



Investing in rural people

Improved Learning Initiative for the design of a
Participatory Impact Assessment & Learning Approach (PIALA)

Insights and lessons learned from the reflections on the PIALA piloting in Vietnam

**Adinda Van Hemelrijck
(PhD interim report)**

With inputs from the Core Design Team, in particular Irene Guijt and Jeremy Holland
With support from Edward Heinemann and Richard Caldwell

November 2014

Project: Improved Learning Initiative (ILI)
Funding: IFAD – Innovations Mainstreaming Initiative (IMI)
BMGF – Measurement, Learning & Evaluation
Managers: **Edward Heinemann** – Senior Policy Advisor, IFAD (e.heinemann@ifad.org)
Richard Caldwell – Head of Measurement, Learning and Evaluation, BMGF
(Richard.Caldwell@gatesfoundation.org)
Adinda Van Hemelrijck – Consultant Impact Assessment & Learning in Collaborative Settings, IDS (a.vanhemelrijck@ids.ac.uk)
Core Design Team: **Irene Guijt** – Consultant Learning by Design (iguijt@learningbydesign.org)
Jeremy Holland – Consultant Participatory Research, IDS (j.holland@ids.ac.uk)
Andre Proctor – Consultant Constituent Voice, Keystone Accountability
(andre@keystoneaccountability.org)
Adinda Van Hemelrijck (also co-manager of the project)

The opinions expressed in this report are those of the authors and do not necessarily represent those of the International Fund for Agricultural Development (IFAD). The designations employed and the presentation of material in this report do not imply the expression of any opinion whatsoever on the part of IFAD concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The designations “developed” and “developing” countries are intended for statistical convenience and do not necessarily express a judgement about the stage reached in the development process by a particular country or area.

This report or any part thereof may be reproduced for non-commercial purposes without prior permission from IFAD, provided that the publication or extract therefrom reproduced is attributed to IFAD and the title of this publication is stated in any publication.

Table of Contents

Glossary	4
Introduction	5
1. PIALA purposes and principles	6
2. PIALA components and processes/methods	8
2.1 <i>Components</i>	8
2.2 <i>Processes/methods</i>	8
3. Assessing the PIALA Pre-Pilots as Rigorous Innovation	9
4. Methodological Reflections on the first PIALA Pilot in Vietnam	11
4.1 <i>Defining impact</i>	11
Processes/Methods applied	11
Critical Reflections	11
4.2 <i>Framing the approach</i>	12
Processes/Methods applied	12
Critical Reflections	12
4.3 <i>Describing what happened</i>	13
Processes/Methods applied	14
Critical Reflections	14
4.4 <i>Understanding causes and synthesising findings</i>	18
Processes/Methods applied	19
Critical Reflections	20
4.5 <i>Assessing the outcomes of the Vietnam pilot impact assessment</i>	22
Processes/Methods applied	22
Critical Reflections	23
Conclusion	25
<i>Summary of positive outcomes and required adaptations</i>	25
<i>Capacity & budget considerations for the next pilot in Ghana</i>	27
Annex I. Overview of prevalent Impact Assessment Approaches	30
Annex II. PIALA Quality Assurance Framework	34
Annex III. PIALA Impact Analytical Framework	36
Annex IV. PIA-DBRP Stakeholder Consultations	38
Annex V. DBRP Theory of Change	39
Annex VI. PIALA Generic Research Questions	42
Annex VII. Intended PIALA Uses and Users	44
Annex VIII. PIA-DBRP Specific Research Questions and Methods	45
Annex IX. PIA-DBRP Sampling Structure and Research Schedule	47
Annex X. Table of impact domains, indicators and target groups	53
Annex XI. PIALA Report Outline as applied in DBRP, VietNam	55
Bibliography	57

Glossary

BHH	Business Household
BMGF	Bill & Melinda Gates Foundation
CDT	Core Design Team
CIG	Common Interest Group ¹
DBRP	Doing Business with the Rural Poor (name of the IFAD-funded project in Vietnam)
DPMO	District Project Management Offices of the DBRP project
DSG	Design Support Group
ERG	External Reference Group
FA	Farmer Association
FGD	Focus Group Discussion
IFAD	International Funds for Agricultural Development
HH	Household
ILI	Improved Learning Initiative
IA	Impact Assessment
IE	Impact Evaluation
IMI	Innovation Mainstreaming Initiative
KIIs	Key Informant Interviews
M&E &L	Monitoring & Evaluation & Learning (= MEL)
MEL	Monitoring, Evaluation, Learning (= M&E &L)
PIALA	Participatory Impact Assessment & Learning Approach
PPSC	Provincial Project Steering Committee of the DBRP project
PPMU	Provincial Project Management Unit of the DBRP project
PRA	Participatory Rural Appraisal
RCT	Randomized Control Trial
RIMS	Results & Impact Management System
SME	Small & Medium Enterprise
ToC	Theory of Change
WEAI	Women's Empowerment in Agriculture Index
WU	Women Union

¹ This is a group of farmers, processors and traders that work together to develop short value chains and obtain collective loans with shared risks.

Introduction

1. Under the 9th Replenishment, IFAD committed to moving 80 million rural people out of poverty cumulative from 2010 onwards to 2015, and conducting 30 rigorous impact assessments. Hence the urgent need for appropriate methodologies for impact assessment. To respond to this need, a few piloting initiatives have been launched, one of which is the **Improved Learning Initiative (ILI)**². This initiative aims to develop a *potentially scalable* **Participatory Impact Assessment and Learning Approach (PIALA)** that can help IFAD and its partners collaboratively *assess, explain* and *debate* its contributions to rural poverty impact. The PIALA design and piloting is funded by IFAD's DFID-financed Innovation Mainstreaming Initiative (IMI) and BMGF's Measurement, Learning and Evaluation Unit in the Agricultural Development Program; and with important contributions from IFAD's Country Program Offices and partners in the pilot countries (Vietnam and Ghana), and its Strategy & Knowledge Management and Program Management Departments.
2. According to the IMI's definition, an innovation must be useful and cost-effective³, and demonstrate potential for wider adoption through pilot testing. Although initially developed for summative impact assessment of projects, PIALA's various components and processes are thought of in ways that should allow for flexible adaptation to also serve other purposes of results- or impact-oriented Monitoring, Evaluation & Learning (M&E &L or MEL) of various types of intervention (e.g. projects, programs, policy & advocacy initiatives) in different contexts, and help strengthen the IFAD-financed M&E functions. The ILI seeks to demonstrate the potential for wider adoption of these PIALA components and processes by pre-piloting them in two IFAD-funded projects, drafting a guidance package for further testing, and engaging stakeholders⁴ at the various levels (e.g. grassroots, intermediate, national and global) in critical reflections on the *rigour, utility* and *feasibility* of the pilots in relation to their specific context and MEL requirements. By doing so, it aims to spawn interest for further pilot-testing in various application forms.
3. PIALA has first been piloted in the IFAD-financed project "*Doing Business with the Rural Poor*" (DBRP) in Vietnam in June-August 2013. Reflections have been held during piloting in Vietnam with grassroots stakeholders in every method/process that was applied, and with other project stakeholders and local researchers on the approach as a whole. From September 2013 until January 2014, a series of methodological debriefs and reflections are organised by the Core Design Team (CDT), with the IFAD Design Support Group (DSG) and the External Reference Group (ERG)⁵. Lessons drawn from the first pilot are discussed and compared to other piloted methodologies.

² The ILI is jointly managed by Edward Heinemann – Senior Policy Advisor at IFAD; Richard Caldwell – Head of Measurement, Learning and Evaluation at the Gates Foundation; and Adinda Van Hemelrijck – Consultant Impact Assessment & Learning in Collaborative Settings, IDS. For leading the design and piloting of PIALA, the ILI has established a **Core Design Team (CDT)** that includes as its members: Irene Guijt – Consultant Learning by Design; Andre Proctor – Consultant Constituent Voice, Keystone Accountability; Jeremy Holland – Consultant Participatory Research, IDS; and Adinda Van Hemelrijck. Furthermore, a **Design Support Group (DSG)** has been established to provide the CDT with critical inputs and feedback on the piloting processes and products, and ensure maximum alignment with IFAD's self-evaluation and knowledge management frameworks, systems, methods/tools and initiatives. The DSG includes managers and advisors of IFAD's Strategy & Knowledge Management and Program Management Departments, as well as the Independent Office of Evaluation. Finally, an **External Reference Group (ERG)** was established for supporting the CDT with strategic guidance and feedback, and advocating for PIALA in the wider development and evaluation community. The ERG is comprised of four external advisors with relevant expertise related to the ILI objectives, namely: Robert Chambers – Professor Participation Power and Social Change, IDS; Marie Gaarder -- Public Sector Evaluation Specialist, IEG of the World Bank; Carlos Barahona –Professor and Deputy Director Statistical Service Centre, University of Reading; and Kent Glenzer (Professor Organizational Behaviour and International Development, Monterrey Institute of International Studies, CA).

³ According to IFAD's innovation strategy, this implies that it "*must have positive value for its users. In the case of IFAD, it needs to empower the rural poor to overcome poverty better and more cost-effectively than previous approaches.*" (IFAD, 2007, p. 2, Para 8; see also the IMI ToR 2012)

⁴ Stakeholders are individuals, groups, or organizations who are affected or can influence the achievements of the intervention under concern. At the global level, these are the members of the Design Support Group that include senior members of IFAD's Strategy and Knowledge Management (SKM), Program Management Department (PMD), Independent Office of Evaluation (IOE), as well as the head of impact evaluation and learning of the BMGF's global program on agriculture. At the national and provincial/district levels, stakeholders include IFAD's country program management, members of the project/program steering committees, management offices and implementing line agencies. Local or grassroots stakeholders, finally, include beneficiaries (rural poor women and men) and their organizations/associations and leaders, other important local social and private development actors, and powerful "key system changers" such as brokers and important market players. (see Box 3.1 in: IFAD, 2013: p. 6)

⁵ See Footnote 2.

4. This report presents main findings from these reflections and debriefs, and main options for improvement and adaptation for pilot-testing in the next pilot in the “*Root and Tuber Improvement and Marketing Programme*” (RTIMP) in Ghana early 2014. Decisions need to be made about which of these should be prioritised relative to the objectives and available budgets.

1. PIALA purposes and principles

5. Given its multidimensional approach of rural poverty, and the complicated or even complex nature of the development processes and partnerships it funds and supports⁶, it is quite challenging to measure IFAD’s share in lifting 80 million rural people out of poverty. A fairly flexible and adaptive management and measurement approach can be helpful for capturing the multiple dimensions, variables and interactions at play in its funded projects. To address the challenge of assessing and learning about IFAD’s influences on complicated/complex development processes, PIALA adopts a *complexity-sensitive* and *participatory* approach in a *contribution analysis logic* that draws on *mixed methods*. In supporting IFAD’s self-evaluation system, these must be complementary or give room to incorporate other methods being used or newly piloted within IFAD⁷.
6. PIALA is designed to serve the following three purposes:
 - First, it seeks to rigorously *assess* an intervention’s influences on processes affecting rural poverty impact in a way that captures different stakeholders’ understanding and unpack multiple dimensions of a sustainable poverty impact on the lives of rural poor women and men. This goes beyond a quantitative measure of poverty impact based on income per head.
 - Second, it seeks to help *explain*⁸ why certain impacts occur as expected (or not), and to what extent and why an intervention is able to influence these impacts (or not), taking account of the complexity of interacting actors and factors in the given socio-economic and -political context. Without deeper understanding of context and change patterns, the assessment of whether and how much an investment makes a difference may lead to erroneous conclusions about its perhaps unnoticeably small yet crucial *influences* on processes that may generate impact. It also may miss the opportunity to identify and reach out to other agencies for complementary investments needed to effectuate and sustain these influences.
 - Third, it also seeks to *facilitate debates*⁹ between aid/service providers, their funders and their clients/beneficiaries around impacts and the processes generating these impacts. The purpose is to create space where stakeholders collectively critically reflect on assumptions and debate and learn

⁶ IFAD’s mission is to enable poor rural people to improve their food security and nutrition, raise their incomes and strengthen their resilience by addressing the structural causes that disempower and exclude them from policies and markets. Therefore it works in partnership with governments, civil society, people’s organisations, private sector actors and other donors to increase institutional responsiveness to rural poor people’s constraints and priorities. (IFAD, 2011, pp. 9, 29, 13–14; <http://www.ifad.org/governance>)

⁷ E.g. from the RIMS surveys to participatory M&E, and from RCTs to in-depth qualitative case studies.

