangulating with other documents (e.g. MTR reports, impact studies and supervision reports), IFAD’s websites (e.g. x-Desk) and other sources (e.g. a World Bank report on an IFAD project it is cofinancing); cross-checking with independent evaluation

reports (when available), conducting interviews with relevant project staff including the current and past CPMs; and

- (iii) Adopting theory-based approaches; considering the conceptual model and logical framework of the project/programme, and verifying whether crucial assumptions hold up.

An increasing international emphasis on impact is posing challenges for all evaluations, also for IFAD. Providing sufficient evidence to be able to attribute impact to any given project is indeed methodologically challenging and a vast body of literature exists on methods and approaches to measure impact (see also chapter 8 on impact evaluations). Most require stronger M&E data than typically found in IFAD-funded projects, in particular older projects, and more resources than what are available in an average PCR; these therefore should be prudent in their conclusions with regard to impact.

Criteria and questions

PCRVs will follow the evaluation criteria outlined in this manual; each PCR will a priori validate and rate all evaluation criteria adopted by IOE, as per chapter 3. The final PCR document will provide a clear explanation of why any criteria are not assessed and/or not rated.

The Manual includes a set of questions to help evaluators form their final judgements. In particular, each PCR should use the core questions for each criterion, which may be seen in chapter 3 as a reference for the analysis. The outcome of the analysis will be used as a basis for rating the project's performance across each criterion. This will ensure consistency and comparability across evaluations done by IOE as well as enable the aggregation of ratings in the ARRI and other evaluations. Moreover, evaluators may use additional evaluation questions to further strengthen the analytic underpinning of the evaluation, especially to generate lessons learned (suggestions are included in annex II).

PCRVs will also assess the quality of the PCR, based on five inter-related criteria: (i) scope (e.g. whether the PCR has adhered to the IFAD guidelines for PCRs); (ii) data (in terms of the quality, etc.); (iii) lessons (e.g. the inclusion of lessons about project design and implementation); (iv) process (e.g. the use of a participatory approach during the PCR process); and (v) candour (in terms of objectivity in the narrative, and whether both positive and negative results are highlighted). Each of these criteria will be rated, and an aggregate rating of the quality of PCR will also be determined in the validation by IOE. A checklist for each of these criteria is provided in box 2 for ease of reference.

BOX 2 PCR checklist

Scope

- To what extent have all the chapters including relevant annexes foreseen in the PCR Guidelines been adequately covered?

Data

- To what extent is the PCR based on robust data?
- Are data collected adequate for their intended use and have their reliability been ascertained?
- If M&E and RIMS data were limited, were supplementary surveys and research carried out?
- Are data systematically analysed to provide information needs in a valid manner?

Lessons

- To what extent have lessons been drawn from a) project design and b) project implementation?
- Are lessons based on explicit evaluation findings?

Process

- To what extent has the PCR process been inclusive of all relevant stakeholders?

Candour

- To what extent is the narrative objective?
- To what extent have both positive and negative results been reported?
- To what extent is the rating coherent with the narrative?

Ratings

As mentioned above, PCRVs will assign a score to each evaluation criterion covered, based on the six-point rating scale adopted by IOE (where 6 is the highest score and 1 the lowest score). All criteria will be scored individually.

Once a PCR is produced by the government and accordingly submitted to IFAD, PMD assigns ratings (self-assessment) for each evaluation criterion covered. For accountability purposes, both the PMD ratings and independent PCRv ratings assigned by IOE should be included in a table in a dedicated annex at the end of the PCRv report. Differences between the PCR and PCRv ratings are also shown criterion-by-criterion, together with an average difference (or disconnect). Should there be a disconnect between PMD and IOE ratings, the PCRv will provide the factors behind a difference in evaluation judgement.

Independent evaluation ratings are important because they will be used as a basis for performance assessment in CLEs, CSPEs, synthesis reports and the ARRI. IFAD Management will also use IOE ratings to report against selected criteria and indicators in the corporate results measurement framework and the annual Report on IFAD's Development Effectiveness (RIDE).

As a general rule, the evaluator will exercise his/her own judgment in assigning ratings, supporting his/her arguments with all information collected and analysed during the validation process. It is important to note that there may be circumstances where, because of limited information and data, IOE is unable to rate a specific evaluation criteria and make a sound independent evaluation assessment. This may be the case even in circumstances where the PCR contains a rating for the concerned evaluation criteria. In such situations, IOE shall assign a "not applicable (N/A)" score, and the PCRv should clearly explain the reasons for this choice.

Process

As per the financing agreements between IFAD and recipient governments, PCRs are to be submitted to IFAD within six months of project completion.⁵⁷ PMD is responsible for transmitting the PCR to IOE; once that and the corresponding ratings are received, the PCRv is carried out immediately.

The process from drafting to finalization of a PCRv consists of four main phases, described below. More detailed guidance is contained in the Evaluation Process Guidelines.

Phase 1: Preparation

The first phase is to identify and collect relevant documentation, data and information. These include the final project design document, the President's Report summarizing the project's design submitted to the Board for approval, baseline survey, MTR report, supervision report, data from the RIMS, project status report, impact study and PCR. Additional supporting documents include the COSOP, any special studies available, and background documentation on the country and project area. Project, country strategy and programme and CLEs will be consulted, whenever relevant. PMD will share with IOE any relevant documents and data.

Phase 2: Desk review, analysis and drafting

The desk review is conducted taking into account the suggestions contained in the methodological section above and in chapters 2 and 3. In terms of drafting, as a general rule, the PCRv is prepared by an IOE staff member. It is his/her responsibility to review the documentation involved, and to produce a draft of the PCRv report within the time frame agreed with his/her supervisor. All PCRvs are only produced in English to minimize costs.

Phase 3: Review process

All draft PCRvs will undergo a rigorous internal review within IOE by the immediate supervisor of the lead evaluator assigned to do the PCRv. The internal review will follow IOE's internal peer review

57 PCRs as submitted by governments to IFAD do not include ratings.

guidelines,⁵⁸ which may be updated from time to time. Comments on the draft PCRV may be provided in writing or verbally, or both. The review will focus primarily on the overall rigour of analysis, robustness of the evidence trail and justification for ratings.

Following the internal review, the revised draft PCRV will be forwarded to the IOE Deputy Director for further review and comments. His/her role is to review the overall quality of the deliverable, promote consistency between analysis and ratings, as well as ensure that inter-evaluator variability is minimized across IOE products.

Once the Deputy Director's comments have been included, the revised draft PCRV is sent to PMD for review and comments, in line with the internal communication protocol established by IOE (see Evaluation Process Guidelines). The relevant regional division of PMD will be requested to provide feedback within 15 working days from submission of the PCRV by IOE.

Phase 4: Finalization, dissemination and follow-up

Upon receipt of PMD's comments, the draft is finalized and an audit trail produced, and the audit trail and final version are cleared by the lead evaluator's immediate supervisor. If any ratings have been changed by IOE based on PMD's feedback, the IOE Deputy Director will review the revised PCRV before it is finalized.

If no feedback is received from PMD on the draft PCRV, IOE will send a reminder to PMD, and if no comments are received by the extended deadline, IOE will consider the PCRV final and inform PMD accordingly.

The final version of the PCRV and the audit trail are transmitted to PMD based on the established communication protocol, and at this stage are considered final. To ensure maximum learning from the exercise, a meeting may be held with concerned PMD and other staff to share and discuss the findings and their implications for future operations or policies. Furthermore, the PCRV is forwarded to the Evaluation Communication Unit of IOE for dissemination as described in chapter 10.

Reports

The PCRV should be concise (maximum 9-11 pages, excluding annexes), and reviewers should limit their presentation to information essential both for the analysis and for the validation of ratings. Being a validation, and as mentioned earlier, it does not include recommendations.

BOX 3 Outline for PCRV

- I. Basic project data
- II. Project outline
- III. Review findings
 - A. Rural poverty impact
 - B. Core criteria
 - C. Other performance criteria
 - D. Overall project achievement
 - E. Performance of partners
- IV. Assessment of PCR quality
- V. Final remarks and lessons learned

5. Project performance evaluation

Background

IOE introduced project performance assessments (PPAs) following the 2010 Peer Review of IFAD's Office of Evaluation and Evaluation Function by the ECG, which recommended that IOE discontinue the way project evaluations were done in the past, and introduce PCRVs (see chapter 4) and PPAs, subsequently renamed project performance evaluations (PPEs). This means that IOE is now aligned with similar functions in other IFIs, which also conduct project evaluations similar to PPEs.

IOE developed guidelines on PPAs in 2010, which were updated in 2013. This chapter builds on IOE's experience in conducting PPEs, and supersedes past guidelines issued by IOE for the purpose. Preparation of PPEs is also embedded in IFAD's Evaluation Policy, which prescribes that IOE shall undertake project performance assessment (now project performance evaluation) for a selected number of projects previously exposed to PCR validation by IOE.

Definition

In line with the main principles of the IFAD Evaluation Policy, the overall purpose of a PPE is to promote accountability and learning for better development effectiveness.

A PPE is a project evaluation with a limited scope and resources. It is generally based on the PCRv, but the analysis is more complete, being based on additional information and data collection by IOE at the country level through a short mission. It also includes ratings by IOE for each evaluation criterion covered.

PPEs include findings on performance and ratings across the evaluation criteria adopted by IOE, and evidence-based conclusions and recommendations. In addition to providing ratings for each evaluation criteria covered, PPEs focus on identifying the proximate causes of good or less good performance. PPEs contain recommendations, addressed respectively to IFAD Management and the concerned governments.

PPEs feed into and inform the design of new projects as well as the implementation of ongoing projects. As such, they contribute to enhance the quality of IFAD's operations in general and constitute an important building block for the preparation of key IOE evaluation products, such as the ARRI, evaluation synthesis reports, and other evaluations such as CSPEs and CLEs.

PPEs are led by IOE staff, and generally require some input from consultant(s), but on a more modest scale than past project evaluations. PPEs normally require 35 staff days to be undertaken, and one or two consultants. In line with the IFAD Evaluation Policy, IOE is fully responsible for the underlying process of preparing PPEs as well as for their contents.

Selection

There are a number of factors that IOE considers in selecting projects to cover through PPEs, but as a general rule, and in order to make informed and strategic choices, IOE will undertake PPEs on projects where a PCRv was carried out the previous year. Other factors that are considered include: (i) the availability of PCRv; (ii) information gaps in PCRv identified through the PCRv; (iii) projects of strategic

relevance that offer enhanced opportunities for learning; (iv) a need to build evidence for forthcoming CLEs, CSPEs or evaluation synthesis reports; (v) provide a regional balance of IOE's evaluation programme; (vi) projects with innovative features that merit deeper analysis and documentation; and (vii) PCR ratings that are inconsistent with narrative. IOE generally undertakes about 10 PPEs in a given year, but this may vary, taking into account IOE divisional priorities and the resources available.

Objectives and methodology

This section provides guidance on the evaluation design and building blocks of PPEs, on the evaluation criteria, and on how to validate and assign ratings.

The objectives of a PPE are to: (i) assess the results of the project under consideration; (ii) generate findings and recommendations for the design and implementation of ongoing and future operations in the country involved; and (iii) identify issues of corporate, operational or strategic interest that merit further evaluative work.

Evaluation design and building blocks

Good practice recommends that evaluators start by analysing the context for an evaluation, and then develop an evaluation strategy appropriate for the situation; this also applies to PPEs.

The two basic building blocks for PPEs are the PCR (see chapter 4) and a field visit, which will include discussions with multiple stakeholders (e.g. government officials, project staff, beneficiaries, cofinanciers) and visits to selected project sites.

In preparing for the PPE mission, simple data collection exercises may be conducted by a local consultant, when sufficient resources and time are available. These could include mini-surveys (see table 2, chapter 2 for more information) and/or spot-checks of typical data ranges or patterns described in the PCR by means of case studies.

Data collection in PPE missions often provide useful reference and insights in addition to data from desk reviews. Sites for field visits are selected by IOE

in consultation with the government concerned and efforts should be made to organize field visits that cover different sites from the ones covered in the PCR. Government staff may accompany the PPE mission on these visits as observers.

While missions provide PPEs with additional evidence and information, they also face challenges. In particular, country visits are often conducted under time and budgetary constraints, and as a consequence it may be difficult to collect new primary data related to impact.

Therefore it is important to recall the purpose of the PPEs, and note that they are expected primarily to fill important information gaps; they are not expected to investigate all activities financed under projects/programmes or to undertake in-depth impact assessments.

Through independent evaluations IFAD aims to establish to what extent certain results are due to a development intervention rather than to external factors. This requires extensive data on household or community-level impact indicators, of sufficient quality and methodological rigour. This is often a challenge and PPEs may help address such weaknesses by:

- (i) Following the logical chain of the project, identifying key hypotheses and reassessing the plausibility chain;
- (ii) Conducting interviews with non-beneficiaries sharing key characteristics (e.g. socio-economic status, livelihood, farming system), which would give the mission an idea of what would have happened without the project (counterfactual);
- (iii) Concentrating field visits on components or criteria in need of verification, or that are innovative or problematic;
- (iv) Undertaking case studies⁵⁹ in a number of sites, for example by visiting areas not adequately covered by the PCR, MTR or supervision missions;
- (v) Concentrating field visits on interaction with the beneficiaries and their communities. Instruments for such interaction (individual interviews, focus groups, key informants with

59 Stake, R.E. 1995. *The Art of Case Study Research*; Hancock, D. and B. Algozzine. 2011. *Doing Case Study Research: A practical Guide for Beginning Researchers*, Second Edition.

thematic knowledge, participant observations, participatory rural assessment techniques) are designed to fill information gaps highlighted in the PCR; and

- (iv) Considering “rival explanations”, that is, alternative factors that may have generated similar results, as a way of exploring attribution issues.

Evaluation criteria and questions

PPEs will follow the evaluation criteria outlined in this Manual; each PPE will evaluate and rate all evaluation criteria adopted by IOE. The approach paper for PPEs will provide a clear explanation of why any criteria have been dropped (i.e. not assessed and/or not rated), and the reasons for focus on specific criteria.

The Manual includes a set of questions to help evaluators form their final judgements. In particular, each PPE must include an analysis across a set of core questions for each criterion, which is found in chapter 3. The outcome of the analysis will be used as a basis for rating the project’s performance across each criterion. This will ensure consistency and comparability across evaluations done by IOE as well as enable the aggregation of ratings in the ARRI and other evaluations. Moreover, evaluators may use additional evaluation questions to further strengthen the analytic underpinning of the evaluation, especially to generate lessons learned (see annex II for examples).

PPEs will also assess the quality of the PCR, based on four inter-related criteria: (i) scope (e.g. whether the PCR has adhered to IFAD guidelines for PCRs); (ii) quality (in terms of data, methods and participatory processes followed); (iii) inclusion of lessons about design and implementation; and (iv) candour (in terms of objectivity in the narrative, and whether both positive and negative results are highlighted). Each of these criteria will be rated, and an aggregate rating of the quality of the PCR will also be determined by IOE.

Ratings

As mentioned above, PPEs will assign a score to each evaluation criterion covered, based on the six-point rating scale adopted by IOE (where 6 is the highest score and 1 the lowest score). All criteria will be scored individually.

Once the PCRs are produced by governments and accordingly submitted to IFAD, PMD assigns ratings (self-assessment) for each evaluation criterion covered. For accountability purposes, the PMD ratings, the independent PCR ratings, and the new PPE rating assigned by IOE should be included in a table in a dedicated annex at the end of the PPE reports. Differences between the PCR and PPE ratings are also shown criterion by criterion, together with an average difference (or disconnect) between the PMD rating and the new PPE rating. As and where there might be a disconnect between a PMD and IOE rating, the PPE should underline clearly the explanatory factors that are at the basis for a difference in evaluation judgement.

Independent evaluation ratings are important because they will be used as a basis for performance assessment, for example in CLEs, CSPEs, synthesis reports and the ARRI. In this regard, it is important to note that the PPE ratings (and not the PCR ratings) will be considered the final ratings by IOE of an IFAD-supported project, since PPEs benefit from in-country work and are based on more detailed analysis of data and information. IFAD Management will also use IOE ratings to report against selected criteria and indicators in the corporate results measurement framework and the RIDE.

As a general rule, the evaluator will exercise his/her own judgment in assigning ratings, supporting his/her arguments with all the information collected and analysed during the evaluation process. It is important to note that there might be circumstances when IOE is unable to provide a rating to specific evaluation criteria, given limited information and data available to make a sound independent evaluation assessment. This might be the case even in circumstances when the PCR contains a rating for the concerned evaluation criteria. In such

situations, evaluators shall assign a “not applicable (N/A)” score, and the PPE should clearly explain the reasons for this score.

Process

As per the financing agreement between IFAD and recipient governments, PCRs are to be submitted to IFAD within six months of project completion.⁶⁰ PPEs are generally conducted one to three years after project completion. While such time frames may pose some challenges in tracing former project staff and other key informants, it does allow for the opportunity to better discern impacts and sustainability of the operation.

PPEs will take a maximum of six months to carry out from start to finish. This includes, inter alia, the preparatory work, field mission, analysis and report writing, internal review process within IOE, comments by Management and the government, preparation of IFAD Management’s Response on the final report, and publication and dissemination.

The PPE process is guided by the Evaluation Process Guidelines. Once IOE has informed the concerned government and PMD of the intention to conduct a PPE, the process will follow the six main phases below.

Phase 1: Preparation

At the outset of the process, IOE formally informs PMD and the government of the PPE, specifying the main objectives, process and tentative timeline for the evaluation. It is important that the timeline take into account the availability of the CPM, who is required to participate in person in the PPE wrap-up meeting in the country, at the end of the field work by IOE. The CPM and IFAD country office (where applicable) assist IOE by providing contact details of key government officials.

The lead evaluator will use the PCR as the starting point for identifying issues for further investigation and for preparing the draft approach paper for the PPE, using the IOE template. PPEs follow an objectives-based evaluation, and apply all

evaluation criteria adopted by IOE (see chapter 3). The preparation of the approach paper is critical, as it provides an opportunity for customizing the PPE, depending on its specific country and project circumstances. The approach paper will first be reviewed by the lead evaluator’s immediate supervisor. Once his/her comments are added, the revised approach paper is shared with the IOE Deputy Director for review and comments.

Following the inclusion of the IOE Deputy Director’s comments, the revised draft approach paper is shared with PMD and the concerned government for review and comments, in line with the established communication protocol. PMD and the government will be given ten working days to provide their feedback. Thereafter, IOE will revise the approach paper, which will be considered final.

The lead evaluator then contacts PMD and government authorities to prepare a mission schedule. IOE informs the relevant IFAD country office of the evaluation and establishes direct contact with concerned government authorities to agree on meetings and develop a programme for the field visit. The lead evaluator maintains close contact with the CPM involved and the IFAD country office; however, all communication with governments regarding the PPE will be sent directly from IOE.

While the CPM and the IFAD country office may provide suggestions for the meeting programme, they should not be asked to prepare the programme of meetings nor determine the field visits to be undertaken by the PPE mission, which is IOE’s responsibility. Preparing the PPE mission well is essential and may require telephone interviews with key stakeholders prior to the mission, as well as identification of former project staff and partners.

If a consultant is recruited to support the IOE lead evaluator, briefings should be held at IFAD headquarters, if possible. This includes meeting(s) with the IOE Director and Deputy Director, the regional division director and CPM, and other IFAD staff concerned (e.g. technical advisors in the Policy and Technical Advisory Division).

60 PCRs as submitted by governments to IFAD do not include ratings.

Phase 2: Desk review

This phase follows the preparation of the draft approach paper. The desk review informs further work, which identifies a list of issues and questions to be further analysed during the field mission, as well as the availability of data and any information gaps that need to be filled in by the evaluation.

Phase 3: Field mission and data collection

The PPE field mission will last from seven to ten days in the country, including meetings in the capital and the project area to collect primary data through qualitative and quantitative techniques. It is important that a meeting be held at the outset of the field work with the main government counterpart in the capital. Similarly, a wrap-up meeting must be held with the government counterpart and other concerned authorities at the end of the mission. The IOE lead evaluator takes part in the wrap up meeting, together with the CPM and IFAD county office staff. The aim of the wrap up meeting is to share the emerging findings from the PPE and provide a briefing about the next steps. No aide-mémoire is prepared, but IOE gives a PowerPoint presentation summarizing the emerging findings of the field mission.

Phase 4: Analysis

Once the field mission is over, the lead evaluator, together with the consultant(s), will analyse the data and information collected from the field and feedback from key informants at the country level, including beneficiaries. This critical phase is based on triangulation. IOE will form its evaluation judgments based on the evidence collected throughout the process, including the review of critical project documents.

Phase 5: Preparation of draft final report and review process

After the analysis, the PPE report is prepared under the overall responsibility of the IOE lead evaluator. S/he will receive written inputs from the consultant(s) for selected areas of the main report, depending on the specialization of the consultant(s) hired. As for

PCRVs and all IOE products, the IOE lead evaluator is fully responsible for preparing the final report and its contents. The indicative table of contents of the PPE is shown in box 4.

The same process will be followed for the internal reviews of PPEs as for PCRVs. See the previous chapter.

After IOE comments are duly incorporated, the draft is forwarded to PMD and the government for review and comments as described in the Process Guidelines. Comments provided should highlight: (i) factual inaccuracies; and (ii) areas of difference of opinions, with supporting evidence that may be helpful in finalizing the evaluation report. If comments are not received by the set deadline, IOE will write to PMD and the government granting an extension of five working days for their comments. In this communication, they will be informed that should no comments be received by the deadline, IOE will consider the report final and proceed with its publication and dissemination.

Phase 6: Finalization, dissemination and follow-up

Upon receipt of PMD's and the government's comments, the lead evaluator finalizes the draft report and prepares an audit trail within five working days. The audit trail and final version are reviewed by the evaluator's immediate supervisor. The IOE Deputy Director will review the revised PPE report should there be any changes to ratings, as compared to the draft final report. After including the Deputy Director's additional comments, the revised PPE is considered final and forwarded to PMD and the government for information, together with the audit trail, in line with IOE's communication protocol. A learning event may be organized within IFAD at this stage to share and discuss findings and their possible implications for future assistance in the same country or thematic area. On an exceptional basis, and resources permitting, a learning event may also be organized at the country level.

No agreement at completion point (ACP) is prepared for a PPE. However, IFAD Management will prepare a written Management's response on

the PPE report, which will be included in the final report at the time of publication by IOE. Under normal circumstances, Management's response should provide IFAD Management's views on the main findings and their agreement or otherwise on the recommendations, as well as an overall appreciation of the evaluation process. Moreover, IFAD Management will report in the PRISMA on the implementation status and follow-up to PPE recommendations made to IFAD (and not on those directed to the government) as it does for the ARRI, CLEs, CSPEs and IOE comments on the RIDE.

Once finalized, the PPE is formatted and submitted to the Evaluation Communication Unit of IOE for editorial quality assurance, web publishing and dissemination (see chapter 10).

Reports

The PPE should be concise (maximum 30 pages excluding annexes) and contain both descriptive, analytical and evaluative material. Box 4 shows an indicative contents page for a PPE report.

PPEs are written in the IFAD official language (Arabic, English, French or Spanish) of the recipient country. The executive summary will be translated into English if the language used for the PPA is Arabic, French or Spanish. Both the original language and English versions of the executive summary and IFAD Management's response will be included in the final report.

BOX 4 Example of contents page for PPE reports

Executive summary

IFAD Management's response

I. Evaluation objectives, methodology and process

II. The project

A. Project context

B. Project implementation

III. Main evaluation findings

A. Rural poverty impact

B. Core criteria

C. Other performance criteria

D. Overall project achievement

E. Performance of partners

F. Assessment of the quality of the PCR

IV. Conclusions and recommendations

A. Conclusions

B. Recommendations

V. Annexes (including table of PMD and PPE ratings)



6. Country strategy and programme evaluation

Background

This chapter outlines the IOE methodology for CSPEs, which were previously called country programme evaluations (CPE). This name change will demonstrate that the overall scope and terminology has been aligned to GPSs for country-level evaluations, as adopted by the ECG. Given that investment projects are one of the main building blocks of IFAD-supported country programmes, this chapter should be read in conjunction with chapters 4 and 5, which are devoted to the methodology for project evaluations, as well as with chapters 1-3, dedicated to methodological fundamentals. Usually conducted before IFAD and the concerned government prepare a new results-based COSOP,⁶¹ CSPEs are expected to provide an overarching assessment of the performance of the country strategy and programme funded by IFAD. This includes IFAD-

funded lending operations, non-lending activities and country programme and COSOP management processes in a given country. CSPEs are also intended to provide guidance for the preparation of future COSOPs.

The proposed methodology takes into account IOE's experience in designing and implementing CSPEs.⁶² It also draws upon the experiences of other bilateral and multilateral development organizations⁶³ as well as the latest ideas in the international community of practice on development evaluation.⁶⁴

The CSPE methodology builds on the definitions contained in the results-based COSOPs guidelines adopted by the Board in September 2006.⁶⁵ Since COSOPs are documents jointly owned by IFAD and the government, CSPEs address their distinct accountabilities and reciprocal obligations in COSOP design and implementation and are thus an important instrument for mutual accountability.

61 As per the Executive Board's decision, if a CSPE has been undertaken by IOE, its agreement at completion point (ACP) must be included as an annex to the COSOP submitted for consideration by the Board. This allows the Board to assess whether the CSPE findings and recommendations have been adequately included in the COSOP. In addition, when a COSOP is presented to the Executive Board and a CSPE has been conducted prior to that COSOP, IOE will also present the CSPE to the Executive Board.

62 IOE undertook its first country portfolio evaluation in 1992, which focused on assessing the results of IFAD-funded projects and programmes. Country programme evaluations (CPEs) were introduced in mid-1999 and were given the broader remit of assessing the results of both IFAD-funded projects and programmes and, albeit in less detail, non-lending activities.

63 For example, the Guidelines for the Preparation of Country Assistance Program Evaluation Reports of Asian Development Bank (February 2006), the Country Assistance Methodology of the World Bank (2004) and the Good Practice Standard for MDB Country Strategy and Programme Evaluation (June 2008).

64 The ECG Good Practice Standards for Country Strategy and Program Evaluation stresses five strong benefits of CSPEs; they can: (i) identify and assess broad and long-term issues and concerns better than other forms of evaluation; (ii) provide valuable information about the country strategy process, whether project selection was based on merit, impact of non-project forms of assistance, aggregating results of activities across all sectors and providing input into, and strengthening, subsequent country strategies; (iii) identify overall programme and project delivery weaknesses, institutional difficulties, capacity utilization constraints, borrower's acceptance, commitment and compliance to conditions and impact of other aid agencies and external factors; (iv) provide a framework for rating overall performance in meeting development goals and objectives, and better assess impact and sustainability issues for long-term aid effectiveness; and (v) provide a valuable instrument for improving aid coordination among institutions and bilateral agencies and for the broader participation goal of increasing the role of national and local governments, civil society and the private sector in the developmental process.

65 EB 2006/88/R.4, approved by the Board during its 88th session in September 2006.

Definition

A CSPE is an evaluation of the results of partnership between IFAD and the concerned government for reducing rural poverty and promoting rural transformation.⁶⁶

Purpose

CSPEs generate an overall appreciation of the partnership between IFAD and the concerned government in reducing rural poverty. CSPEs are thus undertaken for both accountability and learning purposes, and to strengthen IFAD's development effectiveness.

Selection and timing

Careful consideration is given to selecting countries for CSPEs. In order to ensure full transparency, CSPEs are selected in accordance with the selectivity framework (see annex V). Factors such as the link to COSOPs, the regional and country coverage of CSPEs, the size of the portfolio in terms of total investments and number of operations, Debt Sustainability Framework classifications and lending terms are considered when making the selection.

In terms of timing, CSPEs are normally conducted before IFAD and the government concerned prepare a new results-based COSOP, but the country's budget and policy processes should also be considered. CSPEs are also timed so that the results may inform government decisions on the use of external assistance. CSPEs would also be undertaken in those cases where no past COSOP is available or future COSOP envisaged.

Objectives, scope and approach

Objectives

CSPEs have two main objectives: (i) assess the results and performance of the IFAD-financed strategy and programme; and (ii) generate findings and recommendations for the future partnership between IFAD and the concerned country for

enhanced development effectiveness and rural poverty eradication.

Scope

CSPEs will cover the full range of IFAD support to a country, including lending and non-lending activities (knowledge management, partnership-building and policy dialogue), including grants, South-South and Triangular Cooperation and reimbursable technical assistance as well as country programme and COSOP management processes. Specific attention will be devoted to assessing strategic dimensions of the IFAD-government partnership in promoting rural transformation.

CSPEs will generally cover a ten-year period of IFAD's engagement in the concerned country and include an assessment of all closed operations and other activities (e.g. non-lending activities) in the period. However, this time frame would need to be reconsidered if a previous CSPE (or CPE, as it was known in the past) had been conducted by IOE in the same country. In such instances, the evaluation period will cover all activities following the completion of the previous CSPE/CPE, including operations that were ongoing at the time. In any case, the exact period of coverage of a CSPE and activities to be evaluated will be established when preparing the evaluation approach paper.

CSPEs focus closely on evaluating the performance of the programme in a partnership between IFAD and the government, rather than on the borrowing country's general agriculture and rural development efforts at large. While IFAD-funded projects represent only one, albeit important, segment of government actions in the sector, an understanding of the latter is needed in order to appraise the contribution of the IFAD-government partnership to development effectiveness in the sector.

Approach

IOE's approach to CSPEs is broadly consistent with the main principles contained in the latest GPS for CSPEs (see the ECG document, 2008) and with the broader guidance provided in this Manual.

66 While other partners such as international organizations, non-governmental and civil society organizations intervene as well, the overall cooperation agreement is established between IFAD and the government.

The main unit of analysis of CSPEs is the country strategy and programme. CSPEs are designed to evaluate what has been achieved in a particular country, while also supporting forward-looking operational improvement and strategic guidance for the future partnership. Depending on the circumstances, CSPEs may also reveal broader issues of interest that may have implications beyond the individual country programme being evaluated.

Therefore, CSPEs will have a strategic focus building on the assessment of lending and non-lending activities. In particular, while recognizing that IFAD's assistance represents only a small segment of government actions in the agriculture sector, CSPEs will analyse wider issues related to IFAD-government partnership, such as IFAD's strategic positioning in the country in relation to government priorities and the work of other development partners, its role in promoting institutional and policy transformation for better impact, as well as South-South and Triangular Cooperation (when applicable), and efforts in innovative scaling-up approaches to achieve sustainable and inclusive smallholder agriculture development.

Methodology

Following what is rapidly becoming good practice among peers, IOE will apply principles of theory-based evaluations (see chapter 2) and anchor CSPEs in a strong context analysis ensuring that they address contribution to development results and attribution issues to the extent possible. This means that a first step in the evaluation is to review the logical chain that underpins IFAD's partnership with the country during the period under review, and to identify the underlying hypothesis and assumptions.

IFAD's increased focus on results has led to the introduction of results-based (RB) COSOPs (since 2006), with more results, information and focus on baselines, indicators and quantification, as well as annual reporting.⁶⁷ The new RB COSOPs have a results management framework which maintains a logical structure for the country programme while

allowing COSOP strategic objectives to be aligned with national poverty reduction strategies and an IFAD agenda for policy dialogue. While not a full TOC, it nevertheless presents basic elements that allow IOE to review the logic chain and identify key assumptions.

Another feature of the RB COSOP that deserves highlighting is the joint ownership, which entails greater involvement by in-country stakeholders in the design and implementation of COSOPs. This means that IFAD now gives more emphasis to the process of consultation with key stakeholders, harmonization and alignment with country poverty reduction policies and activities of other donors (including harmonization with United Nations reforms), and ownership of the COSOP by both the government and IFAD, thus calling for evaluation approaches that are able to deal with the attribution/contribution issue.

In theory, the most accurate measure of IFAD's contribution would be a comparison of the situation prevailing with and without its assistance. In practice, such counterfactuals are difficult to derive and defend for a country programme as a whole. However, it may be possible to derive reasonable counterfactuals for specific projects, particularly when formal impact evaluations have been undertaken in advance of the CSPE.

Formal attribution (i.e. separating IFAD's role from that of other internal or external players) is extremely difficult in a CSPE because of the multiplicity of factors that affect development outcomes and impacts at the country level. Therefore, the assessment of results will focus on determining whether IFAD has made a contribution to key results or outcomes that is both plausible and meaningful, and identifying the main drivers of the outcomes. Suggested reading to help address the attribution/contribution issue and TOC is found in annex VI.

Structure

An indicative table of contents of the main report of a CSPE is presented in box 5.

67 The COSOP result-based framework is also expected to be linked to individual project results framework, to be reported through the RIMS.

BOX 5 Example of contents page for CSPE reports

Executive summary

Agreement at completion point

Main evaluation report

- I. Introduction, summary of evaluation objectives, methodology and processes
- II. Background on the country and on IFAD-supported programme and operations
 - A. Country background
 - B. IFAD's strategy and operations for the CSPE period

III. Performance of the country strategy and programme

- A. The lending portfolio
 - B. Non-lending activities
 - C. Performance of the partners in managing the programme
 - D. Synthesis of the country programme strategy performance
- IV. Conclusions (including storyline) and recommendations
- V. Annexes

The first chapter of a CSPE will be introductory. It will briefly outline key characteristics of IFAD operations in the country (e.g. total number of loans approved, amounts, total costs, government counterpart financing and external cofinancing, grants). It will present the objectives, methodology and process of the CSPE as well as the scope of the CSPE (i.e. time frame, COSOP and operations considered). It will highlight key data collection methods, including main self-evaluation material available and salient limitations and constraints in data or analysis. The recommended length should be about three pages. As in the past, key points will be presented in a box at the end of each chapter.

The second chapter will include two main sections (country background and outline of IFAD's strategy and operations considered by the CSPE). The rationale for clustering these sections in a single chapter is to keep country background information more concise and directly connected to IFAD's strategy and supported operations.

The country background section will be kept short (indicatively no more than five pages). It will include: (i) a synthetic review of country economic and social indicators, particularly those that refer to rural poverty and development, followed by a brief review of key agriculture rural development issues that are directly relevant to IFAD's strategy

and operations; and (ii) a review of the main country policies and strategies of relevance to IFAD's programme within the evaluation time frame, the size of international cooperation in agriculture, an overview of the activities of the main international partners who work in the agriculture sector in the country. If the country plays a significant role in South-South or Triangular Cooperation, this will be briefly described. Finally, a section should be included on the results and lessons from evaluations of other key development partners working in the country's agriculture sector, including the United Nations Rome-based agencies.

A variety of data sources may be drawn upon to inform this chapter of the report: in 2013 IOE developed a data source kit for country context which is available on IFAD-IOE's intranet.⁶⁸

The section on IFAD's strategy and funded operations will concentrate on the CSPE time frame. A brief historical perspective may be provided to set the context. Findings from previous CSPEs or major evaluations that had a bearing on the strategy or operations may also be briefly presented. A description of the strategy as presented in the latest COSOP will be provided. A summary description of the main operations covered by the CSPE may be provided in this section or in an annex to the main report, depending on their number. This section

68 These typically include reports from the ministries of finance, agriculture and planning; World Bank country assistance strategy, country dataset and World Development Indicators; the UNDP Human Development Report and Human Development Index; the Economist Intelligence Unit; Country Profile, FAO; the World Health Organization global dataset on child malnutrition; IFAD COSOPs, country briefs and project documents; International Food Policy Research Institute working papers and studies; and others.

will typically be four or five pages long. As in the past, key points will be presented at the end of the second chapter.

Chapter three presents the main analysis of the evaluation. The focus of the CSPEs is on the results achieved at the country strategy and programme level. The length of this chapter will tentatively be 25-30 pages. Evaluation questions which the CSPE addresses are presented in box 6. However, to answer these questions, typically CSPEs will need to undertake inductive (bottom-up), deductive (top-down), and contribution assessments to gather evidence on the extent to which strategic objectives were achieved and to test the consistency of evaluation findings.

the agriculture and rural sectors. This will require an analysis of the COSOP's strategic guidelines but also a review of how the strategy has been played out in practice. The top-down analysis should deliver a systematic assessment of the results achieved at the overall programme level. Finally, the contribution assessment is aimed at determining how the main partners, particularly IFAD and the government, have performed. This will include their contribution to the overall programme (not only loan portfolio) management and whether adequate resources were allocated to achieving the programmatic objectives, and the level of synergy and complementarity with other partners.

BOX 6 Broad evaluation questions for CSPEs

- Did the country programme strategy work?
- Were the most appropriate approaches and interventions deployed to achieve the desired results?
- What are the explaining factors for performance?
- What made it work or fail?
- How can it be made better?

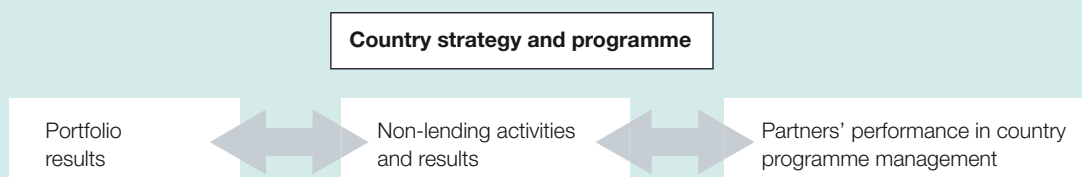
The bottom-up assessment focuses on the results of all IFAD-supported activities in the country, building on previous project evaluations and other evaluative evidence collected during the CSPE such as MTRs and monitoring reports. This is part of the analysis of the performance of the IFAD-supported loan portfolio, but also, to the extent that evidence permits, the results of non-lending activities and their longer-term effects.

The top-down assessment focuses on the Fund's strategic selectivity and the extent to which IFAD positioning took adequate account of its comparative advantage and the role played by other partners in

In chapter three the CSPE will review and assess three dimensions that are crucial for achieving results on the ground:

- Portfolio-level analysis of IFAD-funded past and current projects included in the CSPE scope;
- Non-lending activities: policy dialogue, knowledge management, partnership-building. A sample of grants (national, regional and global benefiting the country) and South-South and Triangular Cooperation and reimbursable technical assistance (when applicable) will also be reviewed;
- The performance of partners (IFAD and the government) in managing the country programme, including process aspects.

Conceptual framework of chapter 3



The chapter will conclude with an assessment at the programme (strategic) level which is informed by the above three dimensions. Within each dimension and at the programme level, ratings will be assigned consistent with guidance on ratings (see chapter 3) and as further explained below.

The project portfolio assessment will cover the loan-funded projects that are included in the CSPE scope. CSPEs will devote special time and attention to the project portfolio assessment, as the bulk of IFAD activities at the country level are based on loan-funded investment projects and programme. The project-level criteria will be adopted. As a general “rule of thumb”, a CSPE will apply the full set of project-level criteria when a project has either been completed (at the time of the evaluation) or approaching completion, or is at an advanced stage of implementation of its main activities.⁶⁹ For projects that are at an initial implementation stage (e.g. during the first four years of implementation) or have experienced serious delays, only relevance will be assessed and the CSPE will examine the implementation progress and the main explanatory factors. In some cases, the CSPE may also apply a sub-set of the criteria. Evaluability issues will first be examined in the approach paper and eventually will be explained in the first chapter of the main report.

The analysis of the portfolio assessment in the report will be organized following the project-

level evaluation criteria (see box 7).⁷⁰ Within each criterion-specific section, the presentation may be organized according to subsectoral, thematic, chronological or geographic clusters if this simplifies exposition and if it helps convey the key messages. Project-specific findings and evidence will need to be presented so as to back the final ratings. Under normal circumstances, the analysis of the criterion will conclude with an overall rating across the lending operations examined (e.g. one overall portfolio rating for relevance, effectiveness, etc.).⁷¹ Individual project ratings will not be discussed in the main report but presented in an annex table. Any further discussion on individual project ratings (if required) will be presented in an annex.

Assessment of non-lending activities

CSPEs also assess the process, performance and results of non-lending activities supported by IFAD and the government, defined as knowledge management (KM), partnership-building (PB) and policy dialogue (PD). Taken together, non-lending activities are expected to help enhance the programme’s development effectiveness in a given country. The assessment also includes a review of a sample of global, regional and country-specific grants and reimbursable technical assistance (when applicable) as well as achievements and synergy with the lending portfolio.

BOX 7 Criteria to be applied at the loan portfolio level

- Rural poverty impact
- Relevance
- Effectiveness
- Efficiency
- Sustainability of benefits
- Gender equality and women’s empowerment
- Innovation and scaling up
- Natural resources management and adaptation to climate change (one section with two separate ratings)

⁶⁹ The status of disbursement of funds from IFAD, the government and other major cofinanciers may be used as an indicator of state of progress in implementation. However, this should be reviewed with IFAD-PMD and with the government.

⁷⁰ The project evaluation methodology set forth in chapters 4 and 5 of the manual will be the reference in evaluating projects included for assessment in the CSPE. The evaluations should be undertaken based both on a desk review of documents and on field work. In some cases, IOE may have carried out specific project evaluations (i.e. PPEs or PCRVs) as separate exercises preceding the CSPE which will be summarized in the CSPE without the need for further field visits.

⁷¹ Evaluators will exercise his/her own judgement in assigning ratings and supporting those arguments with evidence. It is essential that the narrative on findings be consistent with the rating assigned. All ratings (except project performance) will be a rounded number, with no decimal points. Ratings across the lending operations examined will not be an arithmetic average of individual project ratings.

TABLE 7 **Matrix for the evaluation/evaluability of non-lending activities**

Non-lending activities evaluation matrix	Policy dialogue	Knowledge management	Partnership-building
Inputs	_____	_____	_____
Outputs	_____	_____	_____
Outcomes	_____	_____	_____

The assessment will consider to what extent the IFAD policies and strategies, inter alia, on grants (2015), on country-level policy dialogue (2013), partnership (2012), knowledge management (2008) and innovation (2007) have been implemented in the concerned country. It should also assess the type and amount of resources that have been committed to these areas.

A tool for assessing non-lending activities is the Matrix for the Evaluation/Evaluability of Non-lending Activities. It suggests a set of key evaluation questions and facilitates the identification of synergies (e.g. outputs of KM used as inputs for PD). Furthermore, it should be noted that in the mentioned matrix, grants will appear as sources of inputs for either PD, KM and/or PB.

Key evaluation questions (KEQ) suggested by the matrix (complementing those indicated below as “guiding questions”, which are framed in terms of relevance and effectiveness) are the following:

KEQ on non-lending activities concerning the design of the COSOP

- (i) Which inputs, if any, were earmarked in the COSOP (PD, KM, PB)?
- (ii) Were some of those inputs to be provided through grants?
- (iii) Which outputs, if any, were considered in the COSOP?
- (iv) Were there any outputs conceived as inputs for other non-lending activities?

- (v) Which non-lending activities outcomes, if any, were considered in the COSOP?
- (vi) Did the COSOP anticipate any synergies between non-lending activities and the lending portfolio?
- (vii) Which non-lending activities indicators, if any, were considered in the COSOP?

KEQ on non-lending activities concerning the implementation of the COSOP

- (i) Which inputs, if any, were used for non-lending activities (PD, KM, PB)?
- (ii) Were some of those inputs provided through grants?
- (iii) Which outputs, if any, were generated?
- (iv) Were there any outputs used as inputs for other non-lending activities?
- (v) Which outcomes resulted from the use of the outputs?
- (vi) Were there any synergies between non-lending activities and the lending portfolio?
- (vii) Which non-lending activities indicators, if any, were used during implementation of the COSOP?

If the COSOP included indicators for KM, PD and/or PB they must be used, comparing what was planned (if anything) with what was achieved. If non-lending activities indicators were not included, the evaluators should take into account the COSOP’s “theory of change” and the available information in

order to develop relevant and contextualized non-lending activities indicators. To facilitate this task, a set of references are provided in the footnote⁷² and the following general indicators may be taken as a starting point.

Examples of indicators for KM

- Number of different types of knowledge outputs generated by the country programme (e.g. diagnostic reports, policy notes)
- Number of knowledge-sharing events (e.g. workshops, meetings)

Examples of indicators for PD

- Number of partners engaged in policy dialogue (e.g. central government, donors)
- Number of policies or policy-related measures adopted by government and to which the policy dialogue in which IFAD participated contributed

Examples of indicators for PB

- Number of partnerships that were enhanced through the implementation of the country programme (e.g. with local governments, private sector organizations, etc.)
- Resources leveraged through partnerships

During the interviews it is important to search for evidence of the use or influence of knowledge products, PD and PB, incorporating such evidence, including “beneficiary feedback”⁷³ in the narrative of the evaluation (whenever possible, some of this evidence could be presented as boxes), combining quantitative and qualitative information.

Although PD requires essentially a qualitative approach, some PD indicators that could be useful are the following (complementing and further developing those presented in the preceding paragraph: (i) number of PD events

(e.g. workshops, high-level policy meetings), planned/actual; (ii) number of PD partners, planned/actual (indicating which partners such as national policy makers, regional policy makers, private sector representatives, international organization and/or bilateral donors representatives); and (iii) number of PD studies or documents that were produced (these would also be KM outputs).

Furthermore, in addition to the KEQ indicated above for all non-lending activities, in the case of PD (or “policy engagement at the country level”), it is convenient to distinguish between the PD that took place, the country policies, if any, that were adopted (or approved) to which the PD contributed, and the outcomes, if any, of those policies, if they were implemented.

- Policy dialogue → Policy adoption
- Policy adoption → Policy implementation
- Policy implementation → Policy outcome

Reference materials that can be used for assessing PD are Courdunou-Huci (2013) and Holland (2007).⁷⁴

While non-lending activities are expected to be included in the COSOP at the design stage, there may be situations where a CPM, or the country office, identifies opportunities for conducting policy dialogue jointly with other development partners originally not foreseen in the COSOP, or where a new contact provides opportunities for engaging with a knowledge institution. Such cases should be included and documented in the CSPE. Box 8 provides a list of guiding questions for non-lending activities, in terms of the evaluation criteria “relevance” and “effectiveness”, complementing the KEQ already provided.

⁷² Mansfield, Walter and Philipp Grunewald. 2013. *The use of Indicators for the Monitoring and Evaluation of Knowledge Management and Knowledge Brokering in International Development*, Loughborough University and the Institute of Development Studies Knowledge Services <https://www.ids.ac.uk/files/dmfile/Knowledgeindicatorsworkshopreport8thMarch2013.pdf>.

⁷³ Groves, Leslie (2015) *Beneficiary Feedback in Evaluation*, DFID. http://r4d.dfid.gov.uk/pdf/outputs/Evaluation/Beneficiary_Feedback_in_Evaluation.pdf.

⁷⁴ Corduneanu-Huci, Cristina, Alexander Hamilton and Issel Masses Ferrer (2013). *Understanding Policy Change: How to Apply Political Economy Concepts in Practice*, World Bank, <https://openknowledge.worldbank.org/bitstream/handle/10986/11879/9780821395387.pdf?sequence=2>. Holland, J. (2007) *Tools for Institutional, Political, and Social Analysis of Policy Reform*, World Bank, esp. chapter 3 http://siteresources.worldbank.org/EXTTOPPISOU/Resources/1424002-1185304794278/TIPs_Sourcebook_English.pdf.

BOX 8 Guiding questions for assessing the relevance and effectiveness of IFAD's non-lending activities

A. Relevance

- Are policy dialogue, partnership-building, and knowledge management objectives clearly outlined in the COSOP? Are they relevant to the IFAD programme as a whole? If activities that were not originally foreseen have been carried out, were they relevant?
- Were resources earmarked for non-lending activities and explicitly outlined in the COSOP (e.g. in the form of grants and/or the IFAD administrative budget)?
- How were the work and role of other partners taken into account in selecting the focus of non-lending activities?

B. Effectiveness

- Did the foreseen activities, if any, take place? If not, why not?
- To what extent and in what way did non-lending activities achieve the objectives (as explicitly articulated, or as implied)? Could the same objectives have been achieved with different (cheaper) means?
- Did non-lending activities contribute to the replication and/or scaling up of innovation promoted by IFAD?
- How well have non-lending components helped ensure a coherent country programme strategy, consistent with the commitments of the Paris Declaration on Aid Effectiveness?
- What have been the roles of the IFAD country representative, where applicable, and of the main government institutions in making non-lending activities effective?

A single rating will be provided against each non-lending activity: KM, PB and PD (relevance and effectiveness will not be rated separately), as shown in table 8. For brevity and to avoid redundancy, there will be no overall discussion on non-lending activities: this will be presented at the end of

this chapter in the overall programme performance discussion, based on the information gathered using the Matrix for the Evaluation/Evaluability of Non-lending Activities and the answers to the evaluation questions included in this section.

TABLE 8 Example CSPE ratings for non-lending activities

Non-lending activity	Rating
Knowledge management	6
Partnership-building	5
Policy dialogue	2

Review of grants

A sample of grants will be reviewed as well. As a general rule, grants that have supported PD, KM and PB will be assessed as indicated in the previous paragraphs. Grants that did not relate to these non-lending activities shall be discussed in a separate section. Typically, a CSPE will review some five to six grants and these will include country-specific grants (if they exist) as well as regional or global grants with foreseen activities in the country. In order to select which grants to include, the evaluators will create a list of grants with foreseen activities in the country and discuss with the CPM, country office staff and other informed staff at IFAD (e.g. from the Policy and Technical Advisory Division-PMD or the Strategy and Knowledge Department) as well as stakeholders in the country. The sample of grants will seek to cover the main grant thematic areas in the country (e.g. agricultural research, rural finance, KM).

In 2014, IOE conducted a CLE on IFAD's Policy for Grant Financing which is a useful reference guide for preparing questions on grants.⁷⁵ The following questions are appropriate for individual grants:

(i) Was there a demand for that specific grant activity either by the CPM or by one of the main partners in the country? (ii) Has the grant achieved or is it likely to achieve the expected results? and (iii) Is there an ongoing plan for IFAD or any of its partners to internalize or use knowledge, technology or other products generated by the grant, and if so, is the plan being followed up? As in the past, no separate rating is assigned to grants, but the discussion on the use and internalization of knowledge, PD, PB, technology or other grant products will be part of the final country programme strategy assessment.

The **assessment of the partners' performance** in managing the country programme, including process aspects, will generally relate to IFAD and the government (other partnerships would normally be considered under non-lending/ partnership-building). The assessment will include two dimensions: (i) partners' performance at the project level (this assessment will be done using the general questions presented in chapter 3 for IFAD and the government and will assign ratings to the partners' performance in each concerned project); and (ii) broader aspects that pertain to the overall country programme management and related processes.

The review of broader country programme management is crucial, as the operating model defined for the country⁷⁶ helps to determine whether the programmatic objectives (as per the COSOP), or other initially unforeseen objectives, have been met. Therefore, the CSPE assesses, among other issues, whether the administrative budget was appropriate for ensuring proper supervision and implementation support; the type of country presence pursued; and whether adequate support, time and resources were provided to the CPM for PD and KM.

The government's contribution to country programme strategy and country programme management will also be reviewed (see box 9). At the end of this section, a single composite rating will be given to IFAD and the government that encompasses performance at the project level and the overall programme.⁷⁷ Individual project-level ratings will be presented in a table in an annex. If required, they will be discussed in an annex.

⁷⁵ http://www.ifad.org/evaluation/public_html/eksyst/doc/corporate/cle_grants.pdf.

⁷⁶ For example, in terms of supervision and implementation support, the role of the country programme management team and country presence arrangements (if any).

⁷⁷ This rating will not be an average.