⁸ For instance, statistical records of poverty based on household income in Ben Tre Province in Vietnam show that, in the 5 years of implementation of the DBRP project (2008–2013), there has been no real difference in the poverty reduction rate between project districts and non-project districts. Hence one may conclude that the project has not made any difference. However, when analysing primary data on changes in local institutions and relations, market access and livelihoods, one starts to understand why there has been no difference in poverty rates and why/how/where the project has or could have been more effective in changing the conditions for attaining greater impact. First, it is almost impossible to say something about the project’s *attributable* share in poverty reduction since it’s intertwined with many other investments that are part of a much bigger and wider government *strategy* for pro-poor rural development. Second, the project intended to develop and test new interactive mechanisms (such as collective loans to common interest groups involved in short value chain development, and participatory market-oriented development planning processes) that require a fundamental shift in behaviours and relationships involving unknown risks and uncertainties related to market instability and climate change. Hence these mechanisms require more time to incubate and show visible effects, while risks need to be closely monitored and managed in a way that allows for appropriate adaptation. This raises the question if/when it is meaningful to conduct impact assessment, for what purpose/use, and in what form/shape (e.g. a methodology for ongoing impact monitoring and reflection, for formative impact assessment, and/or for ex post impact evaluation).

⁹ In the Vietnam pilot, for instance, project managers and beneficiaries were invited to critically reflect together on evidence revealing that some specific project interventions (such as common interest group support and credit) had hardly reached the poor, due to a lack of trust and capital among group members to share and overcome the risks. Although derived from multiple data sources and validated in the villages and the workshop where the discussion took place, the evidence was contested by some participants who are strongly convinced that risk-sharing in common interest groups must work. This makes people critically debate and reflect on assumptions and evidence of impact.

about the interactions and patterns of behaviour leading to failure/success. Without such debate, critical findings are more likely put away and policy recommendations ignored without due consideration.

7. A key area of innovation for PIALA lies in finding creative ways to move beyond extractive data collection towards more dialogue for examining what success looks like for different actors and generating robust insights about the behaviours and interactions leading to success. Participatory processes and methods are used for not only *collecting* but also *analysing* data and critically *reflecting* on the evidence produced. Albeit limited so far in impact assessment, existing experience with such processes and methods for enhancing poor people's voice¹⁰ in policy-making and development planning at national scale suggests that, *if used appropriately*, this fosters the kind of transparency and responsibility that builds confidence in the relationship between governments and citizens, and enhances an intervention's credibility among the development actors.¹¹ PIALA therefore builds on following key principles:

- *listening* to those whose lives are (directly or indirectly) supposed to be improved;
- *avoiding token participation* for merely data extraction, thus involving different stakeholders (incl. beneficiaries, both women and men equally) in the data collection *and* analysis in a way that is useful to them;
- producing evidence of an intervention's influences on rural poverty impact that is sufficiently *rigorous, contested and debated*¹² and helps understand the interactions and processes generating (or hindering) such impact;
- *amplifying* the voices of the poorest and less powerful (particularly women and minorities) in relation to the more powerful in the critical analysis of change processes and an intervention's influences.

¹⁰ Voice is defined in the literature in various ways. Skuse et al (2007, pp. 3–4) consider 'voice' as "*indicating meaningful inclusion and participation within social, political and economic processes, meaning making, autonomy and expression*". In this definition, voice is a means of empowerment that involves gaining influence on decisions. In instrumental terms, voice is rather conceived as the expression of individual client or constituent satisfaction with the services provided to them, with little or no influence on then decisions. (Ackerman, 2011, pp. 324–325; Rosie McGee & Gaventa, 2011, p. 8) By creating space for debate, voice can be enhanced in both instrumental and empowering terms.

¹¹ This is for instance exemplified in Gaventa & Barrett's publication *So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement* (2010), which is a review of 100 case studies of citizen engagement in governance across 20 countries.

¹² In a critique of radical approaches to participation and empowerment in development, Mohan and Hickey (2004, pp. 61–63) plead for a critical engagement based on knowledge produced through participatory research that is "*rigorous, debated and contested*".

2. PIALA components and processes/methods

2.1 Components

8. The PIALA package is composed of the following standardized components for which guidance is developed:
- An *impact analytical framework* (see [Annex III](#)) that helps link intervention-specific impact dimensions to the sustainable livelihoods and RIMS rural poverty indicators¹³, the Women's Empowerment in Agriculture (WEIA) Index¹⁴ and impact indicators used by the M&E Harmonisation group¹⁵.
 - A *Theory of Change (ToC) articulation process* that engages stakeholders in (re)constructing their collective understanding of an intervention's envisioned changes towards impact and helps focus and frame the impact assessment while allowing for a wider complexity-sensitive lens.
 - A *sampling strategy* that enables aggregation and thus trend analysis and pattern recognition at intervention level.
 - A *mixed-methods approach* that sequences participatory methods, micro-surveys and KIIs in ways that help reconstruct the actual changes and causal flows by bringing together multiple perspectives (e.g. women and men, farm laborers and contracting farms) and cross-checking different data sources (incl. primary and secondary).
 - A *mixed-data collation approach* that helps researchers transcribe and organize data from multiple sources/methods in a way that identifies *actual* (as compared to *presumed*) change pathways and *actual* (as compared to *claimed*) project contributions to the observed changes.
 - A *participatory sensemaking model* that involves the different stakeholders in a collective analysis of and reflection on the evidence of changes and contributions emergent from the mixed-data collation, using the ToC as a guiding structure.
 - A *reporting outline* that helps researchers present the evidence and explain causes and influences in a way that reflect the different perspectives of actual changes and relative project contributions to these changes.

These components can serve formative impact assessment informing the design, expansion or adaptation of an intervention, as well as summative impact evaluation conducted at completion. They can be used separately for various impact-oriented M&E activities, yet can form the basis for developing an impact M&E system integrated with the design of a project, program or policy intervention and embedding other methods (such as constituent feedback, sensemaker-based monitoring, or quasi-experiments) in its mixed-methods package. The processes by which they are used will slightly differ depending on the specific M&E function they serve at a particular moment in time of the project, program or policy initiative that is being assessed.

2.2 Processes/methods

9. Starting from a globally agreed impact analytical framework (see [Annex III](#)), an intervention's implicit or explicit Theory of Change (ToC) is articulated and visualised through a process of consultations and group discussions around relevant documents with the stakeholders involved in the management and implementation of the intervention.¹⁶ This entails a reconstruction of how these actors *envision* change happens or should have happened at the moment of the impact assessment. The process and product is not meant to be a perfect reconstruction, but to provide a structure that enables

¹³ E.g. food security, household income and assets, child nutrition and people out of poverty.

¹⁴ The Women's Empowerment in Agriculture Index (WEAI) is a composite tool developed by the United States Agency for International Development (USAID) in collaboration with the International Food Policy Research Institute (IFPRI) and the Oxford Poverty and Human Development Initiative (OPHI), for measuring women's control over critical parts of their lives in household, community and the broader economy.

¹⁵ See: The M&E Harmonization Group of Food Security Partners, 2013.

¹⁶ In Vietnam, these included national and provincial members of the project steering committee, provincial and district-level members of the project management unit/offices, and officials from the provincial line agencies responsible for project implementation (see [Annex IV](#) for an overview of the stakeholder consultations).

all stakeholders to engage in assessing actual (as compared to envisioned) changes and valuing actual (as compared to claimed) contributions/influences of the intervention. The ToC helps understand the broader picture and determine and reach agreement about the focus, purposes, questions/indicators and standards for evaluating the intervention's influences on important changes that affect rural poverty impact.

10. The resulting ToC diagram is then used to guide data collection on actual changes, causes and impacts in intervention as well as non-intervention areas for comparison. The types of data needed for answering the questions are collected through an appropriate selection of methods that complement and strengthen each other.¹⁷ Multistage cluster sampling¹⁸ serves to enable inference at intervention/country level for generating statistics of rural poverty impacts, while prompting participatory analysis of processes that generate these impacts. The assumption is that participatory methods can produce data generalizable to the intervention population, if applied in a standardized manner across a representative sample, thus avoiding bias and capturing variability of populations, locations and impact patterns.¹⁹ If methods are selected and sequenced appropriately, these can generate rich explanations and stories as well as quantified estimations that are sufficiently rigorous and convincing for global reporting, advocacy, and learning and knowledge-sharing purposes.
11. Combined with mini-surveys with primary and secondary beneficiaries and in-depth interviews with market or other systemic actors, and triangulated with relevant secondary data sources, PRA-based methods are applied for engaging beneficiaries and their organisations and leaders in the data collection and analysis at community level. The selected methods help them visualise what and how they think changes occur (or not) and how these affect their lives and their ability to shape their livelihoods. This involves a reconstruction of how change *actually* happens according to those who are supposed to benefit. By systematically linking data to the ToC, missing or incomplete data and information to be compared with other sources, can be identified. To increase confidence in causal inference, the ToC is used for probing assumptions and plausible explanations of impact (or non-impact), cross-checking various data sources (primary and secondary, qualitative and quantitative), and comparing different configurations of case-based factual and counterfactual data units.
12. Finally, actual changes are compared to the envisioned changes in the ToC, and the likely explanations disputed and discussed in a process of participatory sensemaking. The purpose of this process is not to arrive in itself at rigorous causal inference, but to produce an additional layer of feedback data that helps generate *rigorous, contested and debated* knowledge of the complex interactions leading to impact. While the former processes of defining impact, ToC articulation and data collection are happening in the home bases of the different stakeholder groups, participatory sensemaking takes place in collective spaces where both aid/service providers and clients/beneficiaries meet and debate the evidence. The assumption is that this stimulates a greater uptake or use of findings for making projections and decisions, and if done well, creates opportunity for beneficiaries to testify and (directly or indirectly) hold aid providers accountable.

3. Assessing the PIALA Pre-Pilots as Rigorous Innovation

13. PIALA is a methodological and process innovation that fits with IFAD's expectations and needs regarding evidence-based reporting, learning and knowledge sharing with partners and other development actors for improving development practice and performance.²⁰ To ensure its quality while acknowledging 'real world' conditions under which impact assessments (or impact-oriented M&E) generally take place, a new concept of *rigorous thinking* is developed and tested that is broader than the definition of rigour as statistically verifiable attribution related to a specific method's procedure. While useful to measure the direct effects of pure technical or science-based interventions

¹⁷ E.g. Stern et al., 2012, pp. 18, 33.

¹⁸ Defining the units that need to be sampled for evaluating more complex change processes generally require more thought than in conventional surveys of relative simple projects. PIALA adopts a multi-stage sampling approach that involves a purposive sampling of agro-ecological zones, of sub-project administrative areas within these zones, and of systemic units within these areas; subsequently, a stratified random sampling of geographically-bounded communities within the selected systems/areas; and finally, a purposive sampling of social groups and a stratified random sampling of households within these communities.

¹⁹ Holland, 2013; Levy & Barahona, 2002.

²⁰ E.g. IFAD, 2011b, 2012.