BOX 9 Guiding questions on partners' performance and COSOP management

- Did supervision and implementation support arrangements (including country presence, if applicable) support achievement of the COSOP objectives?
- Were sufficient administrative and human resources made available by both IFAD and the government for the implementation of the country strategy?
- How was the COSOP MTR used to adjust or change the direction, speed or approaches to implementation (for COSOPs approved after September 2006) to reflect changes in the country context?
- What is the quality of the COSOP results management framework, project status reports, and aggregated RIMS reports and country programme sheets, and were Management actions in connection with this information system appropriate?
- Was the COSOP M&E performed properly? Were annual country programme reviews undertaken in a timely manner and were the corresponding recommendations implemented within the required time frames?
- Did the country programme management team concept function appropriately and make the necessary contribution to country programme management?
- Did the Fund systematically assess the progress made in COSOP implementation on an annual basis (for COSOPs approved after September 2006)?
- To what extent has IFAD complied with the Paris Declaration on Aid Effectiveness (2005), which was reaffirmed by the Accra Agenda for Action (2008) and the Busan Declaration (2011)? This means that the CSPE should assess progress in five broad areas identified in the Paris Declaration: ownership, alignment, harmonization, managing for results and mutual accountability.

Synthesis of the findings of chapter three: performance assessment of the country programme strategy

This section assesses the relevance and effectiveness at the country strategy level, building on the previous analysis in the same chapter. The section will include two mandatory parts (strategic relevance and effectiveness) and a third one which is optional and may be used to cover any other broader strategic issue which could not be embedded in the previous two.

To assess the relevance of the country strategy, the CSPE will examine the COSOP and its reviews, as well as the actual operationalization of the same. In several instances, key parts of a strategy or a country programme could be unwritten but based on (tacit) consensus and negotiation between the main partners. Importantly, the CSPE, in analysing

the main logic underlying the COSOP, will pinpoint the main assumptions that have been made in the logical framework. This is consistent with a theory-based approach.

As far as the COSOP is concerned, the CSPE will consider the alignment and coherence of the: (i) strategic objectives; (ii) geographic priority; (iii) subsector focus; (iv) main partner institutions; (v) targeting approach used, including emphasis on selected social groups; (vi) mix of instruments in the country programme (loans, grants and non-lending activities); and (vii) the provisions for country programme and COSOP management. In the discussion, emphasis will be put on the content of the actual strategy pursued by the country programme, whether clearly outlined in the COSOP or not. Guiding questions are presented in box 10.

BOX 10 Guiding questions for assessing the relevance of the country programme strategy

A. Assessment of the alignment of strategic objectives

- Were the strategic objectives identified in the COSOP aligned with the government's strategies and policies, and consistent with the overarching objectives of the prevailing IFAD strategic framework and relevant corporate policies?
- To what extent has IFAD fostered partnerships with other bilateral and multilateral donors working in agriculture and rural development in the same country?

B. Evaluating the design and coherence of the strategy

- Did the strategy succinctly articulate IFAD's comparative advantage and competencies in the country (i.e. country positioning)* and constitute a coherent and well-articulated country programme?

- Were the most appropriate strategy elements and subsectors for investments chosen, given the context and rural poverty analysis?
- Were the geographic priorities defined in the strategy consistent with the definition of the target groups?
- Were the main partner institutions (e.g. for project execution, supervision and implementation support, community mobilization, cofinancing) the most appropriate for meeting the country strategy objectives?
- Were the strategic objectives and design and implementation properly informed by IFAD's experiences in the country and elsewhere?

* Country positioning is a measure of how well the organization responded to (or even anticipated) the evolving development challenges and priorities of the government, built on the organization's comparative advantages, and designed its country strategies and programmes in a manner that took into consideration the support available from other development partners.

The assessment of effectiveness of the country strategy determines the extent to which the overall strategic objectives (as per the COSOP) were achieved and whether other significant – but originally not foreseen – results have been attained at the programme level, and whether a credible logical nexus can be established between the partners' as well as IFAD-supported initiatives (lending, non-lending, programme management) and the observed results. In particular, the section on effectiveness will seek to explain the role played by the government and IFAD in managing the overall country programme in the achievement of results.

Typical challenges that CSPEs will face include the following: (i) COSOP objectives may be identified in broad terms and ex ante milestone indicators may refer to data that are not available at the desired level

of disaggregation; and (ii) when data are available, they may focus on outputs (e.g. cumulative number of clients served through extension services) and provide little information about the results or other crucial aspects (e.g. what percentage of farmers are adopting improved crop management techniques and what are the results in terms of yields and food security, etc.).

One way to address these issues is to map the stated COSOP objectives against activities reviewed by the CSPE – lending, non-lending and other programme management activities. Then achievement indicators (if available) or other findings from the previous analysis in the chapter may be used as evidence of certain results. The CSPE will establish the likely level of influence of IFAD-supported activities on the achievement indicators.

TABLE 9 **A scheme for assessing strategic (COSOP) objective achievement**

Statement of objective	Target	Mapping of IFAD-supported activities	Achievement indicators (if existing) or CSPE relevant findings (lending and non-lending)	Level of influence by IFAD-supported activities and programme management activities
Strategic objective X	Target outcome indicators	<ul style="list-style-type: none"> • Project 1 • Project 2 • Non-lending activity A • Selected programme management activities 		e.g. low/partial/high
Strategic objective Y	Target outcome indicators	<ul style="list-style-type: none"> • Project 3 		e.g. low/partial/high
Strategic objective Z	Target outcome indicators	<ul style="list-style-type: none"> • Project 1 • Project 4 • Non-lending activity B 		e.g. low/partial/high
Other significant results not foreseen in COSOP	Not available (by definition)	<ul style="list-style-type: none"> • Project 2 • Project 5 		e.g. low/partial/high

A schematic overview of this process is shown in table 9. Note that a single lending or non-lending activity may span across several strategic objectives. Moreover, some programme achievements may correspond to unforeseen objectives and, finally, perhaps in less frequent cases, there might be a COSOP strategic objective without any traceable operation to pursue them.

Once the mapping exercise has been done, the CSPE will be able to consider the effectiveness of the country programme strategy, notably:

- To what extent were the COSOP's main strategic objectives achieved?
- Are there other originally not foreseen results that have been attained and how were they achieved?
- What context changes have influenced or are likely to influence the fulfilment of the strategic objectives? Was the COSOP properly adapted mid-course to reflect changes in the context?

- What has been the level of influence of IFAD-funded activities?
- How have efforts made by the government and IFAD in managing the programme contributed to the achievement of the results?

Chapter 3 of the CSPE report will conclude with a synthesis of the country strategy performance. This will pull together key findings from the lending portfolio, non-lending activities and performance of partners in managing the programme. CSPEs should provide individual ratings for the country strategy relevance and effectiveness and a composite rating for the country strategy performance, which will be the ultimate rating of the evaluation. The composite rating should not be an arithmetic average, but take into account findings on lending, non-lending and partners' performance using the informed judgement of the evaluators (see table 10).

TABLE 10 **Example CSPE ratings for country strategy performance**

Country strategy	Rating
Relevance	5
Effectiveness	3
Country strategy performance	4

TABLE 11 **IFAD-government partnership assessment**

	Rating
Portfolio performance	6
Non-lending activities	2
Country strategy performance	3
IFAD-government partnership	4

Finally, based on the assessments and ratings of portfolio performance, non-lending activities, and country strategy performance, the CSPEs will generate a composite assessment and rating for IFAD-government partnership in reducing rural poverty. The final rating will not be based on a simple arithmetic average of the individual ratings, but on a holistic consideration of all the evidence available and the analysis conducted by IOE. See table 11 as an example.

Optional subsection on broader strategic, subsectoral or thematic issues

The evaluators have the option to conclude the discussion on the country programme strategy performance with a subsection on outstanding issues that, for any reason, may not fit well under country programme strategy relevance or under effectiveness. This might be particularly the case of findings on special themes or subsectors that were

not emphasized under the COSOP objectives but were found important or emerging in the course of the conduct of the CSPE.

Similar to the preceding chapters, a table with key points will also be presented at the end of chapter 3.

Evaluation methods

In line with international evaluation practices and IOE practice, a mixed-method approach is applied for a CSPE evaluation: (i) a desk review of existing literature and IFAD documents, including IOE evaluations, project-level information and self-assessments; (ii) interviews with relevant stakeholders in IFAD and the country; and (iii) direct observation in the field (also see chapter 2 on methods and triangulation and the previous section on “bottom-up and top-down” analysis in this chapter).

For the field work, a combination of methods is generally used for data collection. These include:

(i) focus group discussions with a set of questions for project user and comparison groups; (ii) government stakeholder meetings (national, regional/local, including project staff); (iii) sample household visits using a pre-agreed set of questions for household members to obtain indications of levels of project participation and impact; and (iv) key non-government stakeholder meetings, e.g. civil society and private sector representatives (see chapters 2 and 3 for further reference). The methodology will be further developed in the approach paper and notably in its evaluation framework.

Self-assessments

Consistent with IFAD's Evaluation Policy and the ECG GPS, self-assessments by those involved in the design and implementation of the COSOP and IFAD-funded operations are one element of IOE's approach to country evaluations. Typically, self-assessments are made by PMD and the government concerned. Where available, COSOP completion reviews⁷⁸ are treated as self-assessment documents for CSPEs and a dedicated self-assessment will not be needed in such cases. Although a separate document is not required in such cases, CPMs and the relevant government authority may be invited to respond to questions in the CSPE framework not covered by the completion review.⁷⁹ Where a client survey has recently been undertaken, this may also serve the purpose of a self-assessment, and may be complemented by other questions.

Self-assessments are conducted before the main CSPE mission, and a discussion on this topic should be organized at IFAD headquarters between IOE and the relevant regional division to review the focus and dimensions to be covered.

Special performance and impact evaluations

Depending on the quality of available M&E data, and especially in those countries with no (or limited) access to independent evaluative evidence, IOE may commission special performance and impact evaluations of selected IFAD-funded projects and

programmes. This should be done before the main CSPE mission is fielded so that deliverables are available to the CSPE team before it begins its country work.⁸⁰ Such assessments would follow the IOE project performance evaluation or impact evaluation methodology and process.

Process

Consistent with IFAD's Evaluation Policy, the first step for a CSPE is for IOE to prepare an **approach paper**, for which an indicative contents page is provided in table 12. Normally, a CSPE approach paper is approximately 15 pages long, excluding annexes. The draft approach paper may be drafted before or after the preparatory mission but will be finalized after the latter so that information that has been collected can be reflected in the overall design of the evaluation. The approach paper will provide an opportunity to customize selected phases in the process, should a COSOP completion review be available in the country where IOE plans to conduct a CSPE.

A CSPE approach paper will also include an evaluation framework in an annex. This illustrates the link among CSPE's objectives, the evaluation criteria, key questions that need to be addressed to achieve the evaluation's objectives, and the main sources and instruments of data collection.

The IFAD Evaluation Policy states that IOE shall identify the members of the **core learning partnership** (CLP), the main users of the evaluation, in the approach paper. The role of the CLP shall be to enhance the quality of the evaluation, as well as to build ownership among key partners in the evaluation process and its outcomes. Specifically with respect to CSPEs, after completion of the independent evaluation report, IOE "facilitates a process, involving PMD, recipient countries and other partners, through which the main users of the evaluation can deepen their understanding of the evaluation findings and recommendations and make them more operational." In this respect, IOE is committed to organize an in-country national workshop.

⁷⁸ Prepared by IFAD in collaboration with the government.

⁷⁹ For projects that are closed, the completion report will be considered as the self-assessment.

⁸⁰ For example, in the Pakistan CPE recently undertaken by IOE, a local NGO was contracted to perform such assessments for two IFAD-funded projects in the country. These activities produced primary data which proved extremely useful for the CPE team in conducting its work.

TABLE 12 **Sample contents page for a CSPE approach paper****Contents page**

- I. Background and rationale
- II. Country background
- III. Overview of IFAD assistance to the country
- IV. Evaluation objectives, methodology and collection of evidence
- V. Core learning partnership
- VI. Consultants team
- VII. Communication and dissemination
- VIII. Proposed schedule

The IFAD country programme management team should be properly represented in the CLP. Normally, the following people should be included: (i) the CPM and PMD division director, as well as the country officer, as applicable; (ii) directors/ coordinators of projects included as part of the CSPE assessment; (iii) senior government officials (from the main coordinating ministry working with IFAD and the technical ministry associated with IFAD operations); (iv) representatives of cofinancing organizations and cooperating institutions (only if applicable); (v) the main NGOs associated with IFAD operations in the country, as well as representatives of pertinent local community and advocacy groups; (vi) members of selected academic and research institutions; and (vii) the IOE director and the CSPE lead evaluator. This list is indicative. The actual CLP composition is informed through consultations with the relevant regional division, government authorities and civil society in the country concerned.

To provide the full range of skills needed, **CSPE teams** should consist of four to six members, including the IOE lead evaluator, who will participate in the main mission the entire time. The team should

be staffed so as to permit competent treatment of major strategy and policy matters, subsector issues, and project design and implementation modalities. It is crucial to include national consultants and women on the CSPE team, who should also satisfy all provisions in IOE's Conflict of Interest Policy (see annex IV) for consultants.

IOE may seek the collaboration of one **senior independent adviser** (SIA) for the preparation of a CSPE, on a case-by-case basis.⁸¹ If a SIA is recruited, her/his report will provide a review of the overall evaluation process as well as the main findings and recommendations in the final report, and on the quality of the supporting evidence.

The CSPE process is conducted in six phases: (i) Preparation; (ii) desk review; (iii) field mission and data collection; (iv) analysis; (v) preparation of draft final report and review process; and (vi) finalization, dissemination and follow-up. Further details on the CSPE process may be found in the Evaluation Process Guidelines. CSPEs will be fully concluded within twelve months from start to finish.

81 The SIA will be an individual with experience and expertise in evaluation and the topic covered by the CSPE, and with an international profile. The role of the SIA is to assess the quality of the evaluation process and contents of the final report. S/he will review the draft approach paper, and draft final report, as well any other major deliverables produced during the process. S/he will prepare a (joint) independent report (2-3 pages) based on the final evaluation report, which will be included in the final document at the time of publication.

Phase 1: Preparation

The approach paper is developed in this phase and an evaluability assessment is made. An important aspect of this is the advance preparation of input documents, including self-assessments where these are to be prepared.

During the preparatory mission the IOE lead evaluator collaborates with the central government in the designation of a focal point for interactions throughout the CSPE on matters primarily related to the mission's programme, logistical issues and communication with projects and other government partners. The preparatory mission will allow IOE to elicit input from the government and other partners in the country which will be used to refine the evaluation questions and process.

The preparatory mission also provides the IOE lead evaluator with an opportunity to form an assessment of the available information and to identify the knowledge gaps to be filled. This contributes to an appreciation of the evaluability of the country strategy and programme. Therefore, following the preparatory mission, a decision is taken whether to embark on primary data collection, statistical surveys, the organization of focus groups, rapid rural appraisals, and other investigative and participatory methods.

Selection of a representative cross-section of partners and beneficiaries drawn from the capital city and from selected project(s) area(s) is essential. It is equally useful to impress upon government and project authorities the importance of using representative data regarding communities, households and project sites. Furthermore, the preparatory mission will also be the appropriate time to identify suitable national consultants for the CSPE team.

Project authorities to be visited during the main CSPE mission are informed in advance and will provide feedback on the feasibility of transportation and security issues. It is mandatory for IOE to contact the local United Nations Security Officer to gain an understanding of the situation on the ground. Security clearance from the United Nations

system, where necessary, must be obtained before starting field visits.

Phase 2: Desk review

As in the case of project evaluations, this phase follows the preparation of the draft approach paper. The desk review informs further work, which identifies a list of issues and questions to be further analysed during the main CSPE mission, as well as the availability of data and any information gaps that need to be filled in by the evaluation. The desk review is completed before the main CSPE mission. It is an internal document, kept on file for accountability purposes, to enable tracking back from conclusions and recommendations over findings to the facts, so as to ensure the robustness of the evaluation.

Phase 3: Field mission and data collection

CSPE's will include a main mission composed of multidisciplinary experts, to be led by the designated IOE staff member responsible for the evaluation. The mission will spend around four weeks in the concerned country for broad-ranging consultations and data and information collection. Two to three weeks will be spent in the field, in particular to collect data and information from beneficiaries, community-based organizations, subnational authorities and other stakeholders at the local level. Visits are normally made to major project sites, and/or projects with particular issues of interest, as identified by the desk review. A PowerPoint presentation summarizing the emerging CSPE findings will be prepared by the evaluation team before the end of its mission, to be presented by IOE at a wrap-up meeting to government authorities and other stakeholders. The PowerPoint presentation should not contain ratings or recommendations. The CPM and country office staff (in countries where IFAD has a country office) will also attend the CSPE wrap-up meeting. In countries where IFAD has a country office, the IFAD CPM may participate in the wrap-up meeting from headquarters using appropriate telecommunications technology, with the understanding that staff from the country office would attend the meeting in person.

Phase 4: Analysis

During this phase, facts and evidence drawn from different sources are brought together to answer the overarching evaluation questions. Triangulating this evidence, identifying any weaknesses, and collecting additional evidence where needed are key activities in this phase and an inherent part of a robust analysis. In addition to findings from the context analysis, the desk reviews of lending and non-lending, and from the country mission, this phase may include information drawn from commissioned thematic working papers on topics of importance to the country programme.

Phase 5: Preparation of draft final report and review process

The primary activity in this phase is the report writing. The draft final report will be peer-reviewed within IOE (and by the SIA, if applicable). After including the IOE (and SIA) comments, the document is shared with IFAD Management and the government for review and comments, in line with the protocol for communication established. They will be given around four weeks (20 working days) to provide their consolidated comments on the draft final report. In the communication with IFAD-PMD and the government, IOE may suggest areas of priority for comments, for example correcting factual information or updating with changes that have taken place since the latest evaluation field visit.

During this period, a meeting may be organized with the concerned regional division and the Associate Vice-President, PMD to discuss the main findings and recommendations of the report. Moreover, under normal circumstances, IOE will also conduct a short follow-up mission (two working days) to discuss with the government and other main stakeholders the key findings and recommendations of the report. In the past, it has been found useful to conduct such a mission after the draft has been shared with the government but before the government has transmitted its comments. This will enhance clarity from the side

of the government on the contents of the report and will help focus comments on a narrower set of outstanding items. The short mission will also be an opportunity to agree on details for organizing the national workshop (see below).

Phase 6: Finalization, dissemination and follow-up

The IOE lead evaluator is responsible for preparing the final report and its contents. Management and government comments are integrated, in line with the Evaluation Policy, into the final report and an audit trail of their comments is also produced by IOE. The final report and audit trail will be shared with IFAD Management and the government.

A **national workshop** of one day in length is usually organized at the end of each CSPE at the national level to discuss and deepen the participants' understanding of the evaluation findings and recommendations. The workshop is held following the finalization of the CSPE report. A concept note is prepared by the lead evaluator, outlining the objectives of the workshop and the main activities involved (including an optional field visit to an IFAD-funded project), and provides a provisional list of participants and other relevant information. Along with CLP members, representatives of major organizations should be invited to participate, including participants from federal/central and provincial/state governments, NGOs and civil society, key private sector actors, research institutions and universities, and multilateral and bilateral donor organizations.

Thereafter, IFAD Management and the government will prepare the **ACP**. An ACP is included in each CSPE, as stipulated in the IFAD Evaluation Policy. The ACP will contain a summary of the main evaluation findings and recommendations that IFAD Management and the concerned government agree to adopt and implement within specific time frames. The ACP will also document any recommendations that are not found feasible by either IFAD Management and/or the government.

FIGURE 3 Country programme and strategy evaluations: phases, processes and key deliverables

Phases	Processes	Deliverables
Preparatory phase	<ul style="list-style-type: none"> • Evaluability assessment • Preparatory mission • IOE peer review 	<ul style="list-style-type: none"> • Approach paper • Input documents (e.g. self-assessments by PMD and government)
Desk review phase	<ul style="list-style-type: none"> • Collection and review of relevant documentation and data 	<ul style="list-style-type: none"> • Desk review notes (for internal use)
Field mission and data collection	<ul style="list-style-type: none"> • Primary data collection • Main mission: interaction in the capital, discussion with beneficiaries and site visits 	<ul style="list-style-type: none"> • Performance and impact assessments, as required • PowerPoint presentation
Analytical phase	<ul style="list-style-type: none"> • Identifying data gaps • Triangulation of evidence • Limitation identification • Additional evidence collection, where needed 	
Preparation of draft final report and review process	<ul style="list-style-type: none"> • Report-writing • IOE peer review • Draft final report shared with SIA for review and then to PMD and government for comments • Follow-up mission to the country to discuss the report and the preparation of the workshop 	<ul style="list-style-type: none"> • Draft final report • [SIAs comments if applicable] • PMD and government comments
Finalization, dissemination and follow-up	<ul style="list-style-type: none"> • Preparation of ACP • Communication 	<ul style="list-style-type: none"> • Final report and audit trail • National workshop and final ACP • Distribution of final CSPE report • Evaluation Profile • Evaluation Insight

The process and template for ACPs agreed by IOE and PMD will be used for the preparation of ACPs (see annex III). The ACP is an instrument of mutual accountability, consistent with Paris Declaration commitments, as it reflects the joint response of both the government and IFAD Management to the evaluation. This is important because the government of the relevant country is ultimately responsible for the execution of the IFAD-funded operation.

All final CSPE reports, including the ACP, are presented to the Evaluation Committee, as well as to the Executive Board. IFAD Management will report on the implementation status and follow-up to CSPE recommendations in the PRISMA. Chapter 10 summarizes the main products to be generated, including the process for their dissemination and communication.

Reports

A standard CSPE reporting format is adopted to foster uniformity in overall presentation, but reports may be tailored to the needs of a particular country case depending on the scope and evaluation questions. The report should include a brief and mainly descriptive coverage of the country context, country strategy and programme. Subsequently, a more analytical approach is taken in presenting programme implementation, programme outcomes and impacts, partnerships and thematic issues. Findings that are based on clear evidence are presented to support the conclusions and recommendations.

The CSPE report will be presented in plain language and will focus on those key issues that could be evaluated conclusively, rather than on all issues that have been examined; this is essential to ensure credibility and utility of the evaluation.

The total recommended length of the main CSPE report is about 55 pages. In addition, a CSPE report will contain a foreword, an overview (around 5,100 words),⁸² the ACP, as well as annexes. In its table of contents, the report will list the technical working papers (e.g. on subsector issues, policy and strategy matters, institutional arrangements, etc.), that can be made available by IOE upon request: these will not be included in the report.

CSPE reports are usually written in the IFAD official language (Arabic, English, French or Spanish) of the recipient country. The acknowledgements, foreword, overview and ACP will be translated into English if they are written in another IFAD official language. Both the original language and English versions will be included in the final report.

82 Eventually all CSPEs will be shared with the Executive Board; a ceiling of 5,500 words is imposed on the main document, including table of contents and front page.



7

7. Corporate-level evaluation

Background

Since around 2000, IOE has conducted a fairly large number of CLEs. These are the most complex and challenging types of evaluations done by IOE, and contribute to far-reaching changes and reform in the Fund. They address systemic issues that cut across individual operations, country programmes or regions and contribute to strengthening accountability and learning for better institutional transformation and results on the ground.

Definition

The Evaluation Policy states that “Corporate-level evaluations are conducted to assess the results of IFAD-wide corporate policies, strategies, business processes and organizational aspects. They are expected to generate findings and recommendations that can be used for the formulation of new and more effective corporate policies and strategies, as well as improve business processes and the Fund’s organizational architecture, as required.” CLEs done by IOE can be grouped into the following categories, depending on their focus: (i) thematic areas of interest (e.g. gender, rural finance); (ii) corporate policies and instruments (e.g. grants, direct supervision); and (iii) corporate processes, such as those related to institutional efficiency (efficiency or replenishments CLE).

For the thematic type CLEs there might be some similarities with evaluation syntheses (see chapter 9). However, while the evaluation synthesis is a knowledge product based on evaluative evidence rather than an evaluation, the CLE is clearly an evaluation. What they have in common is that both include a review of evaluation findings,

but where this is the core of the report in evaluation syntheses, it is only one of several building blocks in CLEs. In general, CLEs are far more complex and entail a vast amount of data collection and analysis. In some cases, an evaluation synthesis has preceded a CLE, thus providing a very strong basis to build on.

The main aim

CLEs are intended to provide timely, credible and useful information that can be acted upon to increase IFAD’s institutional effectiveness, development effectiveness and efficiency. Over time, CLEs have had significant impact in terms of providing guidance at the policy and strategic levels and have had a strong influence on IFAD’s corporate development, its operational approaches, and its business model. This was clearly shown and documented for instance in the Corporate-level Evaluation on IFAD Replenishments.⁸³

Selecting topics for CLEs

Many CLEs have a direct link to important corporate processes (e.g. institutional efficiency) or have been suggested as part of replenishment exercises, or undertaken explicitly at the request of the Executive Board (e.g. grants policy).⁸⁴ Indeed, the Corporate-level Evaluation on IFAD Replenishments found that “there seems to be a mutually reinforcing process whereby reform is driven to some extent by evaluations and reinforced by the replenishment process, which in turn is informed by the evaluative work.” Also, evaluation topics stem from global trends that affect all multilateral development banks (private sector, gender, fragile states, etc.) and CLEs

⁸³ http://www.ifad.org/evaluation/public_html/eksyst/doc/corporate/replenishments_full.pdf, paragraph 97.

⁸⁴ As for other IOE evaluations, all CLEs are eventually approved by the Board as part of IOE’s annual work programme.

may also cover subjects that can help IFAD identify and address various types of risks, in terms of operational, financial, strategic, etc.

To enhance the transparency and formality of the selection of topics for CLEs, a set of criteria is developed in the selectivity framework (see annex V). Based in part on the selectivity framework, IOE will consider a topic that could: (i) contribute to IFAD's strategic priorities and replenishment commitments; (ii) contribute to strengthening IFAD's institutional accountability; (iii) contribute to filling a critical knowledge gap in IFAD; (iv) generate results to feed into pertinent corporate policies, strategies and/or processes in a timely manner; or (v) help minimize critical corporate risks. A set of criteria for prioritizing and selecting CLEs has been developed, which is included in the IOE selectivity framework.

Timing and timeliness of CLEs

Whatever the “driver of change” as described above, the timing and timeliness of CLEs are critical, as they feed into a specific policy, strategy or development of corporate business processes. Therefore, the decision to conduct a CLE at a particular juncture, and the time required to finalize a CLE should be carefully considered in consultation with IFAD Management and the governing bodies. In general, CLEs take around 12 months from start to completion (excluding presentation to the Executive Board). This will ensure and enhance the usefulness of the CLE.

Who leads CLEs?

Given their importance and scope, CLEs are usually carried out under the leadership of the IOE Deputy Director or a senior evaluation officer. They are supported by other evaluation officers and a team of consultants with expertise and experience in selected topics. As with all other evaluations, IOE is responsible for the contents of the final CLE report and the overall evaluation process.

Senior independent advisers

IOE will seek the collaboration of one or two SIAs for the preparation of a CLE. The SIAs would normally

be individuals with experience and expertise in evaluation and the topic covered by the CLE, and have an international profile. The role of the SIA is to assess the quality of the evaluation process and contents of the final report. S/he will review the draft approach paper, the draft final report and any other major deliverables produced during the process. S/he will prepare a (joint) independent report (2-3 pages) based on the final evaluation report, which will be included in the final document at the time of publication. The SIA's report will provide their reflection on the overall evaluation process as well as the main findings and recommendations in the final report.

Objectives and methodology

The point of departure for a CLE is the articulation of the rationale, purpose and intended use of the evaluation, addressing why the evaluation is being undertaken at this particular point in time; why and for whom it is undertaken; and how will the evaluation be used for strengthening accountability and learning?

Evaluation objectives

While the purpose statements address the “why are we doing this” question, the specific objectives clarify what the evaluation aims to discover. Therefore, while purposes are general in nature, objectives are specifically constructed to reflect the particular issue/theme being evaluated.

Objectives for CLEs would generally include three aspects: (i) assessing performance and results; (ii) generating lessons and identifying areas of improvement; and (iii) proposing recommendations. In sum, their main aim is to strengthen the organization's accountability framework and learning loops for enhanced rural transformation and better livelihoods.

Box 11 shows how these generic objectives were made specific in a particular CLE, using an example the CLE completed in 2010 on IFAD's Private-Sector Development and Partnership Strategy.

BOX 11 Example of specific objectives

- Assess the relevance and evaluate the implementation of IFAD's 2005 private-sector strategy;
- Evaluate the emerging results of IFAD-supported projects designed after the adoption of the private-sector strategy;
- Assess the evolving approaches, as well as good and less than satisfactory practices, to IFAD's private-sector development efforts;
- Examine the instruments and experiences of other development organizations in engaging the private sector in agriculture and rural development, with the aim of identifying lessons that could be pertinent for IFAD; and
- Generate a series of findings and recommendations that might serve as building blocks for IFAD's future engagement with the private sector.

Evaluation scope

Clarifying the evaluation scope means defining the unit of analysis, the time period, geographical area, target groups, organizational set-up, implementation arrangements, policy and institutional context and other dimensions to be covered by the evaluation. Being very clear on the boundaries of the evaluation is critical to develop an appropriate design, allocate an adequate budget, and manage expectations of what information and findings can be expected from the CLE.

Serious consideration should be given to how far back in time it is useful to go, given available resources: What can history teach us? How relevant is it today? Do we need, for accountability or learning purposes, to document the evolution in IFAD's approach to the issue being examined? How relevant are lessons from or benchmarking with peers, and which peers? Answering these questions has resource implications and therefore merits serious reflection. Too wide a scope compared to time and resources available risks jeopardizing the rigour and hence credibility of the exercise.

Since the focus of CLEs varies, the analytical framework and the design will vary from CLE to CLE and therefore a standard design cannot be applied; it has to be developed on a case-by-case basis depending on the subject of the CLE, available resources and timing.

Evaluation design

Once purpose, scope and objectives have been clarified, the actual design phase starts. Taking into account the main evaluation objectives, the design will include articulating a clear notion of the results chain (e.g. in the form of a TOC) including the underlying causalities, hypothesis and assumptions. The analytical framework should be developed and discussed in partnership with stakeholders to ensure maximum engagement and capturing of the most important views and perceptions.

Consistent with IFAD's Evaluation Policy, a first step for every CLE is for IOE to prepare an approach paper, which will outline the evaluation's background and rationale, objectives, key questions and methodology and process, time frames, human resources, and communication and dissemination activities. Conducting an evaluability assessment is also part of developing the approach paper and the findings of that assessment are an important input to the design.

At the time of developing the approach paper, the evaluation criteria to be used in a particular CLE must be developed carefully. The specific evaluation criteria should be decided based on the issue being evaluated, and sound judgement exercised in determining the criteria based on the evaluation questions (see next paragraph and chapter 3 for detailed guidance on criteria and questions related to criteria). The three standard criteria of relevance,

efficiency and effectiveness must always be applied in CLEs, or if they are not, the reasons must be clearly stated. Additional criteria can be applied and the rationale should be given in the approach paper.

Sufficient time and effort, including for dialogue with stakeholders and experts, should be allocated to clearly identify the evaluation questions, as these will determine the ultimate utility of the evaluation. The evaluation questions are set out in the approach paper and further elaborated in the inception phase (see next section on CLE process), and will usually consist of a set of higher-level questions, with sub-questions (see chapter 3 regarding evaluation questions).

Each approach paper will include an evaluation framework (in the form of a matrix). The latter will coherently map the evaluation objectives with the evaluation questions and illustrate the key activities to be undertaken and the instruments to be used to answer the evaluation questions.

Evaluation building blocks

Once the evaluation criteria and key questions have been defined, the evaluators must choose the specific methods and activities (or building blocks) that will be conducted during the evaluation to collect the required data and information for analyses. CLEs generally require different types of activities which allow for the collection of different types of data and information (qualitative and quantitative).

Such building blocks may have separate terms of reference and deliverables in the form of separate working papers. Box 12 shows an example of building blocks used in the 2010 CLE on gender.

One building block is the self-assessment, enshrined in IFAD's Evaluation Policy, which stipulates that before initiating independent analysis for a CLE, IOE will invite IFAD Management to provide a self-assessment. The purpose of the self-assessment is to allow Management to provide their views to IOE on the subject being evaluated early on in the process. Given the differences in scope and focus of CLEs, there is no standard format for self-assessment.

Various methods and techniques are used for data collection and analysis. In the gender evaluation mentioned above, the enquiry methods were a mix of document reviews, a quantitative and qualitative survey, interviews and focus group discussions with IFAD headquarters staff and the staff of partners in borrowing countries, and country visits, consistent with the recommendation in chapter 2 to apply a mixed-methods approach.

Chapter 2 has more detailed guidance on methods and techniques. The following are some of the most commonly used methods in a CLE:

- Desk review; which includes a review of evaluation and research reports from relevant sources, documents from cofinanciers and peers, academic studies where relevant, and

BOX 12 Example of building blocks

IFAD's performance with regard to gender equality and women's empowerment

- An analysis of the evolution of gender-related concepts and development approaches, and a comprehensive documentary review of the policy and evaluation documents prepared by other development organizations (i.e. a benchmarking review);
- An assessment of IFAD's key corporate policy and strategy documents;
- A meta-evaluation of past IFAD-funded operations based on existing evaluative evidence, plus a review of recent country strategies and project designs, and five country visits to collect the perspectives of in-country partners and evidence on the ground about the evolving approaches and results of IFAD-funded projects; and
- A review of selected corporate business processes that have implications for IFAD's performance in promoting gender equality and women's empowerment in partner countries.

IFAD documents (strategic and policy framework and directives, COSOPs, findings from IOE evaluations, management reporting, Board papers, RIMS data, etc.);

- Interviews and focus groups to validate and triangulate findings from the desk review (with IFAD staff and stakeholders as well as with peers where pertinent and feasible);
- Surveys and self-assessments;
- Benchmarking studies of comparators and peers; and
- Case studies.

Value for money

To ensure value for money in each CLE, it is important to consider the cost of each building block and the evaluation instruments suggested, and their contribution to answer the evaluation questions. In particular, when country visits are proposed, the cost of these in relation to the criticality of the evidence that will be generated should be given serious thought, and alternative methods always be considered. The rationale for country visits and criteria for their selection should be clearly given in the approach paper.

Limitations

All evaluations have limitations and risks. These should be clearly identified in the approach paper and addressed to the extent possible in the evaluation design. They often relate to the following clusters of issues:

- Methodology – complexity, lack of standards;
- Resources – time and resources vis-à-vis the scope and timeline;
- Evidence – access to informants, low response rate on surveys and questionnaires, insufficient data (e.g. on costs); and
- Time lags – activities too recent to show any clear impact.

In the final report, the implication of the limitations faced by the evaluation must be clearly acknowledged, as must any changes vis-à-vis the approach paper. Indeed, evaluations are dynamic

and as the work progresses and new knowledge is generated, there is often a need to adjust the evaluation questions or aspects of the framework – such adjustments are desirable to ensure the most useful and pertinent evaluation outcome, but for accountability purposes they must be explained in the limitations section of the final report.

Process

Consistent with the definition, CLE's have corporate-wide implications and therefore often receive more attention in the organization than other types of evaluations with a narrower scope. This calls for a process with a strong element of inclusion, dialogue, feedback and stakeholder management. It is generally acknowledged that the more stakeholders have felt consulted during an evaluation, the more likely they are to use the evaluation findings and implement recommendations. When there is close coordination throughout the process with the most immediate users, as was the case for example with the replenishment evaluation, insights gained early in the process may also be acted upon even before the evaluation report is finished, thus enhancing the subsequent utility.

In terms of process, a CLE is generally conducted in six phases as listed below (see also figure 4). Further information on CLE process may be found in the Evaluation Process Guidelines.

Phase 1: Preparation and design

This phase includes the preparation of the evaluation approach paper and an inception phase. The approach paper is prepared well in advance with the aim of ensuring agreement on the objectives, scope, and evaluation design (including articulating the TOC, if this approach is used in the CLE) and key questions, process and timelines. The draft approach paper will be peer reviewed within IOE and by the SIAs. Thereafter, the document (including IOE and SIAs comments) is shared with IFAD Management for review and comments. Management will be given two weeks (ten working

FIGURE 4 **Corporate-level evaluation: phases, processes and key deliverables**

Phases	Processes	Key deliverables
Preparation and design	<ul style="list-style-type: none"> • Preparation of approach paper <ul style="list-style-type: none"> - Draft approach paper - IOE and SIA peer review - IFAD Management review - Evaluation Committee discussion • Inception phase 	<ul style="list-style-type: none"> • Draft approach paper • IOE and SIA comments • IFAD Management comments • Inception workshop for IOE staff
Desk review	<ul style="list-style-type: none"> • Documentation and data collection • Internal and external interviews 	<ul style="list-style-type: none"> • Desk review notes (for internal use) • Self-assessment report by IFAD Management
Data collection	<ul style="list-style-type: none"> • Country specific data collection 	<ul style="list-style-type: none"> • Country note (for internal use)
Analytical phase	<ul style="list-style-type: none"> • Information and data analysis • Triangulation of findings 	<ul style="list-style-type: none"> • Emerging findings workshop
Preparation of draft final report and review process	<ul style="list-style-type: none"> • Report-writing • IOE and SIA peer review • Draft final report shared with IFAD Management for comments 	<ul style="list-style-type: none"> • Draft final report • IOE and SIA comments • IFAD Management comments
Finalization, dissemination and follow-up	<ul style="list-style-type: none"> • Report finalization • Communication 	<ul style="list-style-type: none"> • Final report and audit trail • Management's response • Distribution of final CLE report • Evaluation Profile • Evaluation Insight

days) to provide their consolidated comments on the draft approach paper. Once IFAD Management comments are included, the document is discussed with the Evaluation Committee and enhanced based on the feedback from members.

The purpose of the inception phase is to further develop the evaluation's design, key questions, instruments for data collection, plans for country visits and timelines. An inception workshop may be organized with concerned IOE staff and consultants to develop a common understanding about the evaluation's methodology, overall process and the like. The approach paper may be fine-tuned, as and if necessary, at the end of the inception phase.

Phase 2: Desk review

During this phase the evaluation team will: (a) collect and analyse relevant external and internal documents and data; and (b) conduct interviews with IFAD Management and staff, and with external informants such as members of the Evaluation Committee and Executive Board. In addition and in line with the IFAD Evaluation Policy, during the desk review phase, IFAD Management will prepare a self-assessment report on the CLE's topic, which will serve as a critical input for IOE's independent evaluation analysis. IOE and IFAD Management will agree on the key questions and format for the self-assessment report.

Phase 3: Data collection

The focus of this phase is to capture evidence from key partners at the country level to strengthen the CLE's partnership dimension, analysis and conclusions. Country visits will be undertaken in most CLEs, although there might be instances when such visits would not be required, for example, if a CLE covers a very specific or mainly management-related topic (e.g. IFAD's investment policy). IOE will select the countries to be covered, following consultation with IFAD Management. The rationale for country visits and criteria for selecting countries – which will differ depending on the topic being evaluated and resources available – will be included in the approach paper.

Before launching the country visits, a communication is sent by IOE to the concerned governments announcing the evaluation exercise, describing the process, and proposing dates for the country visits. IOE should establish direct contact with concerned government authorities to agree on meetings and develop a programme for the field visit. The CPM and IFAD country office should be kept informed of any country visits; they may be called upon to facilitate mission preparations and follow up on IOE communications to the government. The IFAD country office should not be asked to prepare the programme of meetings or determine the field visits to be undertaken by the mission.

Phase 4: Analysis

Once the information and data have been collected, it is to be categorized, systematized, interpreted and analysed. The data will typically consist of quantitative and qualitative data, and various methods and techniques exist for structuring data to facilitate analysis (see chapter 2). Once analysed and at the stage where findings are beginning to emerge, it is important to triangulate findings to ensure that they are drawn on a sufficiently solid evidence base.

During this phase it is also recommended that an “emerging findings workshop” be organized to (i) engage stakeholders in the process; (ii) ensure

that findings resonate with staff and experts knowledgeable of the issue being evaluated; (iii) identify possible gaps or findings with insufficient evidence; and (iv) triangulate and possibly nuance key findings.

During the analytical phase it is important to keep the context analysis and TOC in mind to ensure that the analysis is not driven by the report template, but is guided by the evaluation questions.

Phase 5: Preparation of draft final report and review process

The primary activity in this phase is the report writing, including triangulation of evidence, identification of any weaknesses, and collection of additional evidence where needed. The draft final report will be peer reviewed within IOE and by the SIAs. Thereafter, the document (including IOE and SIAs comments) is shared with IFAD Management for review and comments, in line with the protocol for communication established. IFAD Management will be given around four weeks (20 working days) to provide their consolidated written comments on the draft final report.⁸⁵ During this period, a dedicated meeting to discuss the main findings and recommendations of the report with IFAD Management may also be organized.

Phase 6: Finalization and follow-up

The IOE lead evaluator is fully responsible for preparing the final report and its contents. Management comments are integrated, in line with the Evaluation Policy, into the final report and an audit trail of their comments is also produced by IOE. The final report and audit trail will be shared with IFAD Management.

Thereafter, IFAD Management will prepare a written Management's response (3-5 pages) for all final CLEs, which will be included in the final report when published.⁸⁶ To facilitate follow-up and to be consistent with good practice among peers, the Management's response should include (possibly as an annex) an action-oriented and time-bound

85 It is critical for IOE to receive consolidated comments from IFAD Management, not comments sent by different departments/divisions bilaterally to IOE. This will facilitate the inclusion of IFAD Management comments in the final report, and also avoids the risk that any possible inconsistent/contradictory comments are received by IOE from different divisions/departments.

86 The evaluation policy foresees that an ACP will be prepared for each CLE conducted by IOE. However, in line with the practice in the past years, a Management's response in lieu of the ACP will be prepared, since the latter normally requires the signature of the government counterpart, which does not apply in the case of a CLE.

matrix which clearly assigns responsibility and deadlines for follow-up. All final CLE reports are discussed at the Evaluation Committee together with the IFAD Management response, as well as in the Executive Board. IFAD Management will report on the implementation status and follow-up to CLE recommendations in the PRISMA. With regard to communication, dissemination and learning, see chapter 10.

Reports

Because purpose, theme, scope and methodology vary in evaluations, the structure of reports will also vary. An indicative outline of a CLE report is shown in box 13.

BOX 13 Indicative outline of a CLE report

- I. Background and the evaluation
 - A. Background
 - B. Objectives
 - C. Methodology
 - D. Process
 - E. Limitations
- II. Context (What are the regional/global context/ the strategic implications?)
- III. Description of IFAD's support to the issue being evaluated (What was intended?)
- IV. Assessment of performance/main evaluation findings (What was achieved?)
- V. Review of experience of other agencies (What can we learn from others?)
 - A. Multilateral organizations
 - B. Bilateral agencies
 - C. NGOs/academia
- VI. Conclusions and recommendations (Where to go from here?)



8

8. Impact evaluation

Background

There is growing global attention to and demand for better evidence of the results from development interventions. This has increased pressure on multilateral development organizations, including IFAD, to measure and report their impact in a rigorous manner.

The OECD's DAC defines impacts as "the positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended." Along the same lines, IFAD defines impact as "the changes that have occurred – as perceived at the time of evaluation – in the lives of the rural people (whether positive or negative, direct or indirect, intended or unintended) as a result of IFAD interventions." This definition is also consistent with the ECG's Good Practice Standards for the Evaluation of Public Sector Operations which defines impact evaluation as "an evaluation that quantifies the net change in outcomes that can be attributed to a specific project or program, usually by the construction of a plausible counterfactual."⁸⁷

The purpose of this chapter is to explain IOE's approach to impact evaluations and the corporate context, and provide guidance to staff and consultants conducting such evaluations. Its aim is not to serve as a detailed sourcebook or toolkit for designing and implementing impact evaluations, given there is a wealth of such publications and literature produced by academics and practitioners that is already available in the public domain (e.g. guidance by the UNEG, publications by the International Initiative for Impact Evaluations, etc.).

Impact evaluation in IFAD

For the first time, in 2011 the Fund made a commitment in IFAD9 to lift 80 million people out of poverty in the period 2013-2015. A similar commitment was made in IFAD10 (end-2014) for the period 2016-2018, reflecting the organization's continued commitment to rural poverty reduction through smallholder agricultural development in all developing member countries.

In order to rigorously measure the impact of IFAD activities and transparently report on the achievement of the above-mentioned targets, IFAD Management introduced an impact assessment programme in 2012. Coherent efforts are underway to conduct impact evaluations in a selected number of projects and programmes as well as mainstream the concept of impact evaluations across IFAD operations.

The approach adopted by Management to undertake impact assessments involves: (a) randomly selecting projects for rigorous ex post impact assessments with non-experimental designs; (b) purposively selecting a limited number of highly innovative projects in defined thematic areas that would be eligible for evaluations with experimental methods (randomized controlled trials); and (c) purposively conducting systematic reviews and meta-studies to strengthen the analysis and assessments of the above evaluations with benchmarks derived from other studies.

Impact evaluations by IOE

In response to the demand for more evidence on the impact of IFAD interventions, IOE introduced impact evaluations of completed operations in 2013 as a new evaluation product. This was agreed

87 ECG GPS, annex III.2: A Note on "Impact" and "Impact Evaluation" in the Public Sector GPS.

with the Executive Board, and is in line with the Fund's Evaluation Policy and the role of evaluation offices in other bilateral and multilateral development organizations.

Impact evaluations by IOE are performed only on completed IFAD-financed operations and from an entirely independent perspective. They cover all the evaluation criteria included in this Manual, with the rural poverty impact criterion receiving the greatest attention and resources. This distinguishes IOE impact evaluations from the impact assessments conducted by IFAD Management. The latter primarily focus on the impact criterion and also include ex ante impact assessments using experimental methods (random control trials), which requires impact assessment methods and processes to be embedded in programme design and the close involvement of Management in data collection and analysis throughout implementation.

With its hands-on involvement in impact evaluations of IFAD-financed projects and programmes, IOE also contributes to the generation of knowledge and good practices on methodologies and processes for conducting rigorous impact evaluations. Moreover, IOE contributes to the in-house debate on impact evaluation, for example, by reviewing and commenting on key guidance documents prepared by IFAD Management,⁸⁸ preparing written comments on selected impact assessment reports, and periodically undertaking comprehensive assessments on Management's broader efforts to conduct impact evaluation.

IOE has other evaluation products (CSPEs and PPEs) that also include the evaluation of rural poverty impact as a core criterion. However, impact evaluations differ from the assessment of impact in other IOE products, both in terms of their methodological approach and the overall evaluation process. In particular, impact evaluations are based on more rigorous and quantitative methods, including use of counterfactuals to address attribution issues. They will normally build on primary data collection, through structured and detailed large impact surveys, especially when the extent and quality of

data from project M&E systems is weak or unusable for rigorous impact analysis.

The assessment of impact in other types of evaluations will largely rely on data collected by IFAD Management and project authorities (e.g. through the RIMS and M&E systems), but also benefit from additional data and information collected by IOE during field visits, site inspections and discussions with beneficiaries and other key informants. Such evaluations will also generate primary data through rapid rural appraisals, semi-structured interviews, mini-surveys and other methods. In all evaluations – including impact evaluations – significant attention will be devoted to analysing the “why question”, in order to discern the explanatory factors of satisfactory or less than satisfactory performance (see chapter 2).

Selection of projects

The projects to be selected for impact evaluations by IOE will be guided by a comprehensive selectivity framework (see annex V). The purpose of the selectivity framework is to enhance the transparency in the selection and prioritization of projects for IOE impact evaluations.

Based largely on the selectivity framework, IOE will normally undertake impact evaluations of projects: (i) within three years of their completion date; (ii) that are not selected for impact evaluation by IFAD Management; (iii) that will also be included as part of the project portfolio analysis in forthcoming CSPEs, to enhance the latter's evidence base; (iv) that have innovative development approaches (e.g. institutional, social, technological) that merit deeper analysis and documentation; and (v) that offer enhanced opportunities for learning, on what works and what does not in promoting sustainable and inclusive rural transformation. As noted above, it is essential that projects selected by IOE for impact evaluations do not overlap with projects covered by IFAD Management, to ensure efficient uses of IFAD's overall resources.

In addition, based on an evaluability assessment, priority will be given to projects that have an adequate amount of usable self-evaluation data to

⁸⁸ For example, a draft Impact Evaluation Sourcebook (2014).

ensure that impact evaluations by IOE can be done in an effective and efficient manner. Availability of data will help reduce the costs and time taken for IOE to undertake impact evaluations.

Lastly, IOE will consider issues relating to external validity in the selection by evaluating comparable interventions across different countries and regions – thus assessing the relative effectiveness of alternative interventions in different contexts. This is particularly important given IFAD's business model and the recognition of the need to adjust to different and changing country contexts.

Objectives and methodological considerations

IOE's impact evaluations are geared towards strengthening organizational accountability and learning, with the ultimate aim of promoting sustainable inclusive development and rural transformation. In particular, impact evaluation by IOE aims to: (i) assess impact in a quantitative manner and be able to attribute impact to operations funded by IFAD, while also paying due attention to qualitative aspects that form the critical pillars of IFAD's development approach; (ii) apply and test innovative evaluation methodologies and processes; and (iii) generate valuable evidence for other evaluations done by IOE, such as CSPEs, CLEs and evaluation synthesis reports.

In addition to the methodological fundamentals described in chapters 2 and 3, this section presents the main methodological fundamentals for undertaking impact evaluations of development projects in the smallholder agriculture sector.

Evaluation criteria

The prime focus of impact evaluations is on assessing rural poverty impact, and in the case of IOE, the four impact domains, as articulated in chapter 3. But as mentioned before, IOE impact evaluations also include assessments and ratings across all other key evaluation criteria consistent with chapter 3.

Evaluation questions

Each impact evaluation will aim to respond to the following main questions:

- How much have beneficiary incomes changed since the project?
- Have household net assets changed due to the intervention?
- Did household food security and nutrition improve and by how much?
- Has the intervention contributed to enhance natural resources and environmental management, and resilience to climate change?

Impact evaluations may also assess other questions that will provide a more holistic appreciation of the results of the intervention. These questions should be defined during the design of the impact evaluation.

Theory of change

Impact evaluations will use a participatory theory-based approach to respond to the above questions and assess not only "if", but also "how" and "why" the programme has, or has not, had an impact on the target population in the programme area. This will be done by testing the validity of the assumptions in the intervention's explicit or implicit TOC (see chapter 2).

Therefore, a critical step in the design and methodological development of impact evaluations will be to either: (i) develop the intervention's TOC, if it is not explicit in the project/programme's design; or (ii) fine-tune the TOC as appropriate at the time of evaluation. The development or fine-tuning of the TOC at the time of designing the impact evaluation should include the active participation of project/programme stakeholders to obtain their insights and generate consensus about the project's intervention logic, and causal linkages along the results chain including the assumptions and hypotheses for impact achievement.

Impact indicator matrix

Given that impact evaluations focus on assessing the rural impact evaluation criterion, an impact indicator matrix should be developed based on the assumptions in the project's/programme's TOC to respond to the evaluation questions and assess the various impact domains enshrined in this Evaluation Manual.

The aim of the indicator matrix is to describe the effects of the programme at immediate outcomes, long-term outcomes and impact levels. The impact indicator matrix is a crucial step in the impact evaluation methodology, as it will guide the design of the baseline and end-line surveys for the collection of quantitative and qualitative primary data as well as the finalization of the evaluation framework that comprises the detailed evaluation questions, along with relevant sources of data/information. An example of the impact indicator matrix may be seen in the approach paper for the India impact evaluation (2014).⁸⁹ In the absence of robust baseline data, ex post impact evaluations should identify a counterfactual to establish attribution.

The counterfactual

A counterfactual is a situation or condition which hypothetically may prevail for individuals, organizations or groups where there was no development intervention. The counterfactual entails the identification of two groups: one that benefits from the project and one that does not; the first is commonly referred to in technical terms as the treatment group (i.e. the project beneficiaries), and the latter as the control or comparison group. It is important that members of the two groups share similar ranges of characteristics (e.g. in terms of socio-demographic indicators) as this provides greater reliability in attributing any possible impacts on the treatment group as a result of the intervention.