(such as immunisation as a result of vaccination), methods that follow a strict scientific-statistical procedure appear cumbersome in areas affected by multiple and overlapping actors and investors in more complex political contexts where it is almost impossible to isolate and control a single project's variables. Moreover, rigour defined in statistical terms does not include all types of flaws and biases (other than those related to sampling/selection and attribution) that may occur in evaluative research using *either* or *both* quantitative and qualitative methods and processes.²¹ The concept of rigorous thinking that is tested in the PIALA pre-pilots, refers to what is also called *quality of thought* and applies to all methods, processes and evidence being produced and used to *assess, explain* and *debate* an intervention's influences on rural poverty impact.²²

14. The premise is that a participatory impact assessment and learning approach must use methods and processes that consistently can generate valid evidence that speaks to the hypothetical change mechanisms and indicators of an intervention's Theory of Change, while enabling equal and meaningful participation to create greater value or utility. This implies thinking carefully about who participates when and where for producing what kinds of evidence related to which questions/indicators, thus setting proper boundaries while dealing with political influences and organisational pressures. Furthermore, it also requires the probing of assumptions, cross-checking of different sources and perspectives, and facilitating participation in a power- & gender sensitive manner.
15. In the PIALA pilots, rigorous thinking is therefore piloted as a way to balance various standards of:
 - *rigor* –referring to the thoroughness, consistency and reliability of methods, processes and evidence for generating knowledge that is gender-sensitive, has sufficient explanatory power and is robust enough for reporting and program and policy decisions;
 - *utility* –implying accessibility, credibility and added value of methods, processes and evidence for generating knowledge that is useful for key stakeholders to influence decisions affecting impact;
 - *feasibility* –referring to the replicability, manageability and cost-effectiveness of methods and processes (incl. time, efforts and competencies needed) for adoption at a wider scale.
16. Based on a review of the literature, an integrated PIALA Quality Assurance Framework (see table in Annex II) is developed that formulates questions for assessing if/how the PIALA pilots perform on these three quality dimensions in four key areas of evaluation planning and management²³ –i.e. (i) *defining impact* (or setting the focus of the impact assessment), (ii) *framing the approach* (or identifying the questions/indicators, purposes and standards for the impact assessment); (iii) *describing what happens* (or defining the sampling structure/size and selecting and sequencing methods and processes for data collection related to the questions/indicators); and (iv) *understanding causes and synthesizing findings* (i.e. data collation, sensemaking and causal analysis). In addition, the fifth area of quality assurance looks at the outcomes of the impact assessment in terms the evidence produced and its use for influencing decisions, as well as the added-value-for-money created for strengthening IFAD's self-evaluation system. This framework was used to guide the reflections on the first PIALA pilot in Vietnam.
17. The remainder of this report consists of a summary of the critical findings from these reflections related to each of these five areas (Section 4). For each area, the main methods and processes that were used are presented, followed by the critical reflections. Upfront a summary statement responds to the main questions for that area (taken from the left column derived in the PIALA Quality Assurance Framework). The report concludes with a brief overview of positive outcomes and key points and ideas for improvement (Section 5).

²¹ Annex I provides an overview of impact evaluation approaches and their specific aspects of rigour. For a description of various types of error and bias, see also: Chambers, 2008; Copestake, 2013; White & Phillips, 2012; Woolcock, 2009.

²² Cf. Mertens et al, 2010; Patton, 2008, 2012; Sweetman, Badiee, & Creswell, 2010; Zelik et al., 2007; Midgley, 2007.

²³ Cf. the [BetterEvaluation taxonomy](http://betterevaluation.org/plan) for planning & managing evaluations (<http://betterevaluation.org/plan>).

4. Methodological Reflections on the first PIALA Pilot in Vietnam

4.1 Defining impact

How and to what extent did PIALA help adequately identify the changes and influences that the impact assessment should focus on?

Reflections with project management and researchers suggest that the ToC approach is essential for creating a shared understanding of the project's envisioned changes and intended contributions at the time of the assessment. It helps project stakeholders see these changes in a broader context, and determine which ones should be the focus of assessment. In Vietnam, ToC articulation however happened simultaneous with the research design and training, and insufficiently engaged the lead researchers, which affected their ability to use the ToC to guide and focus the fieldwork and analysis. Since Vietnam was the first pilot in which design, training, implementation and testing were carried out all simultaneously, it was impossible yet to have all tasks and responsibilities performed in a logically sequenced way as would be the case in a normalised PIALA process. In the next pilot in Ghana, sufficient time should be provided to the researchers to engage in desk review and ToC articulation resulting in a well framed and focused impact assessment design.

Processes/Methods applied

18. An Impact Analytical Framework was developed on the basis of what was agreed with global IFAD & BMGF stakeholders²⁴ in Rome (see Annex III). The Framework serves to link the project-specific ToC and its impact dimensions to the RIMS rural poverty indicators, the Women's Empowerment in Agriculture (WEIA) Index and indicators used by the M&E Harmonisation group.
19. A process of ToC articulation and visualisation took place through stakeholder consultations (see Annex IV) around relevant project documents (incl. Logframe) at national, provincial and district levels. The resulting diagram (see Annex V) was validated by the PPMU and project M&E officers at the districts.²⁵ The ToC helped visualising: (a) the changes that are *expected* to generate impacts; (b) the causes or mechanisms that *presumably* trigger these changes, and the DBRP project's *intended* or *envisioned* contributions to these; and (c) the assumptions underpinning the links between the changes and causes in the ToC. These were probed and validated through the data collection and sensemaking processes together with key stakeholders, by attaching the evidence produced to its 5 impact domains (see Paras 20 and 53).

Critical Reflections

20. Project stakeholders very much appreciated the use of a ToC approach to visualise and situate the envisioned change processes and intended project contributions in a broader context. The data collected in the villages (in particular the causal flow diagrams developed in the focus group discussions) served to probe the ToC that was produced through national and provincial stakeholder consultations. The ToC was also essential for the participatory sensemaking and considered by the Vietnam partners as a potentially powerful instrument for involving key stakeholders not only in impact assessment but also in project design and planning and annual outcome monitoring. Stakeholders' views of impact and changes towards impact were included in the ToC diagram through the sensemaking process (see Section 4.4).
21. The researchers didn't fully grasp the envisioned changes and intended contributions of the DBRP project, due to their limited involvement in ToC articulation. In future PIALA applications, the local research organisation's coordinator and senior lead researchers who are responsible for the final reporting, should lead this process on the basis of a proper desk review of project documents prior to any field work. Since Vietnam was the first pilot in which design, training, implementation and testing were carried out all simultaneously, it was not possible to have all roles and responsibilities performed as would be the case in a normalised PIALA process. Hence the desk review and stakeholder consultations (see Annex IV) were conducted by the CDT; while ToC articulation happened partly

²⁴ See footnote 4 for global stakeholders.

²⁵ These are part of the local Vietnamese government institutions.

during the training interactively with the researchers and project M&E officers. To conclude: an important learning point for the next pilot is to provide enough time and guidance for the researchers to conduct a proper desk review and ToC articulation.

22. The scope of the impact assessment of the DBRP project was probably too wide and thus the focus not clear and narrow enough defined. Consequently, the researchers were overwhelmed by data. This was due to the fact that ToC articulation happened slightly too late to provide enough focus for the framing of the assessment. The framing was more guided by the generic research questions and impact analytical framework in the PIALA research strategy than the project's ToC and consultations with the project designers and managers.
23. The process of ToC articulation has not been systematized yet, and forms a gap in the PIALA guidance drafted so far. This is an important working point if the ToC approach has to be replicable. Its success in the Vietnam pilot was a product of 'art' more than 'procedure'. Although art and competence are essential to ensure quality in particular in hybrid mixed-methods impact evaluation designs, we do aim to develop procedures that enable local researchers to use the PIALA methods and processes in a standardized manner.

4.2 Framing the approach

How and to what extent did PIALA help adequately determine the purposes, questions and standards that should guide the impact assessment?

The PIALA generic questions and impact analytical framework were helpful in formulating specific research questions for the Vietnam pilot and linking them to specific purposes and methods. Reflections with participants showed that processes and methods were found useful if these purposes were clearly explained. Missing in the assessment framing though were standards and weights for valuing project contributions, which should have been defined together with project management in the process of ToC articulation.

Processes/Methods applied

24. Based on the discussions around impact dimensions held with global IFAD and BMGF stakeholders in Rome October 2012 (see Para 18), generic questions and intended uses and users were identified, which were further refined and elaborated in the PIALA research strategy²⁶ (see Annex VI for questions and Annex VII for intended users).
25. Together with the local research team and the project M&E team, more specific questions were formulated in relation to the main impact domains identified in the ToC of the DBRP project, which were then linked to specific methods (see Annex VIII).
26. Also purposes or intended uses of the various methods and processes of the impact assessment were determined iteratively during the training and testing. During fieldwork, the purposes were explained to all participants in the group discussions in the villages and the provincial sensemaking workshop.

Critical Reflections

27. The generic questions in the PIALA research strategy were very helpful as a checklist to ensure that the pilot-specific questions were properly aligned with the overall impact analytical framework. A lot of time was spent with the researchers on framing these specific questions and linking them to the various methods. This was very much appreciated and needed for outlining the sequence in which the various methods would be used and drafting the village research schedule (see Annex IX). Moreover it was crucial to help the researchers internalise the analytical framework before going to the field. The generic questions and analytical framework will be revised though for the next Ghana pilot, as to incorporate changes in global impact indicators and lessons from Vietnam.

²⁶ Cf. IFAD, 2013.

28. From the reflections held at the end of every exercise during the participatory research conducted in the villages, it appeared that, quite consistently, participants found the exercises useful²⁷ when the purpose was made clear. Although these reflections may be affected to some extent by a “political correctness” bias²⁸, from the discussions it was obvious that in general people really liked to participate because it created an opportunity to exchange and discuss new information. Furthermore, it could be observed that there was great energy and enthusiasm among the participants in most focus group discussions as well as in the feedback and sensemaking workshops in the villages and at the province.²⁹
29. Missing in the framing of the assessment approach were standards for valuing project contributions to observed changes, based on a weighting of relative importance of these changes (relative to the investments made) to realize the expected impacts³⁰. Critically valuing project contributions together with key stakeholders was difficult in the absence of agreed standards and weights. Such standards and weights should be defined during ToC articulation.

4.3 Describing what happened

How and to what extent did PIALA help rigorously collect, triangulate and collate the data needed to produce evidence for confidently answering the assessment questions?

A sampling strategy was developed for the Vietnam pilot that is robust enough to enable rigorous causal inference. It can do so, however, in so far impact indicators are relevant to the project context and project contributions are evaluable in relation to these indicators. In the case of Vietnam, this was questionable, since project investments were intertwined with many other investments as part of a much bigger and wider government *strategy* for pro-poor rural development. Moreover, project incubation was insufficiently taken into account in the sampling of the focus villages, while critical aspects of context and project variability were not well enough considered in the selection of the non-focus villages.³¹ Finally, while the survey sample size for data collection on the rural poverty impact indicators was sufficient for establishing generalizability and comparison with the RIMS survey, the participatory-explanatory research sample data was more limited in scope and likely insufficient to cover project variability. To conclude: more time must be dedicated upfront to understand project distribution, incubation and integration with other programs, in order to determine the project’s evaluability and accordingly appropriate sample sizes and criteria.

Substantial training was provided, actively involving researchers in design and fieldtesting of the assessment. Appropriate methods were selected, enabling sufficient data collection on research questions and impact

²⁷ Usefulness is defined in terms of accessibility, credibility and value for the users (see PIALA Quality Assurance Framework in Annex I).

²⁸ “Social acceptability” or “political correctness” bias occurs when “people provide responses which reflect what they regard as being the socially acceptable thing to say. This kind of bias is likely to appear if there is a prevailing consensus concerning how things have occurred, or concerning the role that a particular agency has played in achieving outcomes.” (White & Phillips, 2012, p. 21) Generally, a bias occurs when it systematically happens that data are provided, presented and/or interpreted slightly differently from what is really the case, which may affect findings and conclusions. (Ibid, p. 22)

²⁹ Many participants had to make serious efforts, such as taking unpaid leave while their allowance (30,000 VND) didn’t cover the loss of income (min. 60,000 VND). Moreover, there was little to gain in terms of social status, as people are used to attend meetings in the various village and self-management groups. Participants who were seriously delayed by job or other obligations were visibly disappointed when they realised it was too late to join the group, and some didn’t want to leave, even though they were given the cash allowance.

³⁰ For instance, the DBRP project made serious investments in commune-based infrastructure development, which were in some communes bigger than in others, depending on the communes’ needs and wants, identified through their annual SEDP processes. This presumably would lead to greater access to local markets and stimulate local trade and businesses, creating new job and income opportunities for the local people, particularly the poor. Significant change was observed in access to markets, trade and small business development in the villages where in-depth research was conducted. Project achievements in this domain were higher in some villages compared to others, due to different levels and priorities of investment. But it didn’t equally lead to significant change in the domain of livelihood and job opportunities for the poor, for a number of reasons, which were partly out of the project’s control, but partly would have required a longer term investment, a better direct-targeting strategy, and a more adaptive management approach. In terms of relative importance to realise rural poverty impact, as for the project’s ToC, changes in “livelihood and job opportunities” in poor people’s lives has more weight than changes in “access to markets”.

³¹ Focus villages are geographically targeted by DBRP village and commune-level interventions. Non-focus villages are not geographically targeted but exposed to the DBRP by its interventions at district and provincial levels.

domains in the ToC. Detailed guidance notes and research schedules were drafted, supporting consistent and systematic application of selected methods, questionnaires and processes. Overall the research was conducted in a rigorous manner. Also ethical principles of research in general (such as prior consent, safety and confidentiality) and of participatory-transformative research in particular (such as equality, gender/power sensitivity and reciprocity) were overall well respected. Tensions appeared though between data and process, truth and multiplicity, and independence and ownership that challenged the researchers' ability to apply the methodology in a rigorous manner. Simplification together with more training and guidance is needed, therefore, that can make PIALA more doable in difficult and politicised evaluation contexts. Sequencing participatory and survey methods may help ease the tensions, while providing more time for interim data processing.

Processes/Methods applied

30. On the basis of the initial sampling framework in the PIALA research strategy and discussions held with experts³² in the PIALA external reference group, a sampling strategy was developed together with the stats team of the local research organization and the project M&E team that covered the 26 communes where the project was started in 2008 (see Annex IX).³³ From these 26 communes that were included in the sample, in total 720 households were surveyed in 24 villages, of which 540 factual and 180 counterfactual, and approximately³⁴ 520 village/commune members participated in the in-depth research, of which 390 factual and 130 counterfactual.
31. The research team received 3 days introduction training on PIALA concepts, methods, principles and ethics; 3 days dry-run training on methods; 3 days field-testing training; and finally, 3 days reflection and revision of methods and guidance. In total, this made up for 12 days of intensive training and co-design of the research. The project M&E team attended most of the sessions and accompanied the research team during field-testing.
32. PRA-based methods and a mini-survey were designed to fit the project-specific impact domains and research questions that were identified in relation to the ToC and the PIALA analytical framework. This was done iteratively and interactively through the process of training, field-testing and reflection together with the local research team and project M&E team.
33. Questionnaires were developed for semi-structured key informant and group interviews. A household survey questionnaire of limited size (therefore called 'mini') was drafted and reviewed after field-testing, that included questions on rural poverty indicators (food, income and assets) and crucial leverage point indicators in the ToC (incl. WEI).
34. Guidance notes were drafted, tested and reviewed together with researchers for facilitation and note taking for each single method (including PIALA ethics), and for daily and weekly team reflections on data quality/sufficiency and process/participation issues.

Critical Reflections

35. The sampling had a well thought out structure, in which contextual variability was taken into account and selection and research bias avoided. The sampling frame of 26 communes where the project was started in 2008 covered three agro-ecological zones of Ben Tre (salt, brackish and fresh water). These were identified in a consultation meeting with the DBRP M&E team. Because of budget and time limitations, the district sample was restricted to 3 districts from a total of 8 project districts, i.e. one district randomly sampled from each of the three 3 agro-ecological zones. For each selected zone/district, project communes were first stratified by distance³⁵ from the main inter-communal road

³² In particular with Carlos Barahona, deputy director of the Statistical Services Centre at the University of Reading (UK) and members of the external reference group.

³³ One of the criteria for selecting projects for piloting PIALA was that they should have had been implemented for at least 5 years. Hence the remaining 24 communes that were added to the DBRP portfolio in 2010, were not included in the sample. In these communes, the project had been implemented for only 3 years and presumably were less mature in terms of impact. These criteria were agreed at the consultation workshop in Rome on 17 Oct 2012, and are also in the PIALA research strategy (see: IFAD, 2013).

³⁴ Group compositions slightly differed from village to village due to a self-selection margin of average 2 people per focus group.

³⁵ We classified relatively close if distance was less than 2km, and relatively far if distance was greater than 2km. We ended having about 50-50 distribution of close and distant communes.

as a proxy for market accessibility, before randomly selecting 2 from each group in each selected district.

36. Within the sampled communes, data were collected quite systematically and consistently, using the same research schedule and sequence of methods, facilitating the same group discussions and using the methods and questionnaires similarly across all sampled villages, communes and districts. The methods were sufficient and appropriate to collect all data needed to describe changes and identify explanations related to the impact domains in the ToC. Team reflections were held every evening between the teams working in the same district.
37. Overall the research was conducted in a way that was respectful of general ethical research principles (such as prior consent, safety and confidentiality) and of ethics more specific to transformative research³⁶ (such as transparency, equality, gender/power sensitivity, and reciprocity). Group compositions and processes have been guarded carefully to enable equal participation while ensuring consistency in data sourcing and avoiding external influences. Although the local authorities had lists of the participants in the focus group discussions and the households sampled for the survey, data could not be linked to individuals. Survey and key informant interviews were taking place in closed spaces that were sufficiently safe and private (such as in private houses, or in the village meeting place when there were no group discussions). Key informant interviews and group discussions systematically started with an explanation of the purpose and a request for participants' consent, and ended with a reflection on the relevance of the findings and usefulness of the exercise. Sufficient attention was paid to silent or dominant voices, as to ensure participants were given equal opportunity to speak. Flipcharts with diagrams and matrices produced in these sessions were returned to the participants. Preliminary syntheses of findings from the participatory inquiry in each village were presented back to all participants for validation and reflection in a village feedback workshop.
38. Conceivably a sample bias may have occurred due to insufficient consideration of the project's incubation time in the sampling strategy. This was not apparent from the available project documentation. Incubation refers to the start-up times that various interventions need in order to function well, relative to the time frame in which project interventions are supposed to be completed, and the sequencing of interventions over time in the different geographic units.³⁷ In the sampling of project communes, incubation was insufficiently taken into account due to the restriction of the sampling population to the 26 project communes where the project had started in 2008.³⁸ Project management repeatedly asked for full coverage of all 50 communes in the 8 districts where the project had been implemented. However, the CDT did not receive a clear explanation why it had to deviate from the basic selection criteria and sampling structure agreed in the research strategy³⁹.
39. Also in the sampling of non-focus communes, variables important to enable comparison of poverty impact between focus- and non-focus villages (such as initial poverty rates) were insufficiently taken into account. This was due to inadequate understanding of the relative distribution of the project within districts –i.e the proportionality of implementation/treatment of specific project interventions across geographic units.⁴⁰ To address the challenge of comparability between focus and non-focus

³⁶ Transformative research aims at contributing to social justice through an inclusive and power-sensitive design and conduct. (Mertens, 2009, pp. 5, 66–67, 2010, p. 12) Drawing on the premise that there are substantial power differentials among and between researchers and researched affecting the research processes and outcomes, a transformative approach adheres to validity standards that link ethical principles of participation, such as equality and reciprocity, to epistemological principles, such as critical thinking and reflective practice.

³⁷ For instance in the DBRP project, the early project communes were rather slow in implementation due to resistance and lack of implementation and management capacities at higher levels, while the second phase (post 2010) communes learned from the previous ones and could rely on more matured project management and better trained government officials at district levels. Because we didn't include post 2010 communes in our sample, we can't say anything about the differences in project effectiveness or contribution to crucial change processes (f.i. changes in relationships between local institutions and villagers creating a positive dynamic) leading to rural poverty impact.

³⁸ See footnote 34.

³⁹ See footnote 34.

⁴⁰ This also coincides with what is called geographic and self-targeting in IFAD's global targeting policy. For instance, the amount of DBRP's commune investment fund for each project commune was determined based on a set of criteria and thus different from commune to commune, producing different outcomes related to market access and job opportunities that affect wealth & wellbeing impacts for poor households. Because we didn't take into account this proportionality of project implementation in our sample, we may have had a sample bias.

villages, propensity score matching or some kind of matching process might have been helpful. However, when there is insufficient homogeneity in the project target population, context or treatment (as appeared to be the case in the DBRP), adequate matching of non-focus villages for statistical comparison of data on individual poverty impact indicators may not be possible.⁴¹ The matching sample size would need to be big enough to capture all variability, which significantly would increase the cost of the impact assessment. Moreover, the variables to be taken into account for matching may be too many. Even in such cases of high heterogeneity, it remains important though to enable valid comparison of areas with more/less project presence to better understand how different configurations *systemically* lead to different outcomes. For this, sufficient clarity about critical areas of variability and distribution for appropriate selection is crucial, which could have been obtained in the Vietnam pilot by spending more time on ToC articulation with project staff and stakeholders, before starting with the sampling.

40. A better understanding of the project's distribution of investments and incubation time of the mechanisms it tries to influence, may also have provided a better sense of the DBRP's *evaluability* as a stand-alone project. For the DBRP, it appeared to be quasi impossible to isolate and generate quantitative evidence of the project's *attributable* share in food and income security increase, since it's intertwined with many other investments that are part of a much bigger and wider government *strategy* for pro-poor rural development. Moreover, new mechanisms (such as collective loans and group development) require more time to incubate and show visible effects on rural poverty impact.
41. Participatory numbers generated for individual data points are only generalizable if collected from a sufficiently large (random) sample of individual respondents. This can be done but is not always preferable to a mini survey, as opted for in this PIALA design. In this case, sequencing with the survey data⁴² can ensure that participatory research facilitates group-based analysis of causal mechanisms that explain impact data. While survey-based data collection on rural poverty indicators in Vietnam was sufficient for establishing generalizability and comparison with the 900 household RIMS survey⁴³, participatory data collection on causal mechanisms was more limited in scope, taking place in a sub sample of the mini survey communities (6 focus and 2 non-focus villages). *Was this sufficient to cover project variability and allow for generalizability of causal inference?*
42. Although the numbers generated through the group discussions did have value as perceptual data, eliciting different people's and group's views and estimations of the magnitude of recurring changes and events, these types of numbers are group-generated and not 'private' individual data points, and so and cannot be aggregated to allow quantitative inference for the whole project population.⁴⁴ While there is a case for slightly increasing the number of communities engaged in in-depth participatory research in order to capture the variety of project contexts and mixes of interventions, however, this is not merely a numbers game in the same way as the household survey sampling approach. There is also a trade off here: participatory data collection in a larger sample of communities would be considerably more resource-intensive and risk sacrificing depth of analysis for breadth of coverage.
43. To evaluate project contribution relative to global poverty indicators, these indicators need to be relevant to the project context. Farmers in Vietnam, for instance, emphasize a balanced spiritual and family life that is not merely dependent on income increase but also requires opportunities to create a livelihood that is sustainable, help them shape their lives within their own localities and prevent what they call "social evil" such as alcoholism and intra-household gender-based violence. Although contractual jobs may offer landless farmers an opportunity to increase their income and assets, these often involve financial and physical risks that affect people's health and stability of family life, and thus are not considered positive outcomes. Moreover, food security is not an issue in the Mekong delta, as the DBRP baseline and other studies show that more than 92% of the rural poor in the project area are food secure. Hence it would be helpful to find ways to operationalize the RIMS rural poverty indicators in a contextualised manner.

⁴¹ Cf. White & Phillips, 2012: p 5.

⁴² This sequencing was not achieved in this case, as discussed in para 51.

⁴³ Albeit PIALA cluster-sampled more households per community and fewer communities overall.