Methods

The core of impact evaluations is to identify a proper counterfactual and to be confident that impact can be attributed to IFAD operations. This may be

done by assessing the situation of the beneficiaries “before and after” and “with or without” the project, by comparing them with the counterfactual on a set of observable and unobservable socio-economic characteristics in order to assess impact and attribute it to a given operation. To do so, and given frequent lack of solid baselines and the nature of IFAD interventions – with special attention to empowerment of the rural poor, development of grass-roots institutions, and promotion of gender equality, and participatory processes – IOE in most cases will use a mixed-method non-experimental approach that combines quantitative and qualitative methods (see chapter 2). A common first step is to determine the sample size and sampling strategy to select a representative sample from both the treatment and comparison groups.

The calculation of a robust and representative sample depends on various factors such as: (i) the value of the variable of interest (i.e. the Poverty Head Count Ratio) observed in the target population at the time of the first survey (baseline); (ii) the estimated value of this same variable expected at the time of the final survey (end-line); (iii) the significance of the level of testing, which provides the cut-off point beyond which the null hypothesis is rejected; (iv) the statistical power level of detecting change, which determines the minimum sample size to detect differences in the two groups; and (v) the design effect, which is the ratio of the actual variance, under the sampling method actually used, to the variance computed under the assumption of simple random sampling.

Experimental methods, also called randomized control trials use randomization techniques at the outset of the project (usually through a lottery system) to sample the treatment and the control groups.⁹⁰ There are different methods⁹¹ to randomize a population, but a general requirement is that the socio-economic characteristics and the size of the two groups should be broadly equivalent to maximize the statistical degree of precision of the impact on the target group. Randomized experiments solve the problem of selection bias by generating an

⁸⁹ http://www.ifad.org/evaluation/public_html/eksyst/doc/approach/jctdp.pdf.

⁹⁰ The term “control group” is normally used in experimental designs, while the term “comparison group” is associated with non-experimental designs.

⁹¹ The different methods for randomization are shown at the following link: <http://www.povertyactionlab.org/sites/default/files/documents/Randomization%20Methods%20PDF.pdf>.

experimental control group of people who would have participated in a programme but who were randomly denied access to the programme. The random assignment does not remove the selection bias, but instead balances the bias between the participant (treatment) and non-participant (control) groups, so that it cancels out when calculating the mean impact estimate.

In ex ante impact evaluations based on experimental methods, baseline data are collected on the treatment and control groups before project inception to enable a thorough understanding of the characteristics of the treatment and control groups before the intervention is implemented.

In experimental methods, project authorities are required to continuously collect data and monitor progress on key indicators contained in the baseline survey and as defined in the indicator matrix, from both the treatment and control groups throughout implementation. This will allow them to generate the necessary data sets for thorough impact analysis at different points of implementation (e.g. at project mid-term and project completion). Such data can also be used by external entities interested in evaluating the impact of the project. However, this requires a very strong M&E system.

Given the rigorous approach to selecting the treatment and control groups and the frequency of primary data collection for generating the required data sets, experimental methods are considered the most robust methods to assess and attribute impact to a development intervention. However, they have cost and time implications, and might raise ethical considerations (given the purposive exclusion of a group of people from the project benefits) that need to be dealt with upfront.

In this regard, in order for the experiment to be transparent, ethical and fair, the evaluator must ensure that everyone has an equal chance of being in the treatment group to receive the programme. Methods of fairly selecting participants include using a lottery, phasing in a programme, and rotating participants through the programme to ensure that everyone benefits.⁹²

In **non-experimental methods** used in ex post impact evaluations, the participant and comparison groups are not selected randomly at the outset of the project but the comparison group is reconstructed ex post at the time of evaluation. A combination of data and statistical methods are needed to create a reasonable counterfactual (see below).

In order to determine ex post the changes that might have occurred as a result of the intervention and attribute impact, impact evaluations using non-experimental methods undertake at least two complementary analyses, one “before and after” and one “with and without”.

The “before and after” analysis will require the availability of reliable baseline data, and the collection of primary data on the participant group on similar indicators at the time of evaluation, yet baseline surveys are often missing, or if available, their reliability is questionable. Under such circumstances, the baseline situation should be reconstructed, using proven methods such as memory recall. The main limitation of the “before and after” approach is that it cannot distinguish between the intervention’s effects and other external effects (i.e. other interventions benefitting the same project population). Therefore, this method alone will not allow firm attribution of improvements or worsening conditions only to the intervention being evaluated.

The second analysis, the “with and without”, requires comparing the changes induced by the intervention on a set of observable variables (e.g. food security or income) or unobservable variables, on both the treatment group and the comparison group (i.e. those not benefitting from the intervention).

A variety of methods are used in non-experimental methods to ensure that the treatment and comparison groups are as similar as possible and to minimize selection bias. This includes propensity score matching, regression discontinuity design, difference in difference and instrumental variables.⁹³ A description of the different techniques are found in table 13.

92 <http://www.povertyactionlab.org/methodology/who/who-participates-randomized-evaluations>.

93 The World Bank handbook on impact evaluation provides an exhaustive description of non-experimental methods: <https://openknowledge.worldbank.org/bitstream/handle/10986/2550/599980PUB0ID181BLIC1009780821385418.pdf?sequence=1>.

TABLE 13 Experimental and non-experimental methods

Methodology	Description	Who is in the comparison group?	Required assumptions	Required data
Pre-Post	Measure how program participants improved (or changed) over time.	Program participants themselves – before participating in the program.	The program was the only factor influencing any changes in the measured outcome over time.	Before and after data for program participants.
Simple Difference	Measure difference between program participants and non-participants after the program is completed.	Individuals who didn't participate in the program (for any reason), but for whom data were collected after the program.	Non-participants are identical to participants except for program participation, and were equally likely to enter program before it started.	After data for program participants and non-participants.
Differences in Differences	Measure improvement (change) over time of program participants <i>relative to</i> the improvement (change) of non-participants.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program.	If the program didn't exist, the two groups would have had identical trajectories over this period.	Before and after data for both participants and non-participants.
Multivariate Regression	Individuals who received treatment are compared with those who did not, and other factors that might explain differences in the outcomes are “controlled” for.	Individuals who didn't participate in the program (for any reason), but for whom data were collected both before and after the program. In this case data is not comprised of just indicators of outcomes, but other “explanatory” variables as well.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <u>or</u> do not differ between participants and non-participants.	Outcomes as well as “control variables” for both participants and non-participants.
Statistical Matching	Individuals in control group are compared to similar individuals in experimental group.	<u>Exact matching</u> : For each participant, at least one non-participant who is identical <i>on selected characteristics</i> . <u>Propensity score matching</u> : non-participants who have a mix of characteristics which predict that they would be as likely to participate as participants.	The factors that were <i>excluded</i> (because they are unobservable and/or have been not been measured) do not bias results because they are either uncorrelated with the outcome <u>or</u> do not differ between participants and non-participants.	Outcomes as well as “variables for matching” for both participants and non-participants.
Regression Discontinuity Design	Individuals are ranked based on specific, measureable criteria. There is some cutoff that determines whether an individual is eligible to participate. Participants are then compared to non-participants and the eligibility criterion is controlled for.	Individuals who are close to the cutoff, but fall on the “wrong” side of that cutoff, and therefore do not get the program.	After controlling for the criteria (and other measures of choice), the remaining differences between individuals directly below and directly above the cut-off score are not statistically significant and will not bias the results. A necessary but sufficient requirement for this to hold is that the cut-off criteria are strictly adhered to.	Outcomes as well as measures on criteria (and any other controls).

Quasi-Experimental Methods

	Methodology	Description	Who is in the comparison group?	Required assumptions	Required data
Quasi-Experimental Methods	Instrumental Variables	Participation can be predicted by an incidental (almost random) factor, or “instrumental” variable, that is uncorrelated with the outcome, other than the fact that it predicts participation (and participation affects the outcome).	Individuals who, because of this close to random factor, are predicted not to participate and (possibly as a result) did not participate.	If it weren't for the instrumental variable's ability to predict participation, this “instrument” would otherwise have no effect on or be uncorrelated with the outcome.	Outcomes, the “instrument,” and other control variables.
Experimental Method	Randomized Evaluation	Experimental method for measuring a causal relationship between two variables.	Participants are randomly assigned to the control groups.	Randomization “worked.” That is, the two groups are statistically identical (on observed and unobserved factors).	Outcome data for control and experimental groups. Control variables can help absorb variance and improve “power”.

Source: <http://www.povertyactionlab.org/sites/default/files/documents/Experimental%20Methodology%20Table.pdf>.
 For more information see also <https://www.povertyactionlab.org/research-resources/>; and
<http://www.3ieimpact.org/en/evaluation/resources/impact-evaluation-resources/>.

Propensity score matching uses various single-difference estimators. Difference in difference uses panel data⁹⁴ to account for unobserved characteristics that are time invariant and compare treatment and control groups in terms of the difference in outcomes across treatment and control groups, before and after the programme intervention.

Well-designed non-experimental methods help also to address selection bias (i.e. when programme participants differ from non-participants in characteristics that cannot be observed by the evaluator). In propensity score matching, the comparison group is matched to the treatment group on observable characteristics by using the propensity score (i.e. the predicted probability of participation given observed characteristics),⁹⁵

while the instrumental variables method focuses on observable and unobservable characteristics.

When the comparison group is properly identified, non-experimental methods can also help to minimize spillover (i.e. when the non-target population is indirectly affected by the intervention through social and economic interaction with the target population) and contamination effects (i.e. when the beneficiaries are exposed to other interventions, for instance, from government and other donors). Methods to isolate possible contamination effects include the use of “tagging questions” in the impact survey questionnaire. These questions serve two main purposes: (i) to identify the direct beneficiaries of the intervention being evaluated; and (ii) to investigate whether the beneficiaries were exposed to other development interventions.

94 Panel data refers to multi-dimensional data frequently involving measurements over time.

95 Score matching is a very useful method when there are many potential characteristics to match between a sample of programme participants and a sample of non-participants. Instead of aiming to ensure that the matched control for each participant has exactly the same value of the control variables X, the same result can be achieved by matching on the predicted probability of programme participation, P, given X, which is called the propensity score of X. The range of propensity scores estimated for the treatment group should correspond closely to that for the retained sample of non-participants. The closer the propensity score, the better the match. A good comparison group comes from the same economic environment and is administered the same questionnaire as the treatment group by similarly trained interviewers (World Bank).

TABLE 14 Impact domains and analytical areas of focus

Impact domains	Analytical areas of focus	
	Quantitative	Qualitative
(i) Household income and assets	Data on: overall household income and household assets (including livestock, land)	Value of goods and services received in kind. Access to social services (education, health).
(ii) Human and social capital and empowerment	Data on gender equality and women's (and vulnerable groups in particular tribal communities) and youth empowerment	Changes in grass-roots organizations that have been facilitated by the project and their benefit to the poor, women, vulnerable groups, youth. Improved knowledge on watershed and crop/livestock management and adoption rate by gender. Improved knowledge on tribal rights by vulnerable groups (especially by women). Changes in gender roles, including participation in income generating activities, participation in community-level activities, group membership and leadership, division of workload, household food security and nutrition and land ownership.
(iii) Food security and agricultural productivity	Data on variability and average value of production as well as variability and average yields per hectare Data on commercialized production Data on child malnutrition	Improved access to input and output markets. Changes in pricing policies. Improved research and extension advice.
(iv) Institutions and policies		Changes in policies and pro-poor orientation of public agencies and private sector organizations.

Sampling strategy

The ability to detect changes between the treatment and comparison groups depends on the total sample size of each group. The sampling strategy is therefore a key component in the design of the impact evaluation,⁹⁶ which need to be developed with due consideration to the representativeness of the sample and the above-mentioned potential risk of spillovers or contamination.

In ex post impact evaluations, the sampling strategy, including the total sample size, will be determined at the time of the survey design, using information or estimates on the population's statistical characteristics (e.g. their levels of poverty, income, literacy, land holdings).

In order to take into account potential "spillover" effects, the "comparison group" will be identified for example by selecting communities that were initially earmarked for project support but eventually received no assistance. As an alternative or complementary option, households that did not receive project assistance, but are located in communities supported by the project could also be considered as comparison groups.

At times, spillover effects might be sufficiently important to merit trying to capture them. For instance, in some technology adoption programs beneficiary farmers often transfer technology to others. In such cases, the indirect effects can be larger than direct effects. To properly assess impact,

96 Daniel, J. N. 2011. *Sampling Essentials: Practical Guidelines for Making Sampling Choices*; Henry, G. 1990. *Practical Sampling*; Tourangeau, R., B. Edwards, T. Johnson, et ux. (Eds.), 2014. *Hard-to-Survey Populations*; Levy, P.S. and S. Lemeshow. 2013. *Sampling of Populations: Methods and Application*.

a sample should be designed that includes direct and indirect beneficiaries.

Use of mixed methods⁹⁷

The use of mixed methods – as also advocated by the Network of Networks on Impact Evaluation⁹⁸ – is most appropriate in smallholder agricultural interventions (as compared to health and education, for instance), given that multiple inputs, processes and impacts are required to achieve sustainable economic and social advancement.

Given their nature, it is expected that some impact domains may be better explored through quantitative analysis, others through qualitative, and some through a combination of the two. The focus of impact evaluations will be on quantitative analysis. Table 14 shows the four impact domains covered by IFAD impact evaluations, and the corresponding quantitative and qualitative analytic areas of focus.

Quantitative tools: the impact survey

Experimental methods require the impact evaluation to be embedded in programme design and the baseline to be constructed before the intervention starts. Data collection and analysis are conducted throughout implementation on both the treatment and control groups. In ex post non-experimental designs, on the other hand, the impact survey collects primary data at the time of the evaluation on both treatment and comparison groups. As such, it is the “core” of the impact evaluation. The main aim of the impact survey is to measure the indicators included in the indicator matrix towards the assessment of the four impact domains shown in table 14 (see also chapter 3). The results of the survey will inform the overarching impact evaluation report, which will be prepared once the impact survey data and analysis have been completed. It is worth noting that non-experimental designs can also be applied to impact evaluations planned ex-ante at the time of programme design in some cases, especially when experimental methods are not most suited due to various reasons – such as universal targeting, political considerations and programme design.

Qualitative tools

The qualitative component of the survey will provide information and analysis on topics for which the quantitative analysis is not suitable and will help probe into issues that emerge from a detailed review of existing documentation. The collection of qualitative data and information is particularly important, especially to assess the impacts of activities with intangible benefits (e.g. empowerment, promotion of participatory processes) that are common characteristics in IFAD operations. Qualitative data collection may take the form of a combination of participatory techniques (focus group discussions, participatory ranking exercises, interviews and other techniques that are deemed appropriate). The use of qualitative techniques, especially in presence of spillovers, will help understand the “why” question, providing the in-depth perspectives on the processes of change induced by interventions, how and why they worked, and under what conditions they might be replicated.

In particular, interviews can be individual or in focus groups or both; they can be structured, based on a questionnaire, or semi-structured, and different interview techniques can be applied.⁹⁹ In most impact evaluations interviews will be conducted with different stakeholder groups, including beneficiaries.

Triangulation

As mentioned in chapter 2, triangulation includes the use of three or more sources or types of information, or types of analysis to verify and substantiate an assessment. This is a standard methodological fundamental in all IOE evaluations. To enhance the validity and credibility of the impact evaluation findings, the information obtained from the different methods of data collection should be triangulated, so as to identify inconsistencies in different estimates and gain a deeper understanding of the reasons for these differences.

⁹⁷ For an in-depth description of quantitative and qualitative methods for data collection, see annex 2 of the IFAD9 Sourcebook on impact evaluations by IFAD’s Strategic Planning and Impact Assessment Division.

⁹⁸ See also Leeuw, F. L. and J. Vaessen. 2009. *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*, World Bank.

⁹⁹ Brinkmann, S. and S. Kvale. 2014. *Interviews – Learning the Craft of Qualitative Research Interviewing*.

Value for money

To ensure value for money in each impact evaluation, the cost of each building block and the evaluation instruments suggested, and their contribution to answer the evaluation questions, needs to be considered. In particular, because of the cost implications of the impact surveys, the cost of these in relation to the criticality of the evidence that will be generated should be given serious thought, and alternative methods should always be considered.

Limitations

All evaluations have limitations and risks. These should be clearly identified in the approach paper and addressed to the extent possible in the evaluation design. They often relate to the following issues:

- Methodology – complexity, lack of standards;
- Resources – time and resources vis-à-vis the scope and timeline;
- Evidence – access to informants, low response rate on surveys and questionnaires, insufficient data (e.g. costs); and
- Time lags – activities too recent to show any clear impact.

In the final report, the implication of the limitations faced by the evaluation must be clearly acknowledged, as must any changes vis-à-vis the approach paper. Indeed, evaluations are dynamic and as the work progresses and new knowledge is generated there is often a need to adjust the evaluation questions or aspects of the framework – such adjustments are desirable to ensure the most useful and pertinent evaluation outcome, but for accountability purposes they must be explained in the limitations section of the final report.

Process

Impact evaluations are carried out by an IOE lead evaluator and the collaboration of a SIA, whose role is to provide comments on the draft approach paper and draft final report. S/he will prepare a short report (two-three pages) on the final report to attest to the

quality of the impact evaluation including the overall process followed. The SIA report will be included in the final impact evaluation report published by IOE.

The impact evaluations generally include a number of phases, which are described below. It is important to note that the phases are not strictly sequential, but in many cases iterative, conducted in parallel and partly overlapping.

Impact evaluations will take a maximum of 12 months to carry out, from start to finish, taking into account the time required for primary data collection and analysis and if necessary, the time needed for a bidding process to hire a company for such purposes.

Phase 1: Preparation and design

At the outset of the process, IOE formally advises PMD and the concerned government in writing that it intends to undertake an impact evaluation, specifying the main objectives, process and tentative timelines. The CPM and IFAD country office (where applicable) assist IOE by providing contact details of key government officials.

The preparation and design of the evaluation comprise three key activities: the preparation of the approach paper, the preparatory mission, and the development of the detailed terms of reference for the impact survey.

The approach paper

The evaluation design is described in the approach paper which includes: objectives of the evaluation; evaluability assessment guiding the selection of the quantitative approach to be used; evaluation questions; the TOC; evaluation criteria; overall methodology and specific quantitative methods to attribute impact to the initiative (sampling strategy, data collection and analysis methods, etc.); methodological limitations; and the process, including the timeline.

The approach paper is peer-reviewed in IOE. Once IOE and the SIA comments are added, it is shared simultaneously for review and comments with IFAD Management and the government of the country concerned and finalized thereafter.

Preparatory evaluation mission

The preparatory mission is carried out by the lead evaluator at the outset of the evaluation process to: (i) brief national and local authorities; (ii) meet and discuss with representatives of potential institutions and companies suitable to conduct the impact survey; and (iii) hold discussions with project staff to assess the availability and usability of baseline data and to further develop the TOC.

Terms of reference for the impact survey

A critical part of the design stage is the preparation of the terms of reference for the impact survey and the recruitment of the institution/company to implement it. The company hired by IOE to conduct the impact survey is contracted through an institutional contract or a bidding process. Once the company is recruited, the impact questionnaire is finalized and a joint fine-tuning of the methodology is carried out to ensure the mutual understanding of the terms of reference, including the scope, methodology, timeline and deliverables. The qualitative tools are also developed at this stage.

Phase 2: Desk review

This phase consists in reviewing external and internal documents and may be conducted in parallel with the evaluability assessment and the drafting of the approach paper. In terms of the document review and synthesis, different methods exist (see chapter 2) and should be carefully considered. A well selected method can play a pivotal role in collecting existing evidence to deepen the credibility and validity of an impact evaluation, and contribute to future knowledge building while meeting the information needs of stakeholders. Specifically, these methods can serve two major purposes:

- Strengthen external validity by evaluating comparable interventions across different countries and regions – thus assessing the relative effectiveness of alternative interventions in different contexts; and
- Help refine the hypotheses in the results chain.

Phase 3: Field mission and data collection

The research instruments, and in particular the questionnaire for the collection of primary data, shall be piloted in the programme areas before the data collection starts. The piloting will serve two main purposes: (i) to check the appropriateness of the questions and the length of the questionnaire; and (ii) to ensure the quality and reliability of the tools. In this phase, the IOE lead evaluator organizes a field visit to supervise the pre-testing of the research instrument in the field. The data collection starts once the research tools are finalized and the enumerators are trained. This phase may take from four to eight weeks, depending on the sample size.

In preparing the field mission, IOE informs PMD and requests government authorization to conduct the evaluation, after which a mission schedule is prepared according to their availability. IOE will establish direct contact with concerned government authorities to plan meetings and develop a programme for the field visit. The IFAD country office may be called upon to facilitate the mission preparation and to follow up on the IOE communication to the government. The IFAD country office should not be asked to prepare the programme of meetings or determine the field visits to be undertaken by the mission.

Phase 4: Data analysis

Once the impact survey is completed and other impact information and data have been collected, the information is categorized, systematized, interpreted and analysed. Once analysed and at the stage where findings are beginning to emerge, it is important to triangulate findings to ensure that they are drawn on a sufficiently solid evidence base. For transparency and accountability purposes, staff and consultants, including the institution carrying out the impact survey, should maintain data files and syntax files for the record to be shared with IOE, and clearly explain the analytical steps and methods used. This is important to be able to assess the validity of the findings and trace the logic from fact over finding to conclusion and recommendation.

If required, at the beginning of this phase, the lead evaluator might organize another field mission to identify possible gaps or findings with insufficient evidence, validate and nuance preliminary evaluation findings, and further engage with key stakeholders.

Phase 5: Preparation of draft final report and review process

The main activity in this phase is the report writing, including triangulation of evidence, identification of gaps, collection of additional evidence where needed, and internal quality review.

The evaluation report should include a section on limitations and the way in which the evaluators tried to overcome them. As discussed in the methodology section, these might include: lack of or unreliable baseline data, biases in sample selection, difficulties in isolating benefits generated by external factors on both the participant and/or control groups, weaknesses in capacities to collect primary data through impact surveys, etc.

Once the draft report is ready it will be shared with the SIA for his/her comments. At the same time, the lead evaluator will organize an internal IOE peer review, which will follow IOE's peer review guidelines.¹⁰⁰ Comments on the draft report may be provided in writing or verbally, or both. The peer review will focus largely on the overall methodological rigour, robustness of the analysis, evidence trail, and findings and recommendations.

Once the comments from the IOE peer review, including the SIA, are incorporated in the draft report, the report is shared for review and comments with IFAD Management and the concerned government.

IFAD Management will transmit its written comments to IOE in a consolidated manner, rather than individual divisions and/or departments sending comments directly to IOE. Similarly, the main coordinating institution at the country level (e.g. ministry of finance or agriculture) should collect comments from concerned partners and send consolidated comments to IOE on the draft impact evaluation report.

Phase 6: Finalization, dissemination and follow-up

Management and government comments are integrated into the final report and an audit trail of their comments is also produced by IOE. The final report and audit trail will be shared with IFAD Management and the government. No ACP is prepared for impact evaluations. However, IFAD Management will prepare a written Management's response on the impact evaluation report, which will be included in the final report at the time of publication by IOE. Moreover, IFAD Management will report in the PRISMA on the implementation status and follow-up to impact evaluation recommendations made to IFAD (and not on those directed to the government) as it does for the ARRI, CLEs, CSPEs and PPEs.

Once finalized, the impact evaluation is formatted and submitted to the Evaluation Communication Unit of IOE for editorial quality assurance, web publishing and dissemination (see chapter 10 on communication, dissemination and learning).

All final impact evaluation reports, together with the written IFAD Management's response, are discussed by the Evaluation Committee. Upon the request of the Evaluation Committee, such reports may also be discussed at the Executive Board.

Reports

The evaluation report will contain an executive summary covering the main lessons learned. This summary should not be more than two pages long and should be written in the same language as the main document. The report is written in the IFAD official language (Arabic, English, French or Spanish) of the recipient country. The executive summary will be translated into English if the language used for the impact evaluation is Arabic, French or Spanish. Both the original language and English version of the executive summary and IFAD Management's response will be included in the final report.

¹⁰⁰ IFAD 2012, Internal Peer Review Process and Procedures.

9. Evaluation synthesis report

Background

The Peer Review of IFAD's Office of Evaluation and Evaluation Function by the ECG recommended that IOE, in keeping with practice in other IFIs, introduce a new evaluation product to further strengthen the use of evaluation findings, learning and feedback loop, namely evaluation syntheses. Evaluation synthesis reports were introduced by IOE in 2012 and are grounded in IFAD's Evaluation Policy: "IOE shall also prepare evaluation syntheses, which will identify and capture evaluative knowledge and lessons learned on a certain topic from a variety of evaluations produced by IFAD and the evaluation units of other organizations. These syntheses will be supplemented by lessons from academic literature and targeted interviews to promote learning and the use of evaluation findings."

The main aim

Evaluation synthesis reports are very different from other IOE products, as they are prepared to primarily promote learning, collective reflection and improve IFAD's development effectiveness. The report is a knowledge product and a means to consolidate and share acquired knowledge and strengthen IFAD's evaluation feedback and learning loops. Taking stock of findings from previous independent IOE evaluations, they aim to bring together lessons from IFAD evaluations while also capturing evaluation-based lessons from other IFIs, United Nations organizations and bilateral agencies on specific themes.

Evaluation synthesis reports are knowledge products that aim to enhance the general understanding of a particular topic; this level of abstraction makes them more useful in highlighting

the strategic implications of findings and raise strategic issues for further consideration by IFAD Management and the governing bodies. They facilitate wider use of evaluation findings by identifying and capturing accumulated knowledge, lessons learned and good practices on common themes across a variety of situations and sources. Synthesizing existing evaluation evidence allows evaluation synthesis reports to contribute to decision-making processes in an effective way, especially when there is neither adequate time nor resources to undertake a full-fledged evaluation.

It is important not to confuse evaluation syntheses with other similar type of exercises such as systematic reviews or meta evaluations. A meta evaluation is an evaluation of evaluations, and a systematic review implies the use of a rigorous protocol to summarize evidence around a research question, often as part of an impact evaluation.

Specific purposes vary but, being primarily for learning, they are an important part of IFAD's evaluation learning loop and as such a key input to design of new policies, strategies, operations or business processes. Evaluation synthesis reports are key IOE knowledge products and contribute to knowledge generation and learning. They also contribute to assessing IFAD's "learning accountability", that is, how well IFAD internalizes lessons and knowledge generated from previous evaluations in new policies, strategies and operations.