⁴⁴ The distinction between 'private' and 'public' data is important here when considering group-based participants vs individual respondents. Some 'public' data can be generated efficiently and accurately by groups in a way that elicits 100% data coverage of given populations. One example is that of mapping publicly-known household characteristics (such as food security status) for an entire village using a group of key informants (see Barahona, 2013, pp137-146)

44. The prioritisation of household-level surveys and village-level participatory research led to a lack of time spent interviewing market actors and decision makers beyond the household and village in order to understand the upstream and downstream changes in behaviour and relationships at the level of the market system, crucial for testing the Theory of Change and understanding the cause-effect relationships. This meant that the research prioritised local level change processes over higher-level systemic changes in the value chains, market transactions and policy processes. This will need greater consideration in the Ghana pre-pilot.
45. Challenging for PIALA piloting were the standard bureaucratic planning procedures for field research applied in Vietnam. For obtaining official permission to conduct field research, for instance, a day-to-day plan was required detailing all focus group discussions, village walks, household visits and individual and group interviews. Since the methodology has to be consistent across the entire sample in order to ensure research validity, any change had major implications for the agreed schedules and thus needed to be considered and defended carefully. Hence a great deal of diplomacy was needed to convince authorities of the need to change schedules due to the adaptation of a method, process or sequence. The researchers were not always aware of this, which caused some problems for project management.
46. Also challenging to maintain the independence⁴⁵ of the evaluation was the controlled access to villages and households. Although quite collaborative and supportive⁴⁶, local leaders and project staff were omnipresent during fieldwork.
- Village leaders felt responsible for creating good conditions for the research –e.g. collecting the participants, solving language issues and other conflicts, adjusting meeting schedules, and removing disturbing actors and factors. Based on their experience with participatory planning, they sensed what could go wrong or trigger participants to leave or refuse to participate (which happened in a few cases). They were mostly motivated by electoral prestige⁴⁷ vis-à-vis their villagers. Their presence and efforts were much appreciated by the participants.
 - Project M&E staff justified their presence during fieldwork by their requisite aid to the research teams for obtaining formal permissions and administrative documents⁴⁸ from the authorities, and solving logistical hurdles. They wanted to participate in the design, framing, training, field-testing and conduct of the PIALA pilot for learning purposes. Furthermore, they also intended to help explain research purposes and questions to avoid bias and misunderstanding due language issues and cultural differences between researchers and villagers (cf. Para 48).
 - The researchers felt that local authorities and project staff exerted too much control over the field research and requested them to leave the focus group meetings. Leaders and staff responded positively, yet never remained far away from the meeting place. This may (or may not) have influenced the natural unfolding and confidentiality of the conversations, creating a “political correctness” bias⁴⁹, though final outcomes of the impact assessment are unlikely to be affected, since data were triangulated and findings validated within and across villages in multiple and different ways.

Proper guidance and training is needed for the lead researchers on how to deal with power and politics in ‘real world’ impact assessments, and find the right balance between requirements of ‘independence’ for reducing internal project bias and ensuring research ethics on the one hand, and ‘stakeholder

⁴⁵ According to the DAC/OECD (2002, pp. 24–25) this implies research that is free from political influence or organisational pressure. The remain challenging in every impact evaluation context.

⁴⁶ Due to a general change in Vietnam’s political climate in pas years towards more democratic citizen engagement and younger and more female leadership in governance, local and provincial authorities are generally supportive of participatory research and showed genuine interest in the possible added value that PIALA offers in this respect.

⁴⁷ In Vietnam, village leaders are elected by the people, are relatively close to the villagers, and have little power compared to higher level officials. In the 1960s Vietnam’s communist class struggle broke the power of the traditional village elite by electing new village leaders from among the peasants and redistributing wealth (particularly land).

⁴⁸ In the few occasions where inaccurate information was provided, it was likely due to local management deficiencies rather than to deliberate withholding, since the information was obtained fairly easy when requested through the appropriate channels in the hierarchy.

⁴⁹ If this happened systematically across the entire sample, then such a bias may have occurred. See footnote 28.

engagement' for reducing external research bias, enhancing research efficiency and creating ownership for learning purposes on the other.

47. Differences in language use and communication style between researchers and villagers created tensions and noticeably influenced the group dynamics. Naturally, in the beginning, villagers are suspicious about the nature of the research and the motivation of the researchers. Adding to this was the fact that the researchers spoke a more educated and higher class Vietnamese that is common in the Northern capital Hanoi. Moreover, researchers' limited knowledge of project activities and context in the villages made it more difficult to establish rapport and bridge these differences. Questions were often not understood and the researchers were asked frequently by participants and village leaders to speak slower, use simpler sentences and more contextualized examples. These tensions occurred in all the villages where in-depth participatory research was conducted. Researchers' facilitation skills helped overcome these tensions and gain participants' trust. In the next pilot and future applications of PIALA, sufficient attention should be paid to these cultural-political and linguistic differences and preference should be given to research teams who have locally based research assistants.
48. The researchers struggled with the principle of triangulation. Mostly coming from a positivist quantitative research background, they considered it more as a means of verification in search of 'truth', rather than a way to enrich the data. Triangulation⁵⁰ however is a means for getting a richer picture, building off various perspectives for generating more trustworthy findings that help identify and eliminate plausible explanations, while recognising and interpreting the positionality of the different actors. Although the reflection guidance helped them think broader and search for more explanations by looking at multiple data sources, rather than drawing conclusions based on a single source, it remained difficult for the researchers to understand this principle.
49. A tension also occurred in the course of the data collection between different purposes of the assessment –e.g. between process requirements for enabling equal participation and restraint political influence on the one hand, and data requirements for rigorous analysis and reporting demanding a more extractive approach on the other. Under time pressure it was challenging for the researchers to find a balance without losing on rigour of either processes or data –both of which may affect validity of evaluation findings. Although enormous efforts and progress were made with regard to facilitation, researchers' limited experience with participatory research did play out here. When confronted with time limitations, they tended to apply methods in a more mechanical way and ignore power dynamics. In other occasions, researchers lost track of time as they put too much effort and attention into process, ending up with incomplete data⁵¹. Overall, participants found that the questioning was too extractive, overwhelming and leaving insufficient time for discussion.
50. Participatory and survey methods were used simultaneously, without a pause to take stock and scrutinize their respective quantification/qualification requirements. Resource/time limitations meant that data was collected during a single two-week period, limiting the ability to undertake interim analyses and undertake gap-filling follow-up fieldwork. This likely has contributed substantially to the tensions between data and process (cf. Para 49), between 'truth' and multiplicity (cf. Para 48) that hampered the data collation process. A simplification of the methodology is required to make PIALA feasible under 'real world' circumstances where researchers' competencies or profiles don't always perfectly fit the required conditions. This could be done by sequencing participatory and survey methods (instead of using them simultaneously) and making them more lithe, and by providing time for data processing between the two.

4.4 Understanding causes and synthesising findings

How and to what extent did PIALA help analyse the data and produce and present the evidence necessary for confidently answering the assessment questions?

⁵⁰ Triangulation is a principle technique of social science that involves the use of more than one type of information or data source, method and researcher for crosschecking in order to obtain greater credibility of and confidence in findings. It builds on the premise that social research using a single source or method, and even a single theory and researcher, mostly suffer from biases inherent to the researchers' and respondents' limited worldviews and other methodological weaknesses or limitations.

⁵¹ It happened in 2 or 3 occasions where the focus group exercises were left unfinished, and missing data had to be retrieved from additional interviews.

A quite successful model was developed for participatory sensemaking, and guidance was developed for data collation and participatory sensemaking, focused on the assessment questions. However, the researchers experienced great difficulty with processing all data and analysing causal links across the many levels and domains of change, due to their insufficient understanding of the project and its ToC and their limited experience with multi-causal analysis and mixed methods evaluation. As a result, they didn't arrive at systematic data transcription and the data collation tables produced for the participatory sensemaking workshops showed insufficient integration of qualitative and quantitative information, multi-causal discussion and accuracy in triangulating sources and indicating strength of evidence.

The sensemaking model engaged project beneficiaries, village leaders, district and provincial officials, and project staff and management in interactive sessions for validating and analysing the evidence alongside the ToC as a guiding structure, while testing its causal links and assumptions. The sessions were attended with considerable enthusiasm and positive energy, in particular by the villagers who felt their views and judgements were heard. Project management and implementing partners expressed great interest to adopt the sensemaking model together with the ToC approach. Improvements can be made however to the small group compositions and interactions to enable more honest and critical reflection and debate; the formulation of the statements presenting the evidence; the process facilitation of the debates around project achievements and unintended impacts; and the quality of recording/note-taking and translation of group discussions.

Processes/Methods applied

51. Guidance was developed for initial data collation from fieldwork *per village* in preparation of the village feedback workshops. The purpose of this was: (a) to present the data collected from focus group discussions, key informant interviews, household surveys and secondary sources in a way that would enable local stakeholders *to assess significant changes* in their village and commune; and (b) to present the evidence emerging from the village collations in a way that would enable stakeholders *to assess project contributions to these changes*.
52. This guidance was further developed into data collation sheets for fine-grained transcription and organisation of data and sources *per district* around positive and negative changes indicated by the data at the level of the villages (both focus⁵² and non-focus⁵³), the communes and the district. This collation was needed for each of the five impact domains in the ToC and related indicators (see Annex X). The guidance and template focused on describing a specific 'change' in terms of:
 - when it took place and groups affected by the change;
 - causes of and project contribution to/influence on (direct and indirect) to the change;
 - positive and negative impacts and risks directly or indirectly (or not) affecting intended target groups (i.e. poor and near poor rural households, particularly women and minorities); and finally
 - relative strength of evidence of the changes and project contributions (incl. a score and explanation of why the evidence is found "very strong", "relatively strong", "relatively weak" or "very weak").
53. A model and guidance was developed for organising and facilitating an initial village feedback workshop of 3-4 hours at the end of each week's in-depth village inquiry. The model builds off four design principles –namely: creating space for equal and meaningful participation; focusing on changes broader than intended project results; prompting honest critique and debate between different stakeholder perspectives on these changes; and using statements as a mechanism for presenting the evidence of causes and impacts. Village feedback workshops were organised in all the villages where in-depth research had been conducted, creating the opportunity for the villagers and their leaders who participated in the research to:
 - validate the evidence;
 - provide additional information where needed;
 - critically reflect and collectively look forward to address enduring issues of poverty and exclusion; and
 - critically reflect on the added value of PIALA for helping understand and plan change processes.

⁵² Focus villages are geographically targeted by DBRP village and commune-level interventions.

⁵³ Non-focus villages are not geographically targeted but exposed to the DBRP by its interventions at district and provincial levels.

54. Building off the village feedback model, a concept for participatory sensemaking was further developed that formed the basis for organising a 2 days sensemaking workshop, adding four more design principles to the ones of the village feedback model –namely: building the analytical chain of actual changes; mirroring the actual with the envisioned changes in the ToC (thus probing its causal links and assumptions); developing an aggregated perspective despite diversity; developing a replicable model (in terms of flow and design principles) for participatory sensemaking. Village/commune, district, provincial and national level stakeholders (more than 100 participants in total) were brought together in the workshop to:
- validate and analyse the evidence related to the five impact domains in the ToC;
 - probe and validate the causal links and assumptions in the ToC by comparing *theory* with *evidence*;
 - discuss and value the relative project contributions;
 - critically reflect on whether the assessment was worth people’s time and efforts, and if/where the methodology could be improved to produce rigorous evidence, include people who matter, and make it locally implementable.

Critical Reflections

55. The village feedback workshops were generally fairly successful⁵⁴ and generated considerable enthusiasm and positive energy when participants realised they were given the opportunity to look at all the evidence that the researchers had gathered and discuss and analyse it together. People loved the idea that the research was also theirs. It gave them a sense of importance: a voice to express what matters to them, and a feel of confidence in their views and judgements. Lively debates were generated between project beneficiaries, village leaders and district-level officials about the changes observed in the villages, centred around the experiences of poor and near-poor women and men, what had caused these changes and how it had affected their jobs, livelihoods and family lives. The model worked very well to enable all the participants to engage in the discussions and reflections critically and free from pressure. However, we may assume that it is more likely for this model to contribute to the experience of 'empowerment' in a context where participatory development planning is piloted and expanded by the government as part of a democratic and economic opening process, such as is the case in Vietnam.
56. For validating the evidence, participants were first divided in small homogenous groups of beneficiaries, village/commune leaders and district officials, before bringing them together and letting them debate the differences in their findings. The debate was facilitated in a way that centred around the views of the beneficiaries about how they felt the evidence reflected the changes they experienced and the impacts these had on their lives and families. The evidence was presented to the participants in the form of statements about the changes and their causes, inviting the participants to discuss the causal links. Statements could either positively or negatively affirming or challenging the available evidence. This provoked plenty of disagreement fuelling more critical reflection and debate⁵⁵.
57. Also the provincial sensemaking workshop was quite successful, albeit preparation of content happened very last minute due to late submission of collation reports. The workshop had a well thought out design and flow, and generated plenty of good discussions and reflections with a vibrant enthusiasm. It enabled project beneficiaries and their village/commune leaders⁵⁶, district-level officials and project staff, and provincial line agencies and project management to engage in an interactive process of interpreting the evidence and comparing it with the envisioned or expected changes in relation to the project’s ToC. By using the ToC diagram as a guiding structure, participants obtained a broader systemic perspective of change processes that are impacting poverty, which enabled them to look at evidence of changes and project influences in relation to broader trends and change patterns from a more external-critical point of view. This was done in a very powerful way by first letting participants discuss and validate the evidence in small groups organised around the ToC’s five impact

⁵⁴ In some villages it worked better than in others, depending on the experiences people had with participation and the level of trust they had in the project.

⁵⁵ Ideally, if times permits, evidence of important causal links needing more critical analysis could be presented in two or three statements that are conflicting or juxtaposing different explanations.

⁵⁶ These were poor people elected from the group discussions and village feedback workshops, and their respective village and commune leaders (incl. from women and farmer unions), counting for about 30% of the workshop participants.

domains, and then plugging the validated evidence into the ToC diagram projected on a big screen as the groups were presenting the results from their discussions.

58. The models and concepts applied in the village feedback and sensemaking workshops form a good basis for further developing and testing a flexible and replicable model of participatory sensemaking. The IFAD partners in Vietnam expressed great interest to adopt such model and (together with the ToC approach) integrate it with their existing planning and M&E processes. Improvements can be made in the small group compositions for the different sessions, using the concept of “patches & nodes”⁵⁷ for creating equal space, avoiding power dynamics and enabling more honest and critical reflection and debate. The patches are the home bases or the groups were participants are of the same group, thus familiar and at equal terms with each other, hence comfortable to freely express their views. The nodes on the other hand reflect participants’ interactions with members of other groups who are differently positioned in the debate. To create safe and equal space for participation and avoid more powerful groups or stakeholder representatives dominate, it is important to let participants first reflect and develop their thoughts and ideas in their own “patches”, before discussing them with other groups in the “nodes”. This concept was used successfully in the village feedback workshops and can be expanded to all phases of the participatory sensemaking process. Another important improvement that also was tried out in the village feedback workshops, is would be to systematically centre the debates in the “nodes” around the views and experiences of the intended target groups. This would enable a more rigorous and critical mirroring of validated evidence of actual change with the ToC.
59. The statements presenting the evidence at the sensemaking workshop were generally too vague and missing power for triggering discussion. Consequently, the analysis of the statements at the workshop lacked critical depth. This was largely due to the lack of accuracy and logic in the collation tables, which were submitted too late to leave sufficient time for a proper review⁵⁸. Moreover, it turned out that in the sensemaking workshop, when evidence of actual changes were compared with envisioned or expected changes, only the overlapping areas were identified and the divergent ones ignored. Although the facilitation guidelines clearly indicated group discussions of both the overlapping and the divergent area as the basis for critical debate, this didn’t take place and were left unnoticed, due to language constraints and fatigue.
60. The researchers experienced great difficulty with bridging the many levels and domains of change and identify/analyse the many causal links, due to insufficient understanding of the project and its ToC and insufficient experience with multi-causal analysis using mixed methods and triangulation of multiple sources. The lead researchers felt that the guidance provided for data collation was insufficient, and that more detailed guidance would have given them more structure for their line of questioning and observations. This signals a tension between more open-ended *explorative* inquiry to identify and probe unknown causes/explanations and closed *confirmative* inquiry focused on expected change pathways. A better understanding of how to use the ToC as a guiding structure for data collection and its collation should help researchers better focus and quantify crucial data, while creating room for wider probing of plausible change pathways and explanations. Furthermore, requiring the researchers to obtain sufficient secondary information on project activities in each researched village prior to any primary data collection would likely have helped enhance the accuracy and geographic specificity in their aggregated analysis of project contribution. A list was provided for the secondary data collection, but this was done slightly too late. Also explicit guidance around secondary for data collation might have been helpful.
61. Although data were collected in gender-differentiated groups and on gender-specific questions related to the WEIA⁵⁹ permitting gender-disaggregated data collation and analysis, no gender-differentiated analysis took place due to limited experience of the researchers and insufficient guidance from the CDT. The guidance for data collation asked for a specification of those affected by the changes

⁵⁷ See: <http://patchesandnodes.com>; Kauffman, 1995.

⁵⁸ The collation tables were submitted four weeks after finishing the fieldwork and just one day before the provincial sensemaking workshop. They needed to be reviewed by the CDT to improve quality and determine the relative strength of evidence. Statements had to be formulated, around strong evidence for validation, and around insufficient or contradicting evidence for further probing during the workshop. Due to the researchers’ late submission of the collation tables the day before the workshop, this was done over night just before the workshop.

⁵⁹ See Annexes III, VIII & XI.

observed from the data, but not in a gender-differentiated way, thus inadequately guiding the researchers. This is an important improvement to be made to the guidance for the next Ghana pilot.

4.5 Assessing the outcomes of the Vietnam pilot impact assessment

To what extent did the Vietnam pilot produce gender-sensitive qualitative and quantitative evidence that is validated, transparent and robust enough to withstand scrutiny and enable key stakeholders to influence decisions towards sustainable impact? To what extent did the Vietnam pilot create and demonstrate an added-value-for-money for strengthening IFAD's self-evaluation system?

The evidence generated by the Vietnam pilot was extensively validated through the sensemaking that involved hundreds of people at village, commune, district and provincial levels. A process and guidance for aggregated analysis and a quite innovative outline structure for reporting was developed that permit drawing conclusions with sufficient confidence about the cascading influences of the project on the complex multi-level and multi-causal change processes that took place in the context of DBRP. A substantial amount of material (guidance, models, templates, and critical notes) has been developed by this pilot to enable methodological standardization while enhancing rigour, utility and feasibility. Compared to other rigorous impact evaluations, the cost of the actual impact assessment in this pilot (ca. USD \$ 90,000) is reasonably low, in particular when also considering its various outcomes for learning, reporting and advocacy. By taking a theory-based mixed methods approach, with components that fit IFAD's results- & impact-oriented measurement and management approach and its system for project quality assurance and enhancement⁶⁰, this pilot convincingly demonstrates the actual and potential added-value-for-money of PIALA for strengthening IFAD's self-evaluation system. To further increase this value, it is recommendable to integrate its components with project design and M&E. Important improvements remain to be made, however, in documentation and translation of raw and interim data as to ensure maximum transparency and uptake, and in data collation and analysis to ensure sufficient and accurate quant-qual integration and gender differentiation. This is crucial for obtaining sufficient granularity of conclusions needed to answer the question "*what works for whom under which conditions and why*" with greater confidence. Insufficient research capacity necessitated the CDT to hire an additional Vietnamese research and gender expert for this, and substantially hindered and delayed the final reporting. Hence it can be concluded that sufficient research capacity is essential for an efficient, high quality and cost-effective application of PIALA.

Processes/Methods applied

62. Evidence drawn from in-depth participatory research in eight villages was validated with village, commune and district level stakeholders in village feedback workshops, in which a total of about 200 people participated (see Para's 54 & 56). Aggregated evidence drawn from the entire mixed methods research was validated in a provincial workshop in which 100 people participated from the village, commune, district and province level stakeholders (see Para's 55 & 58).
63. Following the provincial sensemaking workshop, a process of aggregated analysis was agreed with the local research team and an outline with guidance (see Annex XI) developed for the final reporting. The lead researchers of the local research team produced an aggregated analysis of the evidence of key changes, relevant causes, and extent of project contribution compared to contribution claims, for each of the five impact domains in the ToC⁶¹. The collation tables produced for presenting the evidence in the feedback and sensemaking workshops, and the participants' comments and feedback from these workshops, formed the primary source for this aggregated analysis. The table with impact domains, indicators and target groups initially developed for the data collation (see Annex X), together with the ToC diagram (see Annex V), served as the guiding tools. Based on this aggregated analysis, the CDT produced a final report that draws conclusions about the cascading effects of the project through building "institutional capacity" affecting "voice & relationships", "access" and "livelihoods/jobs", hence leading to impact on "wealth/wellbeing".

⁶⁰ Cf. IFAD, 2009, 2011a, 2011c, 2012.

⁶¹ These are: "enhanced institutional capacity" + "enhanced voice & relationships" → "enhanced access to training/services/credit" → "enhanced livelihood and job opportunities" → "enhanced wealth and wellbeing".

64. A first draft of the final report was sent to the IFAD's Country Program Manager (CPM) and the Project Management Units (PMU) for feedback/comments. A rapid and elaborate response was delivered, including a request for additional secondary data (incl. from project M&E) to support the contentions for which insufficient convincing evidence was available for changing the impact assessment's findings and conclusions. All comments have been addressed in the final report, either by making adjustments in case of sufficient doubt or evidence; or by including the comments in a management response attached to the final report where no agreement was reached. Finally, both the impact assessment report and this methodological reflections report are presented to the DSG and ERG for feedback and comments (see Para 3), which will be addressed in appropriate ways and described in this Para in due course.
65. As illustrated by many of the methodological reflections in this report, the processes and methods of the PIALA pilot in Vietnam have been designed with the purpose of standardisation while creating maximum value for money in terms of knowledge creation for reporting, learning, innovation and adaptive management. The total cost of the PIALA pilot in Vietnam was about USD \$ 190,000, of which USD \$ 100,000 was spent on the management, methodological design and reflections, while the actual impact assessment (incl. in-country design, research and reporting) has taken about USD \$ 90,000. This compares very favourably with other budgets for independent rigorous impact evaluation (particularly those using mixed methods and serving multiple purposes), which often run into several hundred thousand up to one or two million American dollars.⁶²
66. Finally, to ensure maximum alignment with IFAD's self-evaluation system and complementarity to other methodologies, the CDT has regularly checked in with IFAD's responsible SSD and PMD staff to learn about IFAD's reporting and learning needs and other methodologies being used and pilot-tested as the basis for making adjustments to the PIALA methods and processes. Most happened in the formal meetings with the DSG and ongoing informal meetings with individual staff as part of the methodological reflections at the global level (see Para 3 and Footnote 2). Further alignment and complementarity is sought through a more active and direct involvement of SSD's methods experts in preparation of the second pilot in Ghana. It is hoped by the CDT that a more close collaboration with SSD can help improve the quantitative data collection and analysis aspects of PIALA's mixed methods approach for enhancing the validity of its quantified conclusions.

Critical Reflections

67. The reporting structure that was developed for this pilot is quite innovative, and can be further standardized and used for all impact evaluations using PIALA (or any other mixed-methods, multi-causal and multi-level theory-based approach). What is fairly new compared to existing outlines or structures of impact evaluation reports is the way in which relative project influences/contributions are identified in the 5 impact domains of the ToC. This is done by following the presumed change pathway in the ToC *backwards* from rural poverty impact to enabling institutions, unfolding the evidence of cascading changes towards its causes and the project's relative contributions, and allowing the reader to grasp the complexity of the systemic interactions and the assumptions underneath the causal links at every level.
68. Given what was said in Para's 36, 37, 56, 58 and 63, it can be concluded that, notwithstanding some important improvements to be made for the next pilot in Ghana, the PIALA pilot in Vietnam has produced evidence that is sufficiently validated and conclusions that are robust enough to withstand scrutiny. Arguably, the methodological limitations (Para's 30, 39-41 and 47) and the research capacity constraints (Para's 48, 49, 50 and 61) encountered in this pilot, cannot invalidate the findings and conclusions that were drawn from all the different sources at the multiple levels, involving several hundreds of people's inputs and validations.