Added value

These reports are different from other major IOE evaluations products especially CLEs for three reasons: (i) they collect, consolidate and analyse

existing evaluative evidence to crystallize key lessons and good practices for improving IFAD's operational and corporate performance and effectiveness; (ii) being a knowledge product, they imply closer engagement with Management than traditional evaluation products; and (iii) they are prepared using more limited resources and time as compared to, for example, CLEs and CSPEs. They are, therefore, a key means of enhancing the value for money from IOE activities. Furthermore, these reports may be prepared jointly with evaluation offices in other organizations, including the United Nations Rome-based agencies, seeking to enhance their knowledge on an issue of shared interest, thus also contributing to a stronger partnership and shared learning with peers and partners.

Evaluation syntheses, more than any other IOE product, tend to feed into various corporate or strategic processes or products where knowledge and lessons on a specific topic at a specific moment is needed. In the past, evaluation syntheses have preceded important CLEs (e.g. on the IFAD supervision policy), have fed into other processes such as the replenishment process (e.g. see evaluation synthesis report on IFAD's role in MICs, or into other strategic or global processes, such as the United Nations International Year of Cooperatives (see the evaluation synthesis report on IFAD's Engagement with Cooperatives) where the synthesis was an input to IFAD's joint reflection with FAO and the World Food Programme on approaches to engaging with cooperatives.

Selecting topics

In order to enhance the transparency and formality of the selection of topics for evaluation synthesis reports, IOE developed a selectivity framework (see annex V)

to assist in the construction of its annual work programme. One of the most important drivers in deciding to embark on an evaluation synthesis report is the availability of a minimum amount of existing IOE evaluative evidence on the proposed topic.

The framework takes into account the suggestions by IFAD's Management and governing bodies, and filters the topics against a standard set of guiding questions. Apart from the criteria indicated in the selectivity framework, topics may also be suggested by selected divisions or departments when a need arises for knowledge on a specific topic.

Before a decision is taken on whether to prepare an evaluation synthesis report on a given topic, an evaluability assessment is carried out to ensure there is a fair amount of evaluative evidence to prepare the synthesis. However, on a case-by-case basis and as and when needed, the report will also draw on lessons embedded in evaluations done by evaluation offices of other development organizations.

Objectives and methodology

As a general rule, the objective of an evaluation synthesis report will be to generate lessons learned, good practices and recommendations that can further strengthen the design and implementation of IFAD policies, strategies and operations on the topic chosen for the synthesis. However, as purposes vary, the specific objectives are generally articulated to support this purpose, as in the example in box 14. The recommendations in evaluation synthesis reports should be strategic and add value, and not duplicate the recommendations previously made by IOE in evaluation reports that informed the synthesis.

BOX 14 Example of objectives of a previous IOE evaluation synthesis report

IFAD's Engagement in Middle-income Countries 2014:

- Generate lessons and insights on opportunities and challenges for IFAD's engagement in middle-income countries;
- Identify issues for further reflection on the strategic directions, priorities and instruments for IFAD's future engagement in middle-income countries.

Evaluation approach and design is described in the approach paper prepared at the outset of the process. Evaluation synthesis reports are primarily desk-based, although resources permitting and on an exceptional basis, they may benefit from a country visit.

Generally a qualitative approach is used to synthesize findings from evaluation reports and related sources, but where relevant, and where time and resources allow, more quantitative approaches may also be applied (see also chapter 2 on methodological fundamentals). Given that filtering, categorizing and organizing information is a central element of the design, the rigour of this process is critical for the quality and ultimate credibility of the review. The methodology is usually divided in four phases: (i) screening of a long list of evaluations, (ii) codifying and extracting data for analysis; (iii) analytical phase; and (iv) drafting and presentation of report.

This process should be guided by a well-defined question or set of questions for the synthesis. Furthermore, clear criteria must be established for the inclusion or exclusion of evaluations, and a critical appraisal made of the existing evidence-base to ensure validity of findings.

The approach paper will include the following elements: (i) rationale for selecting a particular topic; (ii) objective; (iii) scope; (iv) key questions; (v) methodology including criteria for the selection of evidence and process to ensure the validity; (vi) process and timeline; (vii) risks and limitations; and (viii) dissemination.

The use of evaluation criteria is different in evaluation synthesis reports as compared to most other IOE evaluations because the focus is squarely on learning. Generally speaking, they will follow four criteria, namely: relevance, effectiveness, efficiency and sustainability, but will not provide ratings, as the evaluative evidence is variable and not always subject to aggregation. There might be instances where one or more of these criteria are not covered or other criteria included (e.g. on innovation and scaling up). In sum, the criteria to be covered in an

evaluation synthesis report will be clarified in the approach paper.

Building blocks

To achieve its objectives, an evaluation synthesis report typically draws on a number of different components. The following are the most commonly used:

- (i) **Review of IOE evaluations:** This step includes reviewing findings first from IOE evaluations, in particular lessons from the project evaluations, CSPEs, CLEs and the ARRI, and secondly, from evaluations done by evaluation units of other organizations.
- (ii) **Review of other evaluations:** This step includes a review of evaluation reports from other development organizations, including multilateral and bilateral. To avoid information overload and ensure the cost effectiveness of the exercise, clear criteria for inclusion must be established, taking due account of the scope and purpose.
- (iii) **Benchmarking with peers:** Given IFAD's position among IFIs, for some topics, a benchmarking or simple comparison with peers may be a relevant building block. The main aim of the benchmarking will be to learn from the experiences of others, and examine how that may be of relevance, taking into account IFAD's specific mandate, development approaches and operating model.
- (iv) **Organizational review:** How IFAD supports organizationally the issue under review may be one of the dimensions studied and therefore a specific review of relevant CLEs and an assessment of resources (human and financial) allocated to address the issue, may be a building block for the evaluation synthesis report.
- (v) **Interviews:** To validate and triangulate findings from the desk review, interviews and focus group discussions are usually held with IFAD staff and stakeholders, as well as with peers where pertinent and feasible.
- (vi) **Under exceptional circumstances,** focused country visits can be organized on a very selective

basis, especially to collect additional data and validate emerging hypothesis and findings.

Limitations

All evaluations have limitations and risks. These should be clearly identified in the approach paper and addressed to the extent possible in the evaluation design. For evaluation synthesis reports they often relate to the following clusters of issues:

- Insufficient evidence – small number of IOE evaluations with findings on the subject being studied and hence over-dependence on external material;
- Quality of evidence – wide variety of quality and independence of evidence from different sources;
- Methodology – difficulty in categorizing and organizing findings from complex analysis; and
- Resources – time and resources vis-à-vis the scope and timeline.

In the final report, the implication of the limitations faced by a synthesis must be clearly acknowledged, as must any changes vis-à-vis the approach paper. Indeed, evaluations are dynamic and as the work progresses and new knowledge is generated there is often a need to adjust the evaluation questions or aspects of the framework – such adjustments are desirable to ensure the most useful and pertinent evaluation outcome, but for accountability purposes they must be explained in the limitations section of the final report.

Process

The process for undertaking evaluation synthesis reports is described below. Evaluation synthesis reports take between six to eight months to complete, from start to finish. It should also be noted that the phases below are not strictly sequential in the process, but to some extent iterative:

(i) Preparation; (ii) Desk review of all relevant documentation; (iii) Interviews with managers, relevant staff and external informants; (iv) Analyse information, validate hypotheses and triangulate key

findings; (v) Preparation of draft final report and review process, including to receive feedback from IFAD Management and staff during one or several “emerging findings” workshops; (vi) Finalization, dissemination and follow up.

Phase 1: Preparation

IOE may hire one or more consultants to help in the preparation of the evaluation synthesis report. However, the designated IOE lead evaluator has the overall responsibility for the process and all deliverables produced, including the final document.

The approach paper is prepared well in advance with the aim of ensuring agreement of the objective, scope, and review design and approach. It also outlines the process and timeline and hence is key in ensuring buy in and consensus among all involved. Comments from the IOE SIA should also be sought. Once the internal peer reviewers’ comments (including those of the SIA) are added to the draft approach paper, the same is shared with IFAD Management for review and comments.

Phase 2: Desk review

This is the fundament for the work, and key activities include filtering, categorizing, and organizing the information to facilitate a clear and strong analysis, and identifying issues that need further exploration through interviews. Methodologies for desk reviews are described in chapter 2.

Phase 3: Interviews

People to be interviewed will have been identified in the approach paper. Interviews can be individual or in focus groups, or both. They can be structured, based on a questionnaire, or semi-structured, and different interview techniques can be applied. In most evaluation synthesis reports, interviews will be conducted mainly with IFAD staff and Management, but may also be with partners and specialists in the areas being evaluated. Depending on the topic, IFAD Member State representatives might also be interviewed in person or through appropriate virtual

means. If country visits are undertaken, beneficiaries, key partners and project staff would also be interviewed by IOE.

Phase 4: Analysis

This is the core of the work, and the analysis is the basis for findings that are nuanced, validated and triangulated using information from the desk review, site visits (if country visits are undertaken) and interviews.

Phase 5: Preparation of draft final report and review process, including emerging findings workshop

Following the above steps, the draft final report, inclusive of recommendations, is prepared by the lead evaluator. Once the draft is complete, a peer review process within IOE is organized to review the document (see internal peer review guidelines). The draft is also shared at this stage with the SIA for comments. The lead evaluator will prepare the revised draft report, inclusive of comments by the internal peer reviewers and the SIA.

Emerging findings workshop

As mentioned above, a fully consultative process will enhance the usefulness of the synthesis and the emerging findings workshop is a critical step, aimed at discussing the draft final report once it has been peer-reviewed within IOE and the corresponding comments included in the draft final report. Such workshops will normally be organized at IFAD headquarters and include the participation of IFAD Management and staff, IOE staff, consultants, the SIA, representatives of the United Nations Rome-based agencies and others concerned. It is important that the Associate Vice-Presidents of PMD and the Strategy and Knowledge Department attend the session. Resources permitting, representatives from developing member countries might be invited to participate. The main background document – the draft evaluation synthesis report – should be shared with all participants ahead of the session. The comments generated at the workshop will be used to finalize the report.

Management comments and response

The draft final report will be shared with IFAD Management, who will provide its written consolidated comments to IOE, rather than individual division/departments sending comments directly to IOE.

Senior independent adviser

Moreover, IOE will seek the collaboration of one SIA for the preparation of the evaluation synthesis report. The role of the SIA is to comment on the draft approach paper and draft final report, and if resources permit, participate in the emerging findings workshop. Lastly, s/he will prepare his/her independent report (2-3 pages) based on the final report, which will be included in the final document at the time of publication.

Phase 6: Finalization, dissemination and follow-up

IOE will integrate Management comments in the final report and produce an audit trail illustrating how these were considered. The final report and audit trail will be shared with IFAD Management for information.

Thereafter, IFAD Management will prepare a written response (2-3 pages) on the final evaluation synthesis report, which will be included in the final report.

Once finalized, the evaluation synthesis report is formatted and submitted to the Evaluation Communication Unit of IOE for editorial quality assurance, web publishing and dissemination (see chapter 10 on communication, dissemination and learning).

All final evaluation synthesis reports, together with the written IFAD Management's response, are generally discussed in the Evaluation Committee. Upon request of the Evaluation Committee, the reports may also be discussed in the Executive Board.

Reports

Because purpose, theme, scope and methodology will vary between syntheses, the structure of the reports will also vary accordingly. The final report should be around 30-40 pages long and include a succinct executive summary covering the salient findings and recommendations included in the main report. The final evaluation synthesis report will also include a foreword by the Director of IOE.

An indicative outline of a report is presented in box 15.

All evaluation synthesis reports will be written in English. However, the acknowledgements, foreword, executive summary and IFAD Management's response will be translated and included in the final report in all IFAD official languages. As with other IOE evaluations, IFAD Management will report in the PRISMA on the follow-up actions taken and implementation of the recommendations from the report.

BOX 15 Indicative outline of an evaluation synthesis report

- I. Background
- II. Evaluation objectives, methodology and scope
- III. Description of IFAD's support to the issue being evaluated, including the relevant aspects of IFAD's strategic and policy framework (What was intended?)
- IV. Main findings from the synthesis (What was learned by IFAD?)
 - A. Objectives and general achievements
 - B. Problems and risks in implementation
 - C. Modifications in approaches and types of assistance
- V. Lessons from "other agencies" (What can we learn from others?)
 - A. Multilateral organizations
 - B. Bilateral agencies
 - C. NGOs/academia
- VI. Conclusions and recommendations (How might we do better?)



10. Communication, dissemination and learning

Learning

A thorough and rigorous evaluation analysis and the production of a good report are not enough for an evaluation to be useful in terms of learning. Far more needs to be done to capitalize on the knowledge generated through the evaluation process.

Today, there is strong awareness of the learning potential in the process itself right from the start – lessons and insights are generated as evaluators ask questions, probe issues, and present emerging findings for discussion with partners and stakeholders. Michael Quinn Patton argues that research on evaluation demonstrates that: “Intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings [and that] they are more likely to understand and feel ownership if they’ve been actively involved. By actively involving primary intended users, the evaluator is preparing the groundwork for use.”¹⁰¹

IOE audiences include: (i) IFAD Management, staff and consultants (ii) IFAD governing bodies; (iii) in-country stakeholders (e.g. government and project staff and beneficiaries); (iv) international partners and evaluation practitioners (e.g. Swiss Agency for Development and Cooperation, ECG and UNEG); (v) rural development practitioners; and (vi) the general public and the media.

To maximize the learning potential, IOE takes a participatory approach. A CLP is established for CLEs, CSPEs, evaluation synthesis reports and impact evaluations. The CLP is composed of the main stakeholders of the evaluation, who contribute to the process from the outset and throughout the different stages of the evaluation. At the same time,

the stakeholders can deepen their understanding of the evaluation findings and recommendations leading to a higher rate of adoption of the evaluation output.

Moreover, IOE’s evaluation approach aims to transform each evaluation into a systematic and operations-oriented learning exercise, emphasizing “the need for evaluations to finish not merely with a publication, but with an understanding among the evaluation’s partners to adopt specific findings and recommendations.” (IFAD 2000). With this objective, at the completion of an evaluation, a Management’s response¹⁰² or an ACP¹⁰³ is prepared. The status of the adoption of the recommendations is reported through the PRISMA, which is submitted to the Evaluation Committee and Executive Board.

Feedback activities and dissemination, however, are integral parts of the whole evaluation process. And dissemination to facilitate the sharing of lessons and uptake of recommendations remains important and is much about identifying tools and platforms to facilitate sharing. IFAD’s Evaluation Policy clearly states that all IOE products should be disclosed to the public and disseminated widely; this is critical to meet IFAD’s accountability obligation, but also to ensure learning from evaluations.

This chapter is designed to: (i) guide evaluators in using communications and outreach to maximize the impact of evaluation products; and (ii) provide a comprehensive overview of the diverse set of communication tools available to disseminate evaluation findings.

101 Patton M. Q, *Utilization-focused Evaluation*, 2008.

102 At the end of a CLE, IFAD Management will prepare a written response. As at 2015, a Management’s response will also be prepared for PPEs and evaluation synthesis reports.

103 CSPEs end with an ACP, which includes the agreement of IFAD Management and the concerned government on the main evaluation findings and their commitment to adopt and implement the evaluation recommendations within specified time frames.

Communication tools and dissemination

Disclosure of evaluation results is a core principle of IFAD's Evaluation Policy: "IOE shall ensure that all evaluation reports and other evaluation products are disclosed to the public and disseminated widely using electronic and other media." And it goes on to stress that accountability through evaluation analysis requires that successes, unexpected results, shortcomings and failures highlighted during the evaluation be disclosed to relevant partners and the general public without interference from any vested interest, thus stressing the independence aspect of IOE's work.

On the more traditional dissemination agenda, two key factors deserve attention when discussing dissemination. One is the emergence of new information technology-based and virtual platforms for sharing lessons from all types of evaluations and evaluation products; the other is the recognition of the need to more systematically share findings with the intended beneficiaries of IFAD assistance and with media and civil society in developing countries. This requires building up networks of contacts and planning ahead to reach relevant partner audiences effectively, as well as translating the evaluation products into local languages. In this respect, the integration in IOE of a dedicated evaluation communications unit (ECU) is a considerable strength.

The ECU undertakes communication functions, with a view to enhancing IOE's profile as a knowledge producer and to reach out and share evaluation learning. The Unit implements a standard dissemination approach for each type of product, as described below (some information on dissemination steps common to more than one type of evaluation report are repeated to render each section as a stand-alone). A targeted dissemination strategy may also be developed for specific evaluations, or for specific audiences, especially the beneficiaries. The most commonly used communications tools are: publications and other material (Profiles, Insights, overviews, reports, newsletters), Internet (website, social media), media relations (press releases),

audio-visual tools that are particularly suited to reaching the main beneficiaries, and events (in-house and international). The ECU collaborates with IFAD's Communications Division and is consistently exploring the use of innovative communications techniques, strategies and instruments to improve IOE communications and learning tools.

Publications

Evaluation Insights

Insights are two-sided brochures of approximately 800 words and focus on one learning issue emerging from an evaluation. They primarily serve to direct attention to critical learning hypotheses and form the basis for further debate among professionals and policy-makers at IFAD and outside the institution, project staff, development practitioners and other partners. Insights are produced for each CLE and CSPE, and evaluation synthesis reports. They may also, on a case-by-case basis, be produced for impact evaluations and PPEs, but not for PCRVs. Insights are disseminated to all audiences, together with the respective evaluation report.

Evaluation Profiles

Profiles are also two-sided brochures (also around 800 words) and contain a user-friendly overview of the main findings and recommendations of the evaluation. A profile is usually prepared for the ARRI, CLEs and CSPEs, impact evaluations and evaluation synthesis reports. IOE may also prepare a Profile for PPEs but not for PCRVs. Profiles are disseminated together with the respective evaluation report. Profiles and Insights are normally only produced in English, although they may be translated into other official IFAD languages, on a case-by-case basis.

Overview booklets

This short format publication is intended for those who do not have time to read the whole report. It comprises an overview or summary and IFAD Management's response or the ACP, depending of the type of evaluation.

Evaluation reports

Project completion report validations. PCRVs are prepared in English only, and no translations are foreseen. Once finalized, all PCRVs are posted on the independent evaluation section of the IFAD website. No other forms of dissemination are normally foreseen.

Project performance evaluations. The communication plan for a PPE is included in the corresponding approach paper. The PPE is usually written in the IFAD official language of the concerned country. The executive summary, preface and IFAD Management's response will be translated into English if they are written in another official language, and should be included in the final report. The lead evaluator may also prepare an evaluation Profile or Insights, on a case-by-case basis.

Once finalized, the report (and Profile/Insights, if available) is submitted to the ECU for editorial quality assurance, web publishing and dissemination. The ECU will post the PPE on the IFAD intranet (log-on screen), the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD Facebook page, and Twitter and Yammer accounts. The link to the report is also included in the UNEG database of evaluation reports. This is a public repository centralizing all evaluation reports of UNEG members, available at <http://www.uneval.org/evaluation/reports>.

The report is shared electronically with internal and external audiences, including IFAD Management and staff, IOE staff and consultants, members of the Executive Board, UNEG, ECG, the concerned IFAD regional network (e.g. IFADAsia, FIDAFRIQUE, KARIANET), evaluation associations and other networks according to the geographical region. A limited number of hard copies are printed and placed in IOE's display case for IFAD staff and visitors. Printed copies are available upon request.

The lead evaluator disseminates the evaluation report to government officials and other in-country partners. If there is an IFAD country office, it supports IOE in disseminating hard copies of evaluation reports, including to subnational authorities and

related partners in remote areas. On an exceptional basis, and resources permitting, a learning event may also be organized at the country level.

Impact evaluations. The communication plan for impact evaluations should be included in the approach paper. Impact evaluations are usually written in the IFAD official language of the concerned country. The executive summary, preface and IFAD Management's response will be translated into English if they are written in another official language, and should be included in the final report. The lead evaluator will prepare an Evaluation Profile in all cases. An Insights may be produced on a case-by-case basis.

Once finalized, the report, Evaluation Profile and Insights (if available) are submitted to the ECU for editorial quality assurance, web publishing and dissemination. The ECU will post the impact evaluation products on the IFAD intranet (log-on screen), the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD's Facebook page, and Twitter and Yammer accounts. The link to the report is also included in the UNEG database of evaluation reports. This is a public repository centralizing all evaluation reports of UNEG members, available at <http://www.uneval.org/evaluation/reports>.

The ECU will share the final documents electronically with internal and external audiences, including IFAD Management and staff, IOE staff, consultants, members of the Executive Board, UNEG, ECG, the concerned IFAD regional network, evaluation associations and other networks according to the geographical region. A limited number of hard copies are printed and placed in IOE's display case for IFAD staff and visitors. Printed copies are available upon request. The lead evaluator disseminates the evaluation report to government officials and other in-country partners. On a case-by-case basis, appropriate audio-visual instruments (e.g. using radio and television programmes) will be used to ensure outreach to the main beneficiaries.

Finally, a learning event is organized at the end of the evaluation by IOE at IFAD with Management and staff to discuss the main findings, lessons and recommendations. Similarly, a learning event is organized at the country level with the participation of government officials, representatives of beneficiaries (e.g. farmers' leaders or coordinators of self-help groups), and other key partners and stakeholders.

Evaluation synthesis reports. The communication plan should be included in the evaluation synthesis approach paper. Evaluation synthesis reports are prepared in English only. However, the acknowledgements, foreword, executive summary and IFAD Management's response will be translated into IFAD official languages, and should be included in the final report. The lead evaluator will prepare an Evaluation Profile and Insights in all cases.

An in-house workshop is organized to discuss the main findings from all evaluation synthesis reports to share lessons and promote dialogue around the main emerging themes.

Once finalized, the evaluation synthesis report is submitted to the ECU for editorial quality assurance, web publishing and dissemination. The ECU will post the report on the IFAD intranet (log-on screen), the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD's Facebook page and Twitter and Yammer accounts. The link to the report is also included in the UNEG database of evaluation reports, available at <http://www.uneval.org/evaluation/reports>.

The ECU shall share the final report electronically with internal and external audiences, including IFAD Management and staff, IOE staff, consultants, members of the Executive Board, UNEG, ECG, all IFAD regional networks, evaluation associations and other networks according to the thematic area discussed in the report. A limited number of hard copies are printed and placed in IOE's display case for IFAD staff and visitors. Printed copies are also available upon request.

Country strategy and programme evaluations. The communication plan for CSPEs should be included in the approach paper. The final CSPE report should include an ACP. CSPE reports are usually written in the IFAD official language of the concerned country. The acknowledgements, foreword, overview and ACP will be translated into English if they are written in another IFAD official language. The lead evaluator will prepare an Evaluation Profile and Insights in all cases.

As a final step in the evaluation process, national workshops for the evaluation stakeholders are held in partner countries to discuss main issues emerging from CSPEs. The workshops are co-organized by IOE and the concerned ministry in the country, in cooperation with PMD. The ECU develops and implements a communication strategy for these workshops, which includes a media advisory note and a press release. The news on the workshop is widely disseminated in-house and to external audiences. Participants invited to the workshops include IFAD and IOE Management and staff, development partners, government authorities, academics and researchers, representatives of beneficiaries (e.g. farmers' leader or coordinators of self-help groups), and other stakeholders.

Once finalized, the CSPE report, Profile and Insights are submitted to the ECU for editorial quality assurance, web publishing and dissemination. The ECU will post the documents on the IFAD intranet log-on screen, the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD's Facebook page and Twitter and Yammer accounts. The report is also uploaded in the UNEG database of evaluation reports, available at <http://www.uneval.org/evaluation/reports>. It is shared electronically with internal and external audiences, including the IFAD Management and staff, IOE staff, consultants, members of the Executive Board, UNEG, ECG, all IFAD regional networks, evaluation associations and other networks according to the respective geographical region. A limited number of hard copies are printed and placed in IOE's display case

for IFAD staff and visitors. Printed copies are also available upon request.

The lead evaluator disseminates the evaluation products to the members of the CLP and other government officials and in-country partners. If there is an IFAD country office, it supports IOE in disseminating hard copies of evaluation reports, including to subnational authorities and related partners, also in remote areas. As for impact evaluations, on a case-by-case basis, appropriate audio-visual instruments (e.g. using radio and television programmes) will be used to ensure outreach to the main beneficiaries.

Corporate-level evaluations. The communication plan for CLEs is included in the approach paper. The lead evaluator will be required to prepare an Evaluation Profile and Insights. Each CLE should include an overview, which covers the main findings, conclusions and recommendations of the evaluations. All CLE reports are prepared in English. The acknowledgements, foreword, overview and Management's response should be included in the final report. The report from the SIAs of the evaluation should be included as an annex.

An in-house learning workshop will also be organized by IOE to share the main findings, lessons and recommendations with IFAD Management and staff.

The report is submitted to the ECU of IOE for editorial quality assurance, web publishing and dissemination. The ECU will post the report on the IFAD intranet (log-on screen), the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD Facebook, Twitter and Yammer. The report is also uploaded on the UNEG database of evaluation reports, available at <http://www.uneval.org/evaluation/reports>. It is shared electronically with internal and external audiences, including IFAD Management and staff, members of the Executive Board, the UNEG, ECG, evaluation offices of bilateral agencies and regional development banks, all IFAD regional networks, evaluation associations and other networks according to the thematic area discussed by the CLE.

Annual Report on Results and Impact of IFAD Operations. An in-house event is organized by IOE to share the main results with IFAD Management and staff, and discuss the findings with regard to the specific learning theme covered in each ARRI. An Issues Paper on the learning theme is prepared and distributed before the event. The lead evaluator prepares a Profile for each ARRI. Each year, as per the Evaluation Policy, the ARRI is also shared and discussed with the Evaluation Committee and Executive Board.

Once finalized, the report and Profile are submitted to the ECU, which manages the graphic development, publishing and dissemination processes. The ECU will ensure the posting of the report on IFAD's intranet (log-on screen), the video wall in the IFAD lobby, the independent evaluation section of the IFAD website, IFAD Facebook page, and Twitter and Yammer accounts. The report is also uploaded in the UNEG database of evaluation reports, available at <http://www.uneval.org/evaluation/reports>. It is shared electronically with internal and external audiences, including IFAD Management and staff, members of the Executive Board, UNEG, ECG, OECD/DAC network, evaluation offices of bilateral development agencies and regional development banks, all IFAD regional networks and networks of evaluation associations and societies.

IOE newsletter

The ECU publishes a quarterly newsletter (at the end of March, June, September and December), which also includes updates related to the Evaluation Committee and the Executive Board.

The newsletter is shared electronically with internal and external audiences, including IFAD Management and staff, IOE staff, consultants, members of the Executive Board, UNEG, ECG, all IFAD regional networks, evaluation associations and other select evaluation networks.

Internet

IOE website

In cooperation with the Communications Division and within the framework of IFAD's corporate website, IOE maintains a dedicated website (www.ifad.org/evaluation) that is used as the main channel to publish evaluation reports and publications, announce events and post news, newsletters and editorials. It is also used to disseminate updated versions of documents such as the Evaluation Policy and the Evaluation Manual.

Social media

Social media has changed how people communicate and organizations now also use these channels as advocacy tools to inform the public about their work. IOE uses social media to disseminate reports that are published on the IOE website and announce events. It does so by using IFAD's social media accounts (e.g. Twitter, Facebook, IFAD social reporting blog), and is considering developing its independent social media presence in the future, as already done by other IFIs.

Media relations – press releases

The Evaluation Policy states that “on a case-by-case basis, IOE will issue a press release to inform the wider audience of the main results and lessons from key evaluations. IOE will be responsible for preparing the press release and will not require clearance from anyone outside the division for the contents contained therein. It will however seek the support of IFAD's Communications Division to draw on their contacts, expertise and capabilities for the issuance of the press releases.”

Press releases are normally used to announce the launching of selected evaluation reports (in particular CLEs and CSPEs) but may be used in other cases, for instance to communicate important events that may be of interest to the media.

IFAD's Communications Division supports IOE by uploading press releases on IFAD's corporate website, Facebook page and Twitter account. For CSPE in-country workshops, they are widely disseminated to local and regional media.

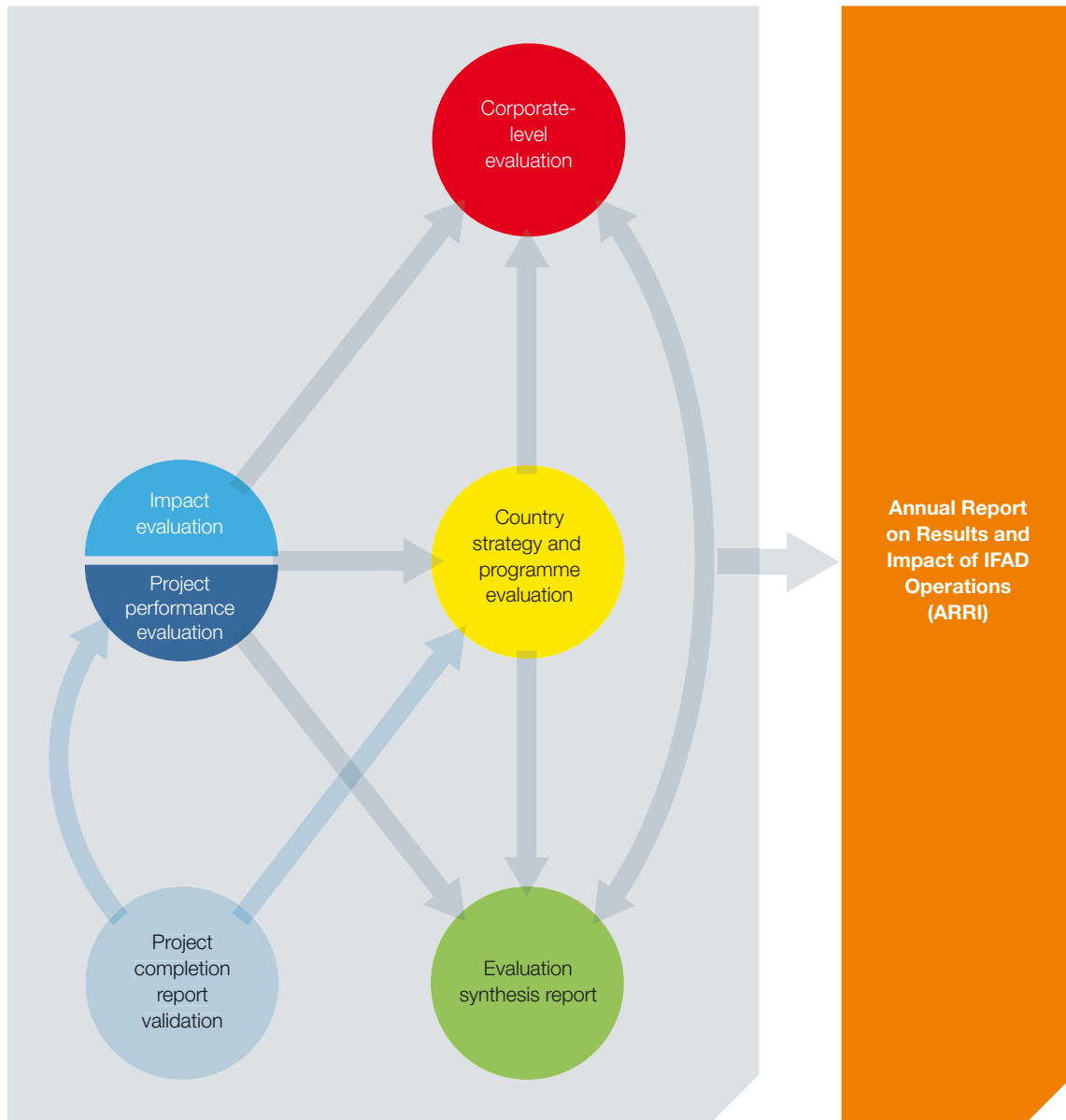
Visual tools

The ECU may produce videos to (i) illustrate beneficiary's perspectives in relation to the evaluation results and activities on the ground; (ii) show the highlights of an evaluation using footage from the field or interviews with selected evaluation officers and stakeholders; and (iii) tell the story of an evaluation using animation, maps and other visual material that can improve the understanding of the evaluation and increase transparency, stakeholders' buy-in and participation. Experts in the Communications Division assist IOE, as needed, in editing and finalizing such videos.

IOE produces visual communication tools that accompany evaluation reports. In particular, it will produce infographics for the ARRI, CLEs, CSPEs, and evaluation syntheses, and will also develop targeted visual sheets (e.g. info notes and fact sheets) on specific issues or special occasions.

Events

Emerging evaluation findings are presented at in-house learning events for ARRIs, CLEs, evaluation synthesis reports and impact evaluations; and at the country-level workshops for CSPEs. Findings from impact evaluations are presented in-house. The purpose is to stimulate discussion around the evaluation and its main findings and recommendations. These events and workshops also present an opportunity for IOE to display and share various publications.

ANNEX I Independent evaluation architecture for promoting accountability and learning at IFAD

ANNEX II Indicative set of additional questions

These questions supplement the core questions outlined in chapter 3. They are suggestions to help inform the evaluative inquiry, and are to be used selectively.

Rural poverty impact: impact domains

Household income and assets

Issues to explore:

- The change in the composition and level of household incomes (income sources, diversification, stability) including intra-household incomes and assets, and financial assets (savings, debt, borrowing, insurance)
- Changes in physical assets (farmland, water, livestock, trees, equipment, houses, bicycles, radios, television sets, telephones, etc.)

Questions:

- To what extent were the rural poor able to access financial markets more easily?
- In what way did the rural poor's access to input and output markets change?
- Can better health and education promoted by the programme explain a change in incomes and more assets?

Human and social capital and empowerment

Issues to explore:

- Changes in rural people's organizations, grass-roots institutions, social cohesion and local self-help capacities
- Changes in access to the information needed for the rural poor's livelihoods
- Questions:
 - To what extent do the rural poor play more effective roles in decision-making?
 - In what way did the project empower the rural poor vis-à-vis development actors and local and national public authorities?

Food security and agricultural productivity

Issues to explore:

- Changes in availability and stability of access to food
- Changes in cropping yield and intensity, land productivity and returns to labour

- Changes in children's nutritional status (e.g. stunting, wasting, underweight)

Questions:

- How did agricultural productivity impact on household food security?
- What was the role of improved access to input and output markets in enhancing the productivity of the rural poor?
- What are the links between productivity and access to food of the rural poor?

Institutions and policies

Issues to explore:

- Changes in rural financial institutions (e.g. in facilitating access for the rural poor)
- Changes in local governance, public institutions, NGOs and private sector that provide service delivery for the rural poor

Questions:

- What were the major ways in which the rural poor were affected by national/sectoral policies and the regulatory framework?
- What were the contributions of the project to changes in market structures and other institutional factors that can explain changes in poor producers' access to markets?

Relevance

Issues to explore:

- Consistency, coherence and complementarity with: national agriculture and rural development strategies, policies and programmes; other development partners' programmes and focus; the COSOP and relevant IFAD sector and subsector policies; and the needs of the rural poor.

Questions:

- Did the project benefit from available knowledge (for example, the experience of similar projects in the area or in the country) during its design and implementation?
- Were project objectives realistic?
- Did project objectives and design remain relevant over the period of time required for implementation?

- What are the main factors that contributed to a positive or less positive assessment of relevance?

Effectiveness

Issues to explore:

- Changes in the context (e.g. policy framework, political situation, institutional set-up, economic shocks, civil unrest, etc.)
- Degree to which objectives have been met and what may still be outstanding

Questions:

- Are adopted approaches technically viable?
- Do project users have access to adequate training for maintenance and to spare parts and repairs?

Efficiency

Issues to explore:

- Cost ratio of inputs to outputs
- Timeliness of loan effectiveness and implementation
- Administrative costs
- Economic rate of return

Questions:

- Was the project implemented in the most efficient way compared to alternatives?
- What are the major factors that account for project efficiency performance?
- What are the possibilities of benchmarking with peers?

Sustainability of benefits

Issues to explore:

- Expressions and indications of government commitment and capacity to sustain the project (provision of funds for selected activities, human resources availability, continuity of pro-poor policies and participatory development approaches, and institutional support)
- Engagement, participation and contributions from local communities, grass-roots organizations, and the rural poor

Questions:

- What is the likely resilience of economic activities to shocks or progressive exposure to competition and reduction of subsidies?

- What are the measures taken in terms of, for example, a specific exit strategy or approach prepared and agreed upon by key partners to ensure post-project sustainability?
- What factors militate in favour of or against maintaining benefits?

Gender equality and women's empowerment

Issues to explore:

- Volume and nature of project resources invested in gender equality and women's empowerment activities
- Changes to women's access to resources, assets and services; to women's influence in decision-making; in workload distribution among household members; in women's health, skills, income and nutritional levels; in gender relations within households, groups and communities in the project area; etc.

Questions:

- Did the initiative contain specific activities for gender equality and women's empowerment, and what was their effect on the rural poor? Did any activities give rise to unintended consequences on gender equality and women's empowerment?
- In terms of design, were corporate objectives on gender adequately addressed and integrated in the project activities/the results-framework of COSOPs? And were gender dimensions adequately addressed in the implementation structure and included in the project's annual work plans and budgets?
- During implementation, to what extent did the project: (i) monitor gender-disaggregated outputs to ensure gender equality and women's empowerment objectives were being met; (ii) adapt project implementation as required to better meet gender equality and women's empowerment objectives; (iii) address and report on gender issues in supervision and implementation support; (iv) systematically analyse, document and disseminate lessons on gender equality and women's empowerment; and (v) engage in policy dialogue to promote changes to government and other partner systems and processes that would improve gender equality and women's empowerment?

- What factors, including strengths and weaknesses of the contributions of IFAD and the government, were the most significant in promoting or hindering the corporate objectives on gender equality and women's empowerment?
- To what extent is the gender-related impact likely to be sustainable after completion of the IFAD-funded project?

Innovation and scaling up

Issues to explore:

- The innovative aspects of the project and examples/potential of scaling up
- Ways in which innovation and scaling up in the project has been documented and disseminated (e.g. workshops, exchange visits, MTRs, project supervision reporting, etc.)
- Strategic partnerships with organizations that could potentially be involved in scaling up of successfully piloted innovation

Questions:

- To what extent did the initiative being evaluated specifically address innovation and scaling up?
- To what extent did the project build on prior successful experiences and lessons with scaling up that may be well-established elsewhere, but new to the country or project area?
- Was an explicit strategy defined, including identifying the origin of innovation and pathways and drivers for scaling up? Was an ultimate scale target included?
- Did project implementation, including through the M&E system, support the development of relevant drivers (e.g. in terms of resource allocation for knowledge management) that are essential for scaling up?
- Through what processes has the project innovations been replicated and scaled up and, if so, by whom? If not, what were the obstacles and what are the realistic prospects that they can and will be replicated and scaled up by the government, other donors and/or the private sector?

Natural resources and the environment

Issues to explore:

- Status of the natural resources base (land, water, forest, pasture, fish stocks, etc.)
- Changes in protection, rehabilitation or depletion of natural resources and the environment

Questions:

- Did the initiative contain specific activities for rehabilitation or protection of natural resources and ecosystem services?
- In what way has the project impacted on environmental vulnerability (e.g. exposure to pollutants, climate change effects, volatility in resources, potential natural disasters)?
- In terms of design, was the rehabilitation or protection of natural resources and ecosystem services adequately addressed, in line with corporate objectives on environment and natural resources management, and included in the project's annual work plans and budgets?
- During implementation, to what extent did the project monitor changes in rehabilitation or protection of natural resources and ecosystem services and address and systematically analyse, document and disseminate lessons on rehabilitation or protection of natural resources and ecosystem services?

Adaptation to climate change

Issues to explore:

- Climate resilience, disaster preparedness measures, for example, in terms of agro-meteorological warning systems, drought contingency plans, response to flooding, weather-indexed risk insurance, etc.
- Volume and nature of funds allocated to measures aiming at adapting to climate-change related risks

Questions:

- Did the initiative contain specific adaptation activities, and what was their effect on the rural poor?

- In terms of design, were issues related to adaptation to climate change adequately addressed, in line with the corporate objectives of IFAD's climate change strategy and included in the project's annual work plans and budgets?
- To what extent were the climate-related considerations integrated in a coherent, consistent and logical manner across the project design?
- To what extent did the project include explicit measures to reduce the vulnerability of livelihoods to climate shocks and stresses?
- During implementation, to what extent did the project monitor changes in capacity to manage climate change, and systematically analyse, document and disseminate lessons on climate resilience?

Performance of partners

IFAD

Issues to explore:

- Volume and nature of resources mobilized (funding, time, technical expertise)
- Where applicable, role and performance of IFAD's country presence team (including proxy country presence arrangements) and support provided by headquarters to its country presence team (resources, follow-up and guidance, delegation of authority)

Questions:

- Has IFAD effectively and efficiently exercised its developmental, project management, and fiduciary responsibilities?
- At design stage, to what extent were (i) specific efforts made to incorporate the lessons and recommendations from previous independent evaluations in project design and implementation, and (ii) the design process participatory (with national and local agencies, grass-roots organizations) thus promoting ownership by the borrower
- During implementation, to what extent did IFAD: (i) take prompt action to ensure the timely implementation of recommendations stemming

from the supervision and implementation support missions, including the MTRs; and (ii) undertake the necessary follow-up to resolve any implementation bottlenecks?

- In what way has IFAD actively created an effective partnership and maintained coordination among key partners to ensure the achievement of project objectives, including the scaling up of pro-poor innovations?
- Has IFAD, together with the government, contributed to planning an exit strategy to ensure continued funding and sustainability of results?

Government

Issues to explore:

- Volume and nature of resources provided (staffing, counterpart funds, technical support, and project management)
- Project management, audit, observance of loan covenants, policy guidance to project management staff, and coordination of the various departments involved in project execution

Questions:

- Has the government assumed ownership and responsibility for the project? Judging by its actions and policies, has the government been fully supportive of project goals?
- During implementation did the government: (i) take the initiative to suitably modify the project design (if required); (ii) take prompt action to ensure the timely implementation of recommendations from supervision and implementation support missions, including the MTR?
- Did the M&E system generate information on performance and impact, which is useful for project managers, and has appropriate action been taken on the basis of this information?
- In what way has the government facilitated the participation of NGOs and civil society, where appropriate, and what were the implications?

ANNEX III **Template and process for agreement at completion point**¹⁰⁴

A. Introduction

Background

As per the IFAD Evaluation Policy and the Evaluation Manual, an agreement at completion point (ACP) is to be prepared for country strategy and programme evaluations undertaken by IOE.

The process and template outlines an efficient and transparent approach to preparing and finalizing the ACP within a specific time frame. This makes it possible to clearly capture any differing views on the part of IFAD Management and/or the government with regard to finding(s) and/or recommendation(s) deriving from the evaluation. It also allows IOE to convey its perspectives on any differences that may be articulated by the government and/or IFAD Management on any particular evaluation finding(s) and/or recommendation(s). The process also outlines the way in which such differences may be resolved.

Structure of the present document

Section B below outlines the main steps in the process of preparing the ACP, the timeframe, and the roles and responsibilities of all stakeholders involved. The role of the Evaluation Committee, if and when required, is also spelled out. The Evaluation Committee will review ACPs together with all CSPE reports they consider in a given year. They will be included in the process for completing an ACP only when differing views emerge among the key partners. The provision for disclosure of the final ACP is also outlined. Section C sets out the template, building upon the revised process outlined in section B.

B. Process for preparing the agreement at completion point

Drafting the ACP

As per the Evaluation Policy, IOE is only responsible for facilitating the process leading to preparation of the ACP and, to that end, it will initiate the ACP process by drafting the sections on Introduction and Main evaluation findings and send the

document to the Programme Management Department (PMD). The latter, working closely with the concerned government, will be responsible for drafting the ACP. This will be the joint response of IFAD Management and the concerned government, and will address the recommendations contained in the final evaluation report. The responsibility for the timely completion of the ACP rests ultimately with IFAD Management and the concerned government. The relevant sections of the draft ACP will first be sent by IOE to the concerned regional division of PMD. Once the ACP has been prepared, IFAD Management will send the document back to IOE, who will then transmit the draft ACP to the government by fax (copied to PMD) for its review and comments, or confirmation of the ACP.

In particular, the template for the ACP includes a section (section C), reserved for IOE comments in the final ACP, especially in the event of disagreement with any of the finding(s) and/or recommendation(s) contained in the evaluation.

Signing the ACP

The ACP will be signed by designated representatives of IFAD Management and the government concerned.

The concerned government will designate a representative of appropriate seniority/authority to sign the ACP on its behalf.

Time frames

The signed final ACP will be included in, and form an integral part of, the main evaluation report to be published by IOE. As such, it is important that the ACP be completed within a specified time frame to ensure timely issuance of the final evaluation report. In particular, ACPs should be signed within three months of the date of the evaluation learning workshop organized by IOE in collaboration with PMD (and, as appropriate, the concerned government).

¹⁰⁴ Agreed between the Independent Office of Evaluation of IFAD (IOE) and the Programme Management Department (PMD) – effective 1 January 2011.

Discussion of the ACP at the Evaluation Committee and Executive Board

ACPs will be discussed in the Evaluation Committee for all CSPEs considered by the Committee in a given year. Following the Board decisions related to new COSOPs, if a CSPE has been undertaken by IOE, its ACP must be included as an annex to the COSOP submitted for consideration by the Board.

For the evaluations that will not be considered by the Committee and in the event of delayed signature of an ACP or disagreement by IFAD Management and/or the government with regard to one or more evaluation finding or recommendation, fully or partly, IOE may request that the ACP in question be included in the provisional agenda of the Evaluation Committee. Taking into account IOE's comments, the aim of the Evaluation Committee discussion is to seek its guidance on the evaluation finding(s) and/or recommendation(s) that IFAD Management and/or the government disagrees with. The Committee will also, by means of its chairperson's report, explicitly recommend that the Executive Board request IFAD Management and/or the government to take action on the recommendation(s) contested, as deemed appropriate.

IOE will inform the Evaluation Committee of delays in the provision of feedback on the draft ACP from either IFAD Management and/or the government, with the aim of informing them and seeking their guidance on the way forward.

Disclosure

As mentioned above, the signed ACP will be included as part of the final published evaluation report, to be disclosed in both printed and electronic form. However, in order to ensure timely disclosure of the main evaluation results, an advance electronic copy of the final evaluation report (excluding the ACP) will be made available through the IOE web pages on IFAD's corporate website before the final evaluation learning workshop is held.

Once an evaluation is fully completed, IOE will inform members of the Executive Board through means of a letter/email that the final evaluation

report inclusive of the ACP is now available on the IOE web pages.

C. Template for the agreement at completion point

Introduction

The introductory section of the ACP will provide an overview of the objectives of the evaluation and of key steps in the process leading to the conclusion of the agreement, including the date of the learning workshop held at the end of the evaluation process.

A short statement will explain what the ACP constitutes and who will sign the document for the government and IFAD, and describe IOE's role in facilitating the process leading to conclusion of the agreement. It will also explain that the recommendations agreed upon will be tracked through the President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA). In all, this section should not exceed half a page.

Main evaluation findings

This section will summarize the key evaluation findings, which provide the basis for the evaluation's recommendations. The indicative length of this section will be around one page.

Agreement at completion point

This section will be drafted by IFAD Management, in consultation with the concerned government, as appropriate. Each recommendation specified in the final evaluation report will be discussed and concrete measures will be provided on how to implement them. They will also propose a deadline for the implementation of each recommendation and indicate the entity (within government, IFAD or both) responsible for acting on them. IFAD Management and the concerned government will specify how each evaluation recommendation will be implemented (e.g. preparation of a new corporate policy or procedure, a COSOP or project design) and possible resource or other implications. For example, if an evaluation has

generated two main recommendations, the following format will apply:

- **Recommendation 1:** [text to be taken from the final evaluation report]
 - Deadline date for implementation: [proposed by PMD]
 - Entities responsible for implementation: [proposed by PMD]
 - This recommendation will be implemented during preparation of the next COSOP for [name of the country concerned].

- **Recommendation 2:** [text taken from the final evaluation report]
 - Deadline date for implementation: [proposed by PMD]
 - Entities responsible for implementation: [proposed by PMD]
 - This recommendation will be implemented in the next project to be designed in [name of the country concerned].

IFAD Management and/or the government will clearly specify if they do not agree with a particular recommendation and underline the reasons for the same, and specify how they intend to proceed alternatively. A difference of opinion between IFAD Management and the government on any of the recommendations will also be captured here.

Comments by IOE

This section is optional. If either IFAD Management and/or the government have expressed disagreement on any of the findings and/or on one or more of the recommendations deriving from the evaluation, either fully or partially, IOE will add a section to the ACP providing its final views on the disagreement. The final ACP will be shared with the Evaluation Committee for its consideration.

ANNEX IV Conflict of Interest Policy for IOE Staff and Consultants

Conflict of Interest Policy for evaluation officers in the Independent Office of Evaluation of IFAD

Background

Management and staff of the Independent Office of Evaluation of IFAD (IOE) are committed to producing excellent independent evaluations with impartiality and integrity. In this process, they are committed to avoiding conflicts of interests in their work.

The Peer Review of IFAD's Office of Evaluation and Evaluation Function, undertaken in 2010, recognized that IOE has comprehensive guidelines for avoiding conflict of interest of consultants it employs for evaluation work. It recommended that the division also develop similar conflict of interest guidelines for IOE staff members.

The aim of this note therefore is to outline those guidelines; It relates to the conflict of interest of Professional staff members who ultimately are responsible for forming evaluative judgements and preparing evaluation reports related to IFAD-supported policies, strategies, business processes and operations. These guidelines build on similar existing guidelines used by evaluation outfits in other multilateral development organizations.

The guidelines

IOE staff will recuse themselves from evaluating any IFAD-funded policy, strategy, or operation they may have worked on at an earlier stage, such as the design, implementation or supervision of an IFAD-financed project. However, such staff may be part of internal peer review processes within IOE, which are undertaken to ensure high quality evaluation deliverables.

Moreover, an IOE staff member who had previously worked in a regional division within IFAD's Programme Management Department (PMD) will generally not be requested to perform evaluations in the same region of that division for a specific period of time. This will be defined on a case-by-case basis when the staff member joins IOE.

IOE staff is invited to participate in in-house design processes, with the aim of conveying lessons and recommendations emerging from previous evaluations. In the situation where an IOE staff member has been designated to evaluate a policy, strategy or project for which s/he may have provided such type of inputs at the design stage, this shall not constitute a conflict of interest.

IOE staff will not be allowed to take up assignments (as staff or consultant) in an IFAD regional division of PMD in which s/he may have had major responsibility for the overall management and contents of an evaluation. The Director IOE will examine the cases of individual IOE staff members requesting to take up assignments in PMD, and on a case-by-case basis, take decisions accordingly.

With regard to the aforementioned, if a staff member plans to seek employment in another IFAD regional division, s/he must disclose this information to the Director of IOE. Thus, the Director will have complete information before assigning (or reassigning) responsibilities for evaluation work. The Director must have complete information before assigning (or reassigning) responsibilities for evaluation work, which also includes the supervision of an evaluation of an IFAD-funded project or if s/he (or an immediate family member) had been previously employed in a decision-making position at a non-IFAD entity (e.g. an NGO) that was included in an IFAD-funded project.

In the case where the potential conflict of interest or perception of conflict of interest is identified after an evaluation has started, the Director of IOE will decide if the assigned evaluator should thereafter recuse him/herself from the evaluation and, if so, whether the evaluation should be continued using the work undertaken up to that point, or whether it should recommence from the beginning.

Causes of possible conflict of interest that may emerge from working with governments and/or partners in borrowing countries will be evaluated on a case-by-case basis, and a decision taken accordingly by the Director of IOE.

The Director of IOE will manage the guidelines and keep track of the issues and their resolution. The documentation regarding these issues will be treated confidentially.