⁶² It would be interesting to know, for instance, the range of budgets for impact evaluations considered rigorous by the 3ie.

69. However, although requested by the CDT, most raw and secondary data was not translated and some of it remained unused due to limited capacity and resources. There is an unfortunate price to that, namely: a lack of transparency in data collation. With more research capacity and budget for data translation and processing, this could have been avoided from the start. Translation is crucial yet often overlooked. It ensures that any authorised reader or user of the evidence can go back to raw data and initial transcription and analysis tables.
70. Also the sensemaking workshops were not documented in formal reports, which the CDT had not asked for, as to not spend additional time on interim products. However, both in its guidance and the technical backstopping, the CDT did require the researchers to take notes and include the outcomes of these processes in the aggregated analysis. Good notetaking increases the value of the feedback provided in the workshops as additional perceptual data for the final analysis and reporting. The quality of the notetaking of group discussions in the sensemaking workshops though was generally low. Hence the collation tables and the final aggregated analysis made rare reference to the workshops. Outsourcing the notetaking and the translation could be helpful, yet may bring outsider bias and thus would require review by the researchers. The best solution here is probably to have a good division of roles and guidelines for managing the sensemaking processes and the notetaking and translations of the discussions during these processes. In any case, the Ghana pilot will require interim reports of any of these critical events.
71. Although the researchers worked hard to improve the logic and quality of their aggregated analysis, their final report displayed the same quality issues as the data collation tables, namely: delayed delivery, insufficient quant-qual data integration, absence of gender differentiation, limited accurate multi-causal analysis⁶³, and limited indication of strength of evidence for specific assertions of change/causality. A greater degree of granularity of conclusions from their aggregated analysis would have helped answer the question what works for whom under which conditions and why with greater confidence. Together with the inaccessibility of much of the raw and secondary data, it was quite challenging and time consuming for the CDT to produce a final report of sufficient quality. To help the CDT to address these quality issues, a Hanoi-based professor in research methods and gender analysis was hired for reprocessing all raw data (both quantitative and qualitative, as well as secondary data, most of which were in Vietnamese) and produce a robust gender analysis. This significantly delayed the reporting process but greatly enhanced the robustness of the final report.
72. The research capacity and language & quality issues caused a five months delay in the final reporting process, which may have led to some missing data and information⁶⁴. However, the time lag in the reporting does not imply any difference in the validity of the conclusions, since the impact issues looked at have a slow rate of incubation and change, and thus presumably will not shift in five months' time. Furthermore, compared to mainstream impact assessments that measure T and T+1 with time lags of up to 5 years, this impact assessment can be considered a real-time evaluation.
73. In an ideal PIALA scenario, the final reporting would be the responsibility of the research coordinator who is leading the local research team, supervising the data collection & collation, and coordinating the ToC articulation and participatory sensemaking processes. Given this is the first PIALA pilot however, the members of the CDT divided these roles. Although the director of Depocen took on the role of research coordinator, he remained largely absent once the field research started, affecting leadership and outputs. A key lesson for the Ghana pilot is to ensure more active involvement and greater ownership by the in-country research coordinator.
74. Important moments of critical engagement and reflection have been created in this first PIALA pilot, contributing to empowerment (cf. Para's 28 and 56). Evidence of critical issues and constraints of the project logic have been presented to the stakeholders, contributing to their enhanced understanding of systemic change mechanisms affecting rural poverty (cf. Para 58). From the comments and reflections of project management and stakeholders (including intended beneficiaries) though it became apparent that a one-time-off impact assessment using PIALA is not sufficient to ensure that the evidence is useful and used for decision-making and improvements in the next project. For this an evaluative

⁶³ Incl. referencing and geographic specificity related to focus vs. non-focus villages.

⁶⁴ Although requested, no information has been obtained from the PMU about what data has been collected in the interim by the project that affect the findings of the PIALA impact assessment report.

learning culture is needed that promotes evidence-based learning, innovation and policy-making.⁶⁵ PIALA should therefore best be embedded in a monitoring, evaluation and learning system that is integrated with program design and management, enabling project implementers and intended beneficiaries to participate in data collection and analysis on a regular basis and learn to critically discuss and use evidence for developing and testing new mechanisms leading to higher and more sustainable impacts. In the context of DBRP for instance, a critical use of the evidence produced by the PIALA pilot could imply a decision to conduct an action research around targeting mechanisms in the new AMD⁶⁶ project (Adaptation to Climate Change in the Mekong Delta), in order to bridge what is signalled in the DBRP as a growing inequality gap.

Conclusion

Summary of positive outcomes and required adaptations

75. From the PIALA reflections, it appeared that most appreciated as positive outcomes of the first pilot in Vietnam are:
- Quality Assurance Framework for assessing the pilot's *rigour, utility and feasibility*
 - the impact analytical framing incorporating rural poverty, sustainable livelihood and women's empowerment indicators, and what was considered as a proper and useful set of impact assessment questions;
 - the Theory of Change (ToC) approach that convincingly demonstrated its potential for multi-causal analysis and was found very powerful for involving different stakeholders in this analysis;
 - a critical approach to sampling addressing variability issues and selection bias, and an adaptable yet standardized set of methods, which together forms a strong basis for testing improvements to the approach in the next pilot;
 - the use of mixed methods plus substantial guidance notes for the various methods and processes that were crucial for guiding the researchers in the data collection and collation;
 - an innovative model for participatory sensemaking at various levels that (together with the ToC approach) raised a strong interest among partners and country management for adoption in other IFAD projects and government programs; and
 - a reporting outline and guidance note that help synthesize and present findings and conclusions for each impact domain in a way that is accessible, traceable, convincing and useful for its users.
76. For ensuring PIALA will be sufficiently rigorous, useful and feasible/replicable for piloting at a larger scale, essential improvements have to be made and tested in the next pilot in Ghana based on the lessons learned from the Vietnam pilot, regarding:
- the provision of adequate guidance, and a well-defined contractual procedure (with clear deliverables) for ToC articulation and evaluability assessment;
 - the provision of adequate guidance for selecting standards and weights to operationalize rural poverty and other key indicators in the impact analytical framework in a context-relevant manner while allowing for aggregation to report on IFAD9;
 - sequencing of secondary data collection, household surveys and participatory methods, and provision of sufficient time for data processing, to enhance quality and rigour of data collection processes and outcomes, and enable better crosschecking and multi-causal analysis;
 - the provision of adequate guidance for using and adapting PIALA methods and processes in difficult and politicized contexts, while balancing quality standards and avoiding internal project and external researcher biases;

⁶⁵ Some comments of project management on the first draft impact assessment report for instance revealed a certain defensiveness, echoing an anxiety that was also observed during the design and conduct of the pilot and during the validation and sense-making workshops. Although the CDT had explained and repeatedly stressed that PIALA is *not* a methodology for performance evaluation, the evaluation seemed to be experienced as a judgement about performance. The CDT experienced solid professionalism and commitment on the part of the PMU in Ben Tre but commenting on this aspect is not part of PIALA.

⁶⁶ E.g. as a part of the monitoring, evaluation and learning activities in Component 2 of the AMD project (i.e. investing in climate change adaptation for vulnerable rural communities). Annex 2 on poverty, targeting and gender in the project design document provides a clear argument for this.

- the provision of clear guidance on formulating change statements presenting the evidence to the stakeholders for participatory sensemaking, and on group composition and facilitation for enabling honest and critical debate and reflection;
 - contractual requirements regarding the translation & submission of raw and secondary data, and the documentation of the sensemaking processes and group discussions;
 - the provision of adequate guidance (incl. quality standards) on multi-causal analysis, data collation and final reporting using the ToC as a guiding structure;
 - required research competences and the active involvement of the in-country research coordinator.
77. A thorough desk review should form the starting point of the ToC articulation process. Involving national and intermediate level stakeholders in this process should help understand relative project distribution and incubation and thus its evaluability. Based on this, the focus of the impact assessment, the specific questions/indicators, and an appropriate sampling hierarchy and stratification for assessing these indicators in relation to the ToC, can be determined. To evaluate project contribution relative to global poverty indicators, these indicators need to be operationalized in a way that is relevant to the project context, which could be done by selecting a set of descriptive standards reflecting key stakeholders' understanding of these indicators.
78. It is important that those who lead the research, and are responsible for the analysis and reporting, also lead this process of ToC articulation, framing and sampling (which constitutes the in-country design phase of the impact assessment). The leading research team needs to internalise both the project and the scope and methodological approach before starting with any data collection. Enough time should be provided for the lead researchers to do this properly. A key deliverable of the in-country design (conditional to first payment) should be a revised proposal presenting the outcomes and products (incl. ToC diagram, rationale for the impact assessment, indicators/questions, standards and sampling structure).
79. The following two criteria should be considered as non-negotiable for selecting a local research organisation:
- *Experience with participatory research:*
While the composition of the research teams should adequately reflect the competences needed to conduct a mixed qual-quant impact evaluation, crucial is also that the lead researchers have sufficient experience with participatory research. The PIALA research strategy indicates 3-5 years of field experience in participatory research⁶⁷. Unfortunately the ideal profile is not always available in 'real world' contexts, which makes it necessary to also think along the lines of further simplifying the methodology.
 - *Understanding of causal analysis and experience with impact assessment:*
A basic understanding of causal analysis beyond statistical correlation is indispensable for complexity-sensitive theory-based impact assessment using mixed methods. This also implies an understanding of the principle of triangulation as a means for probing causes/explanations and eliminating bias/influence. Such basic understanding is indispensable on the part of the lead researchers. The research coordinator moreover should also have proven knowledge of and experience with impact assessment in the mixed methods tradition.
80. Proper guidance is needed for the lead researchers and research coordinator on balancing different quality standards and requirements in politicized evaluation contexts. Basic training and guidance was provided on standards of rigor and utility essential for PIALA piloting. Based on the experiences with the first pilot, this can now be refined and expanded with concrete examples and exercises that reflect 'real world' conditions under which PIALA will be used. These conditions often involve highly politicized evaluation contexts that require researchers to acknowledge and properly deal with value judgments and power issues pervading the evaluation decisions (e.g. regarding 'independence' versus 'stakeholder engagement').
81. Proper guidance for data collation and analysis is also crucial to ensure sufficient and accurate quantal integration and gender & generation differentiation. The generic and specific research questions,

⁶⁷ See: IFAD, 2013: p.25. From the proposals received from local research institutes in the pre-selection, it appeared that none responded to this criterion. Hence the one most fitting and showing greatest ability to adapt and develop the required competencies was selected. Based on the lessons drawn from the Vietnam pilot however, it is clear that this criterion is non-negotiable.

and the tables and guidance for interim and final data collation, need to be gender/generation-differentiated and specify where/how quantitative and qualitative data need to be linked. In the next pilot in Ghana, decisions will have to be made about how to best mix and link quantitative and qualitative *data collection* methods (i.e. sequential, concurrent, or conversational), and how to best link and integrate the *analysis* of quantitative and qualitative data.⁶⁸ Finally, the CDT will search for useful software to help researchers triangulate and integrate secondary and primary quantitative and qualitative data, and ensure gender/generation-differentiated insights. This also serves to systematize mixed-data analysis necessary for mixed-methods impact evaluation. Researchers will need time though to familiarize themselves with the software and what it requires in terms of data input prior to any fieldwork.