In summary, the Director of IOE and staff are required to exercise sound professional ethics and personal good judgement in adhering to these guidelines. The Director of IOE and staff are therefore responsible for conforming with the intent and spirit of the guidelines in all matters not specifically stated above. Should evaluators have any doubts with regard to their proper course of action in any matter related to a conflict of interest issue, they must seek the advice of the Director of IOE.

These guidelines extend to all IOE Professional staff and are effective immediately.

Date: 15 November 2010

Conflict of Interest Policy for consultants in the Independent Office of Evaluation of IFAD

Introduction and definition

The Peer Review of IFAD's Office of Evaluation and Evaluation Function recommended that a review be undertaken of the conflict of interest policy for consultants working in IOE.

For the purposes of this policy, a conflict of interest situation in consultant recruitment is defined as a situation in which interests other than those associated with the duty of conducting an objective evaluation may materially interfere, or be perceived to interfere, with the fulfilment of this duty.

The most relevant cases would be those in which consultants could: (a) influence the analysis or recommendations so that they are consistent with findings previously stated by themselves (upstream or ex ante conflicts of interest); (b) artificially create favourable conditions for consideration in a downstream assignment (downstream or ex post conflict of interest); or (c) work simultaneously for two or more clients whose interests diverge.¹⁰⁵

Evaluation outfits in other international financial institutions have formulated general principles to avoid conflict of interest with staff and consultants. The United Nations Evaluation Group sets out similar

¹⁰⁵ Annex 6, Evaluation Manual, 2009, Independent Office of Evaluation of IFAD.

principles in its Standards for Evaluation in the United Nations System. This document presents a conflict of interest policy for consultants in IOE and is aligned with the principles discussed above.

Purpose, basic principles and operational rules

The purpose of this policy is to provide simple, practical and enforceable rules to identify, and consequently avoid, conflicts of interest with consultants engaged by IOE.

The basic principles are the following:

- (i) Conflicts of interest, either in substance or in appearance, are to be avoided in all IOE evaluations; IOE evaluations must be – and must also be perceived to be – independent, impartial and devoid of any conflict of interest;
- (ii) IOE consultants, together with IOE staff members, share the responsibility to avoid situations and activities that might reflect negatively on IOE, compromise their operations, or lead to real or apparent conflicts of interest; and
- (iii) All evaluators, staff and consultants are responsible for exercising sound professional ethics and good judgement in applying the conflict of interest policy. They are also responsible for sharing any information about conflict of interest they may face and, if necessary, should recuse themselves from the proposed assignment. As a consequence:
 - IOE consultants must inform IOE staff/lead evaluators of any conflict of interest, real or perceived, and sign a declaration to this effect with IOE prior to commencing their assignment; and
 - IOE staff/lead evaluators must inform the Director of IOE of any conflict of interest, real or perceived in relation to the recruitment of the consultant.
- (iv) If in doubt as to whether a conflict of interest exists, IOE staff/lead evaluators should seek the views of the Director of IOE.

From an operational point of view, these principles translate into the following rules:

- (i) IOE will not assign evaluation of projects, country programmes, corporate processes, policies or strategies to consultants who have had prior involvement in their design or implementation, including supervision, related decision-making or financing;
- (ii) Consultants should not accept concomitant assignments with the regional division responsible for the project/country programme being assessed by the evaluation; and
- (iii) Consultants should agree not to take other IFAD assignments for the country programme being assessed by the evaluation for a period of six months.

Procedure for contract approval

All consultants will be requested to sign and submit a declaration stating that they are aware of IOE's conflict of interest policy and that they have no conflict of interest with regard to the proposed assignment. This is a precondition of contract preparation.

Exemptions

This policy does not apply to editors, translators, facilitators or resource persons who are not responsible for analytical evaluation work. Any other exemptions from this policy will be based on the recommendation of the hiring officer and must be approved by the Director of IOE.

Date: 10 July 2013

Sample declaration of absence of conflict of interest

I declare that I have read the Conflict of Interest Policy for Consultants in the Independent Office of Evaluation of IFAD and state that:

- 1) I have had no prior involvement in:
 - any activity (e.g. design, supervision, implementation support) related to the project to be assessed by the evaluation (for project completion report validations, project performance or impact evaluations);
 - any activity (e.g. design, supervision or implementation) related to the country programme to be assessed by the evaluation (for country strategy and programme evaluations);
 - any activity related to the formulation of the policies or strategies to be assessed by the evaluation (for corporate-level evaluations and evaluation syntheses).
- 2) I have no conflict of interest, real or perceived, in relation to the proposed assignment;
- 3) Under normal circumstances, while under assignment with IOE, I will not accept concomitant contracts with the division responsible for the project/ country programme being assessed by the evaluation;
- 4) I will not seek employment within IFAD that is connected in any way with the country programme assessed by the evaluation for a period of six months after the end date of my contract with IOE.

Printed name

Signature

Date

ANNEX V IOE selectivity framework

Criteria for the selection and prioritization of evaluations for inclusion in IOE's work programme

Corporate-level evaluations (CLEs)	Country strategy and programme evaluations (CSPEs)	Evaluation synthesis reports (ESRs)	Project performance evaluations (PPEs)	Impact evaluations
<p>1. Strategic priority. The evaluation contributes to IFAD's strategic priorities and replenishment commitments</p> <p>2. Accountability. Topic selected contributes to strengthening IFAD's institutional accountability</p> <p>3. Knowledge gap. CLEs contribute to filling a critical knowledge gap in IFAD</p> <p>4. Timeliness. Evaluation results feed punctually into pertinent corporate policies, strategies and/or processes</p> <p>5. Corporate risks. The evaluation serves to help minimize critical corporate risks</p>	<p>1. Link to COSOPs. Results feed into the development of IFAD country strategies/ COSOPs</p> <p>2. Coverage:</p> <p>a) Regional and country coverage of CSPEs</p> <p>b) Size of the portfolio in terms of total investments and number of operations</p> <p>c) Debt Sustainability Framework classification (red, yellow, green)</p> <p>d) Lending terms (highly concessional, blend or ordinary)</p>	<p>1. Evaluative evidence. Availability of adequate evaluative evidence by IOE and evaluation functions in other development organizations</p> <p>2. Knowledge gap. ESRs contribute to filling a critical knowledge gap in IFAD</p> <p>3. Strategic priority. The synthesis contributes to IFAD's strategic priorities and replenishment commitments</p> <p>4. Timeliness. The synthesis feeds punctually into pertinent corporate policies, strategies and/or processes</p> <p>5. Building block. The synthesis serves as an input for other IOE products</p>	<p>1. Availability of project completion report (PCR). PPEs will be done only when a PCR is available</p> <p>2. Geographic coverage. PPEs selected to ensure regional balance of the IOE evaluation programme</p> <p>3. Building block. Priority given to PPEs that will provide an input into CSPEs, CLEs or synthesis reports</p> <p>4. Information gaps. PCR does not provide sufficient analysis of project performance and results</p> <p>5. Inconsistencies. PCR ratings are inconsistent with narrative</p> <p>6. Innovative approaches. The project includes innovative approaches that merit deeper analysis and documentation</p> <p>7. Learning from PPE. Evidence needed on what worked and why</p>	<p>1. No duplication. No impact evaluation conducted by IFAD management of the same operation</p> <p>2. Learning from impact evaluations. Evidence needed on what works in a certain context</p> <p>3. Building block. Priority for impact evaluations that will provide an input into CSPEs, CLEs or synthesis reports</p> <p>4. Completion date. Impact evaluations will be done within three years after completion date</p> <p>5. Baseline data. The availability and usability of baselines is essential to determine the methodology to be applied in impact evaluations</p> <p>6. Information gaps. The PCR does not provide sufficient analysis of the effectiveness and impact of certain interventions</p> <p>7. Innovative approaches. The project includes innovative approaches that merit deeper analysis and documentation</p>

ANNEX VI Bibliography and references

- Asian Development Bank. 2008. *Strengthening the Quality of Project Economic Analysis in ADB Operations*. Manila: ADB.
- Bamberger, M. 2012. *Introduction to Mixed Methods in Impact Evaluation*, InterAction, USA.
- Bamberger, M., J. Rugh, M. Church and L. Fort. 2004. Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints. *American Journal of Evaluation*, 3 2004; Vol. 25, USA.
- Bamberger, M., Rugh, J. and L. Mabry. 2006. *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. Thousand Oaks, California: SAGE Publications.
- Befani, B., C. Barnett and E. Stern. 2014. Rethinking Impact Evaluation for Development, *IDS Bulletin*. Volume 45, Number 6.
- Brinkmann, S. and S. Kvale. 2014. *Interviews – Learning the Craft of Qualitative Research Interviewing*. Thousand Oaks, California: SAGE Publications.
- Carletto, C. 1999. *Constructing Samples for Characterizing Household Food Security and for Monitoring and Evaluating Food Security Interventions: Theoretical Concerns and Practical Guidelines*. IFPRI Technical Guide 8. Washington, D.C.
- Chen, H.S. 1990. *Theory Driven Evaluation*. Thousand Oaks, California: SAGE Publications.
- Chris, C. et. al. 2011. "A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009". USA: American Journal of Evaluation.
- Daniel, J.N. 2011. *Sampling Essentials: Practical Guidelines for Making Sampling Choices*. Thousand Oaks, California: SAGE Publications.
- Danish International Development Agency. 2012. *Case studies in Development Evaluation; Validity, Generalization and Learning*, Copenhagen, Denmark (Workshop) <http://um.dk/en/danida-en/results/eval/>. Copenhagen: DANIDA.
- Department for International Development. 2013. *Planning Evaluability Assessments*, London: DFID.
- Donaldson, S.I. (editor). 2013. *Emerging Practices in International Development Evaluation*, Evaluation and Society. USA: Information Age Publishing.
- Earl, S., Carden, F. and T. Smutylo. 2001. *Outcome mapping: Building learning and reflection into development programs*. Ottawa: International Development Research Centre.
- Evaluation Cooperation Group. 2008. *Good Practice Standards for Country Strategy and Program Evaluation*. Manila: ECG.
- _____. 2010. *Peer review of IFAD's Office of Evaluation and Evaluation Function*. Rome: ECG.
- _____. 2011. *Evaluative Lessons for Agriculture and Agribusiness*. Manila: ECG.
- _____. 2012. *Big book of Good Practice Standards*. Manila: ECG.
- _____. 2012. *Evaluating Technical Assistance: Taking Stock of the Practices of International Financial Institutions*. Manila: ECG.
- _____. 2012. *Good Practice Standards for the Evaluation of Public Sector Operations*, Manila: ECG.
- _____. 2014. *Evaluating Support to SMEs: Rationale, Challenges and Opportunities*. Manila: ECG.
- _____. 2014. *Knowledge Management in International Financial Institutions*. Manila: ECG.
- Funnell, S. and P. Rogers. 2011. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco, California: Jossey Bass, USA.
- Gertler, P.J., S. Martinez, P. Premand, L.B. Rawlings, and C.M. Vermeersch. 2011. *Impact Evaluation in Practice*. Washington, D.C.: World Bank.
- Government Accountability Office. (GAO) 1990. *Case Study Evaluations*, Washington, D.C.: GAO.
- Hancock, D. and B. Algozzine. 2011. *Doing Case Study Research: A Practical Guide for Beginning Researchers*, Second Edition (no electronic document). New York, N.Y.: Teachers College Press, Columbia University.
- Henry, G. 1990. *Practical Sampling*. Thousand Oaks, California: SAGE Publications.
- International Fund for Agricultural Development. *Annual Report on Results and Impact of IFAD Operations (ARRI)*. Rome: IFAD.
- _____. Report on IFAD's Development Effectiveness (RIDE). Rome: IFAD.
- _____. 2002. *A Guide for Project M&E*. Rome: IFAD.
- _____. 2005. *Sector-wide Approaches for Agriculture and Rural Development Policy*. Rome: IFAD.
- _____. 2006. *Supervision and Implementation Support Policy*. Rome: IFAD.
- _____. 2006. *Targeting Policy: Reaching the Rural Poor*. Rome: IFAD.
- _____. 2007. *Innovation Strategy*. Rome: IFAD.
- _____. 2007. *Knowledge Management Strategy*. Rome: IFAD.
- _____. 2008. *IFAD Policy on Improving Access to Land and Tenure Security*. Rome: IFAD.
- _____. 2009. *IFAD Policy on Engagement with Indigenous Peoples*. Rome: IFAD.
- _____. 2009. *Guidance note: Poverty Targeting in IFAD-supported Projects*. Rome: IFAD.
- _____. 2009. *Revised IFAD Policy for Grant Financing*. Rome: IFAD.
- _____. 2009. *Sustainability of Rural Development Projects*. Occasional paper 8: <http://www.ifad.org/operations/projects/regions/pi/paper/8.pdf>.
- _____. 2010. *IFAD Climate Change Strategy*. Rome: IFAD.
- _____. 2011. *IFAD Country Presence Policy and Strategy*. Rome: IFAD.
- _____. 2011. *IFAD Evaluation Policy*. Rome: IFAD.
- _____. 2011. *IFAD Policy on Environment and Natural Resources Management*. Rome: IFAD.
- _____. 2011. *IFAD Policy on Engagement with Middle-income Countries*. Rome: IFAD.
- _____. 2011. Strategic Framework 2011-2015.
- _____. 2011. *Updated Guidelines and Source Book for Preparation and Implementation of a Results-Based Country Strategic Opportunities Programme (RB-COSOP)*. Rome: IFAD.
- _____. 2012. *IFAD Policy on Gender Equality and Women's Empowerment*. Rome: IFAD.
- _____. 2012. *Internal Peer Review Process and Procedures*. Rome: IFAD.
- _____. 2012. *Partnership Strategy*. Rome: IFAD.
- _____. 2012. *Private-Sector Strategy – Deepening IFAD's Engagement with the Private Sector*. Rome: IFAD.
- _____. 2012. *Report of the Consultation on the Ninth Replenishment of IFAD's Resources*, Rome: IFAD.
- _____. 2013. *IOE results-based work programme and budget*, Rome: IFAD.

- _____. 2014. *IFAD Replenishments, Corporate-level Evaluation*, Rome. IFAD.
- _____. 2014. *The IFAD 9 Impact Evaluation Initiative: Sourcebook (draft)*. Alessandra Garbero and Statistics and Studies for Development Division. Rome: IFAD.
- Iverson A. 2003. *Attribution and Aid Evaluation in International Development: A literature review*. Evaluation Unit, International Development Research Center, Ontario, Canada.
- Johnson, N., J. Mayne, D. Grace and A. Wyatt. 2015. *How Will Training Traders Contribute to Improved Food Safety in Informal Markets for Meat and Milk?: A Theory of Change Analysis*, IFPRI Discussion Paper 1451. Washington, D.C.: International Food Policy Research Institute (IFPRI). Available at <http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/129293>.
- Khandker, S.R., G.B. Koolwal and H.A. Samad. 2010. *Handbook on Impact Evaluation. Quantitative Methods and Practices*. Washington, D.C.: World Bank.
- Krueger, R.A. and M.A. Casey. 2000. *Focus Groups*. Thousand Oaks, California: SAGE Publications.
- Laubli Loud, M. L., and J. Mayne (eds.). 2014. *Enhancing Evaluation Use*, Thousand Oaks, California: SAGE Publications.
- Leeuw, F. L. 2012. *Theory-Based Evaluation*, ALNAP.
- Leeuw, F. L. and J. Vaessen. 2009. *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*. Washington, D.C.: The World Bank.
- Levin, H.M. and P.J. McEwan. 2001. *Cost Effectiveness Analysis*. Thousand Oaks, California: SAGE Publications.
- Levy, P. S. and S. Lemeshow. 2013. *Sampling of Populations: Methods and Applications*.
- List, J.A., S. Sadoff, and M. Wagner. 2009. *So you want to run an experiment, Now what? Some simple rules of thumb for optimal experimental design*, Carlo Alberto Notebooks No 125, Turin, Italy.
- Mayne, J. 1999. *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*, Mimeo. Office of the Auditor General of Canada, Canada.
- _____. 2015. *Useful Theory of Change Models*. Canadian Journal of Program Evaluation, 30 (2), 119-142.
- Mayne, J. and N. Johnson. 2015. *Using Theories of Change in the CGIAR Research Program on Agriculture for Nutrition and Health*. SAGE Journals. Evaluation, 21(3), 407-428.
- McNall, M. and P.G. Foster-Fishman. 2007. Methods of Rapid Evaluation, Assessment, and Appraisal. *American Journal of Evaluation*, 6 2007; vol. 28, USA.
- Office of Internal Oversight Services. 2014. *Inspection and Evaluation Manual*. Inspection and Evaluation Division. New York: OIOS.
- Organisation for Economic Co-operation and Development/ Development Assistance Committee. *Criteria for Evaluating Development Assistance. Fact sheet*. Paris: OECD/DAC.
- _____. *Evaluating Development Cooperation Summary of Key Norms and Standards Second edition*. DAC Network on Development Evaluation. Paris: OECD/DAC.
- _____. 1991. *DAC Principles for the Evaluation of Development Assistance*. Paris: OECD/DAC.
- _____. 1998. *Review of the DAC Principles for Evaluation of Development Assistance*. Paris: OECD/DAC.
- _____. 2002. *Glossary of Key Terms in Evaluation and Results-Based Management*. Paris: OECD/DAC.
- _____. 2006. *Guidance for Managing Joint Evaluations*. Paris: OECD/DAC.
- Patrizi, P. and M.Q. Patton. 2010. *New Directions for Evaluation. Evaluating Strategy*, Wiley Periodicals, USA.
- Patton, M. Q. 2001. *Qualitative Research and Evaluation Methods*. Thousand Oaks, California: SAGE Publications.
- _____. 2008. *Utilization-Focused Evaluation*. Thousand Oaks, California: SAGE Publications.
- _____. 2010. *Guidelines and Reference Series: Quality Standards for Development Evaluation*. Paris: OECD/DAC.
- _____. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Thousand Oaks, California: SAGE Publications.
- _____. 2011. *Qualitative Research and Evaluation Methods: Integrating Theory and Practice*. Thousand Oaks, California: SAGE Publications.
- Pawson, R. and N. Tilley. 2004. *Realistic Evaluation*. Thousand Oaks, California: SAGE Publications.
- Ravallion M. 1999. *The Mystery of the Vanishing Benefits: Ms Speedy Analyst's Introduction to Evaluation*, Working Paper No. 2153. Washington, D.C.: World Bank.
- Rihoux, B. and C. Ragin. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and related techniques*. Applied Social Research Methods Series, Volume 51. Thousand Oaks, California: SAGE Publications.
- Rogers, P. and G. Peersman. 2015. *Impact evaluation: challenges to address*. [Online] 23 January 2015. Available from http://betterevaluation.org/blog/impact_evaluation_1.
- Savedoff, W., R. Levine and N. Birdsall. 2006. "When Will We Ever Learn? Improving Lives through Impact Evaluation." The Evaluation Gap Working Group. Washington, D.C.: Center for Global Development.
- Schneider, C. and C. Wagemann. 2012. *Set-theoretic methods for the social sciences. A Guide to QCA*. Cambridge University Press. <http://www.compass.org>.
- Scriven, M. 2011. *Evaluating evaluations: a meta-evaluation checklist*. Available at: http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.1.11.pdf.
- Shaffer, P. 2015. *Two Concepts of Causation: Implications for Poverty*. In *Development and Change*, 46(1), pp.148-166. Available at: <http://doi.wiley.com/10.1111/dech.12140>
- Stake, R.E. 1995. *The Art of Case Study Research*. Thousand Oaks, California: SAGE Publications.
- Stern, E. 2015. *Impact Evaluation: A Guide for Commissioners and Managers*. London: Big Lottery Fund, Bond, Comic Relief and the Department for International Development. Available at http://www.bond.org.uk/data/files/Impact_Evaluation_Guide_0515.pdf.
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani. 2012. *Broadening the Range of Designs and Methods for Impact Evaluations*, Final report. London: DFID.
- Swedish International Development Cooperation Agency. 2010. SIDA studies in evaluation. *Measuring Empowerment? Ask them. Quantifying qualitative outcomes from people's own analysis*. Stockholm: SIDA.
- Tashakkori A. and C. Teddlie. (1998), *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, California: SAGE Publications.
- Todd, D. and J. Brann. 2007. *GEF Evaluation Office Approach Paper on Impact Evaluation*, (1), pp. 1-44. Washington, D.C.: Global Environment Facility.
- Tourangeau, R., B. Edwards, T. Johnson, et ux. (Eds.). 2014. *Hard-to-Survey Populations*. Cambridge University Press.

- United Nations Development Programme. 2011. *Handbook on Planning, Monitoring and Evaluating for Results*. New York: UNDP.
- . 2013. *Innovations in Monitoring and Evaluation*. Discussion Paper. New York: UNDP.
- . 2015. *Human Development Report*. New York: UNDP.
- United Nations Evaluation Group. 2005. *Standards for Evaluation in the UN System*, New York: UNEG.
- . 2005. *Norms for evaluation in the UN System*, New York: UNEG.
- . 2010. *Good Practice Guidelines for Follow-up to Evaluations*. New York: UNEG.
- . 2010. *Quality Checklist for Evaluation Reports*. New York: UNEG.
- . 2013. *Impact Evaluation Guidance Document*. New York: UNEG.
- . 2014. *Handbook for Conducting Evaluations of Normative Work in the UN System*. New York: UNEG.
- . 2014. *Integrating Human Rights and Gender Equality in Evaluations*. New York: UNEG.
- United Nations Children's Fund. 1991. *Guide for Monitoring and Evaluation: Making a Difference?* UNICEF.
- United Nations Environment Programme. 2008. *Evaluation Manual*. Kenya: UNEP.
- United States Agency for International Development (USAID). 2013. *Technical Note – Conducting Mixed Methods Evaluations*. Washington, D.C.: USAID.
- Urdan, T. 2010. *Statistics in Plain English*, Third Edition. Taylor & Francis.
- Whitney, D. and A. Trosten-Bloom. 2010. *The Power of Appreciative Inquiry (2nd Ed.)*. San Francisco, California: Berrett-Koehler.
- Wholey, J. S., H.P. Hatry and K.E. Newcomer. (2004), *Handbook of Practical Program Evaluation*, San Francisco, California: John Wiley and Sons.
- World Bank. 2004. *Monitoring and Evaluation: Some Tools, Methods and Approaches*. Washington, D.C.: World Bank.
- . 2006. *Conducting Quality Impact Evaluations under Budget, Time and Data Constraints*. Washington, D.C.: World Bank.
- . 2010. *Handbook on Impact Evaluation*, Washington, D.C.: World Bank.
- . 2011. *Combining Quantitative and Qualitative Methods for Program Monitoring and Evaluation: Why Are Mixed-Method Designs Best?* Prem-notes No 9. Washington, D.C.: World Bank.
- . 2011. *Writing Terms of Reference for an Evaluation: a How-to guide*. Washington, D.C.: World Bank.
- . 2012. *Building Better Policies: the Nuts and Bolts of Monitoring and Evaluation Systems*. Washington, D.C.: World Bank.
- . 2012. *Designing a Results Framework for Achieving Results: a How-to guide*. Washington, D.C.: World Bank.
- . 2013. *Coping with the Attribution Problem in Program Evaluation*. Prem-notes, Number 26. Washington, D.C.: World Bank.
- World Health Organization. Global Database on Child Growth and Malnutrition: www.who.int/nutgrowthdb/database/en/
- Yin, R. K. 2003. *Case Study Research: Design and Methods (Fifth Edition)*. Thousand Oaks, California: SAGE Publications.
- . 2011. *Applications of Case Study Research (Third Edition)*. Thousand Oaks, California: SAGE Publications.

Websites

Evaluation offices in multilateral and United Nations agencies

- African Development Bank:
<http://operationsevaluation.afdb.org/en>
- Asian Development Bank: <http://adb.org/evaluation>
- IFAD: www.ifad.org/evaluation/index.htm
- Inter-American Development Bank: www.iadb.org/ove
- Organisation for Economic Cooperation and Development (OECD/DAC):
www.oecd.org/development/evaluation
- World Bank: <http://ieg.worldbank.org>

Evaluation groups

- Evaluation Cooperation Group: www.ecgnet.org
- United Nations Evaluation Group: www.uneval.org

Evaluation associations

- Active Learning Network for Accountability and Performance: www.alnap.org
- African Evaluation Association: www.afrea.org
- American Evaluation Association: www.eval.org
- Better Evaluation: www.betterevaluation.org
- Canadian Evaluation Society: www.evaluationcanada.ca
- European Commission Better Regulation: http://ec.europa.eu/smart-regulation/evaluation/index_en.htm
- European Evaluation Society: www.europeanevaluation.org
- Innovations for Poverty Action: www.poverty-action.org
- International Initiative for Impact Evaluation: www.3ieimpact.org
- Poverty Action Lab: www.povertyactionlab.org
- Regional Centers for Learning on Evaluation and Results: www.theclearinitiative.org

Other links

- DAC Evaluation Resource Centre (DEReC): www.oecd.org/derec
- Engendering Policy through Evaluation: www.feministevaluation.org
- Gender and Evaluation: <http://gendereval.ning.com>
- International Initiative for Impact Evaluation (3ie): www.3ieimpact.org/en/evaluation/resources/impact-evaluation-resources/
- Mixed Methods International Research Association (MMIRA): www.mmira.wildapricot.org
- My M&E: www.mymande.org
- NONIE, Impact evaluations and development: http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
- Poverty Action Lab: www.povertyactionlab.org/research-resources
- Random number generators: <http://stattrek.com/statistics/random-number-generator.aspx>
- Sample size calculators: www.surveysystem.com/sscalc.htm
- UNEG guidance documents: www.uneval.org/document/guidance-documents



International Fund for Agricultural Development

Via Paolo di Dono, 44 - 00142 Rome, Italy


Tel: +39 06 54591 - Fax: +39 06 5043463

E-mail: evaluation@ifad.org

www.ifad.org/evaluation

www.ruralpovertyportal.org

 ifad-un.blogspot.com

 www.facebook.com/ifad

 instagram.com/ifadnews

 www.twitter.com/ifadnews

 www.youtube.com/user/ifadTV

ISBN-978-92-9072-624-1



9 789290 726241

February 2016