82. Further improvements can be made to the participatory sensemaking model in terms of the different group compositions for different feedback-validation-analysis sessions at the different levels, the formulation of statements for presenting the evidence of changes and causal links, and the process facilitation of critical debates around deviations from expected changes and the project influences on these changes. A contractual requirement should be for the researchers to properly document and produce interim reports in English of the sensemaking processes and group discussions, which will require more time and budget. However this is essential to enhance the validity and credibility of the evidence produced and obtain sufficient granularity of conclusions from the sensemaking and contribution analysis.
83. PIALA as a mixed method approach aims to elicit representative quantified evidence on household poverty impact through a small number of impact indicators, while generating robust qualitative analysis of causal impact at community level. The sample size should therefore permit data collection on rural poverty indicators that are generalizable, and community-level causal analysis that can be confidently stated as representative of other project communities with similar interventions. In the Vietnam pilot, a limited sized survey was conducted on relevant impact indicators in a sample of 720 households in 18 focus- and 6 non-focus villages, which should be sufficient for establishing generalizability and comparing with the 900 household RIMS survey (albeit cluster-sampling more households per community and fewer communities overall). Certainly there is scope for increasing this sample size, although if so this should be linked to a revisiting of the sample size of the RIMS survey.

The participatory data collection on causal mechanisms was more limited in scope, taking place in a sub sample of the mini survey communities: 6 focus- and 2 non-focus villages. This was insufficient to cover project variability and allow for generalizability of causal inference. While there is a case for slightly increasing the number of communities engaged in in-depth participatory research in order to capture the variety of project contexts and mixes of interventions, and enhance the assessments' explanatory power, this is not a numbers game in the same way as the household survey sampling approach. Moreover, participatory data collection in a larger sample of communities is considerably more resource-intensive and at the risk of sacrificing depth of analysis for breadth of coverage.

84. Generally, it has proven to be very important for the quality of the assessment to provide sufficient time (and thus also budget) for the desk review and ToC articulation; the sequencing of secondary data review, household survey and in-depth participatory research; the data processing (incl. cleaning and translation of raw data), and for the final reporting. The bigger the sample, the more resources are needed for all this. Obviously if the entire mixed methods research has to be conducted in a country-representative sample, this will significantly increase the research cost.

Capacity & budget considerations for the next pilot in Ghana

85. Impact evaluation is a relatively new form of evaluation that has been embraced in international development in past 8 years, with a more widespread application since five years. Hence major capacity constraints exist in this area, particularly for innovative and rigorous approaches such as PIALA. Although collaboration with local researchers within the country where the impact assessment is conducted is favourable, there might not be sufficient capacity available to do so in a way that responds to the expected quality standards. In such case, it is suggested to look for a regionally-based

⁶⁸ See Bamberger, 2012: p.18.

qualified research or evaluation institute who works with local research assistants. Important is that the lead researchers are sufficiently qualified and are given enough resources to properly train the local research assistants. Pairing national teams with qualified international researchers or evaluators is an alternative option, contributing to in-country research capacity-building and the development of an impact knowledge base as a public good.

86. Given the demands to sampling for meeting statistical principles on the one hand, and those of facilitation for rigorous participatory data collection and sensemaking on the other, a tension remains between depth/quality and coverage/quantity. Both types of demands are resource-intensive in themselves, if to be done rigorously. Combining participation and statistical analysis in a nested mixed-methods design demands high coverage, enough time, and highly competent researchers, thus is naturally resource-intensive. However, all impact assessments that are rigorous have a serious cost. IFAD could promote good standard practice across its portfolio by determining and making public what it considers a reasonable indicative budget for impact assessment and how it seeks to make impact assessment more cost-effective by better embedding it in an impact-oriented measurement and learning system that is integrated with program design and M&E. Such an indicative budget for rigorous impact assessment needs to take into account:
- the total investments made in an intervention;
 - the value-for-money it should produce in terms of *learning* and *accountability* for both aid/service providers and receivers;
 - the quality requirements it should meet to deliver this value; and
 - the available research and absorption capacity in the countries, needed for meeting the requirements and promoting effective use/uptake.
87. Sufficient budget is needed for the next pilot in Ghana to address the lessons from the Vietnam pilot and make the improvements mentioned above, and to sequence and apply the mix of methods in an appropriate sample size and structure that permits adequate causal inference to serve the purposes or intended uses of the impact assessment. Hence there are three options for the design and testing of these improvements in the Ghana pilot:
- The first option is a *narrow-focused* PIALA version that looks at *one particular change mechanism* in the intervention's ToC, in a *country-representative stratified-random sample*. The purpose is learning about the effects of one particular aspect of a project. Findings are not sufficient though (and should be complemented with other studies looking at the remaining mechanisms) to report on the intervention's total contribution to rural poverty impact.
 - The second option consists of a *broad-focused* PIALA version that looks at the *full range of mechanisms* in the ToC, in a *stratified-random sample* in a *limited geographic area*. The purpose is to learn about the intervention's total contribution to rural poverty impact in this geographic area. Hence findings cannot be generalized and used to report on the intervention's contribution to rural poverty impact for the whole project area.
 - The third option finally is a *broad-focused* PIALA version that takes a systemic perspective and looks at the full range of mechanisms in the ToC, in a *country-representative random sample*. The purpose is learning about *and* reporting on the intervention's *total contribution* to rural poverty impact as a whole, and the specific mechanisms that were more (or less) effective in realising these contributions.
88. Finally, budget is also needed to finalize the knowledge products for dissemination and mainstreaming of this IMI-funded initiative. These include: the individual impact assessment reports on Vietnam and Ghana and the draft PIALA guidance package; a synthesis report on PIALA reflections and lessons from the two pilots; two debriefs and a seminar around the outcomes of the Improved Learning Initiative; and (if time and budget permits) a professional and academic publication to share the approach and the insights from the reflections with the wider development community.

ANNEXES

Annex I. Overview of prevalent Impact Assessment Approaches

Quoted from:

Van Hemelrijck, A. (2014). *Understanding 'Rigor': Challenges in Impact Evaluation of Transformational Development*. PhD Research Outline Paper (revised version). Institute for Development Studies (IDS). Pp. 19-26.

Impact assessment or impact evaluation generally refers to the collection of evidence of an intervention's outcomes and impacts or *causal claims*, and the value judgements that are made based on this. Evaluation differs from regular research in its explicit value judgements about an intervention as to inform planning, policy and funding. Impact however is understood in many different ways, delineating different approaches. Based on a review of contemporary academic and professional literatures, I broadly identify three types of definitions of impact and approaches of impact evaluation. These include: the technical-statistical approaches, the multi-linear theory-based approaches, and the complex realist approaches.

While useful to measure the direct linear effects of technical interventions or program components (such as vaccination in a health program) as to find out if it works (or not) within that particular (program) context, methods following the strict scientific-statistical procedure appear cumbersome and risk to run into predicaments of construct and conclusion validity, credibility, utility and transferability in more complicated or complex program contexts. Yet these statistical-scientific methods seem to be the best we have for accurately measuring an intervention's immediate effects towards broader and perhaps more sustainable systemic change. (Copestake, 2013) To assess such broader systemic change in more complicated program environments, theory-based approaches have expanded the methodological scope of impact evaluation by using Logframe-derived ToC models that visualize and describe the multiple causal strands and actors, and mixing methods that address the validity and utility concerns. However, still framed in a predominantly linear-systemic logic, also this type of approaches is found inadequate to assess more complex causation processes. (Funnell & Rogers, 2011; Nobuko, 2010; Vogel, 2012)

The realist approaches take more of a complexity lens to assess interventions' influences on complex systemic change processes by using ToC models that identify mechanisms and assess plausible outcomes of these mechanisms in relation to each other and to their contexts. The difference with the more linear ToC models is that, instead of merely focusing on planned activities and intended effects, these realist models are evolving constructs that serve to probe hypotheses and assumptions and discover patterns of broader systemic change. *Cross-validation* serves as the realist alternative for *counterfactual analysis* and as a means to deal with politics in complex evaluation contexts. The emphasis is on attaining explanatory power and building generalizable knowledge by systematically analysing and cross-validating multiple independent sources and lines of evidence (both primary and secondary) at multiple levels, and building on evaluations already conducted. Hence realist approaches are most promising for helping to build the knowledge base needed for donors to make appropriate funding and policy decisions that support (rather than reduce or even undermine) transformational development. Yet they are criticised for their lack of rigor in causal inference – since there are no such exact measurements as with scientific-statistical approaches; only trustworthy estimations that infinitely remain subject to verification. (Pawson, 2013; Sanderson, 2002) In return, many authors argue that impact and rigor are defined in too narrow statistical-scientific terms, and a broader notion of rigor or quality assurance is needed that relates to the entire evaluation process and evidence being produced, covering all quality issues, flaws and biases that may occur in mixed-methods evaluations of more complicated or complex interventions. (Chambers, 2008b; Patton, 2012; Stern et al., 2012; Westthorp, 2012)

Stern et al (2012) provide a definition of impact and impact assessment that builds on the DAC/OECD definition but synthesizes all three approaches in a way that creates room for tailored complexity-sensitive designs. In their definition, impact is considered as long-term changes in people's lives, but impact assessment is expected to look at contribution claims that can happen at any moment in time along the impact trajectory and explain *how* it happened for learning purposes. The notion of contribution (of which attribution is considered a special kind) rests on the recognition that most interventions are *neither sufficient nor necessary* in themselves to cause the desired impact, are usually accompanied by many other development processes, and affected by changes in the environment. Looking at *contributory cause* –this is: the kind of causal effect a particular intervention has contributed at a particular moment in the impact trajectory (e.g. triggering/systemic, ground-preparing or supportive)– implies an understanding of context,

actors and processes in the development context. (Stern et al., 2012) This, Stern et al and others argue, enables us to *tailor* the design of each evaluation to the degree of complexity with regard to: (a) the specific kind of impact it has to focus on, (b) the type of intervention (and causal claim) it has to investigate, (c) the questions it has to answer, and (d) the purposes it has to serve. (Funnell & Rogers, 2011; McGee, 2013; Rogers, 2009; Stern et al., 2012) The table below presents an overview of impact assessment approaches and designs that best fit rather *simple*, *complicated* or more *complex*⁶⁹ kinds of interventions, impacts, questions and purposes. Taking a complexity-sensitive approach makes it possible to think of *hybrid* or *nested designs* that *combine* elements of the three types of approaches summarised in this table. Doing so can help overcome limitations, serve multiple purposes/values, evaluate specific causal effects and at the same time also assess more complex systemic changes. (Forss et al, 2011; Stern et al., 2012)

Finally, a major concern expressed in the literature that deserves due attention in light of the trends and debates around transformational development and aid effectiveness, is about the focus on merely *scientific* instead of *practical-communicative* knowledge and *instrumental* instead of *ethical-moral* rationality, which undermines the capacity of development actors to adapt and respond responsibly to complex change situations. These two kinds of knowledge and rationality need to be better balanced in ways that support transformative action, reflection and decision-making, for which critical engagement of stakeholders in development practice and evaluation is essential. (Bawden, 2010a, 2010b; Sanderson, 2002) Although theory-based and realist evaluation designs may have participatory methods in their mixed-methods basket, there is little to find in the literature about how to deal with knowledge-power issues that affect rigor in causal analysis. (Chen et al., 2011; White & Phillips, 2012) The literature on transformative research and participatory statistics provide useful suggestions for joining forces and defining rigor in ways that address both scientific and participatory knowledge concerns. (Chambers, 2003, 2008b; Holland, 2013; Levy & Barahona, 2002; Mertens, 2009) The IFAD and BMGF-funded Improved Learning Initiative seeks to contribute to this.

⁶⁹ This follows Rogers' *simple-complicated-complex* typology for determining an appropriate impact evaluation design. (Funnell & Rogers, 2011; Rogers, 2009) Other similar models are: David Snowden's Cynefin framework that makes a distinction between *simple*, *complicated*, *complex* and *chaotic* problems and contexts, requiring respectively a *procedural*, *expert-led*, *horizontal-collaborative* and *emergency* response (Snowden, 2002, 2007); and Barder's recently posted *simple-complex* cubic that helps determine the complexity degree of any specific project or program – "*ranging from simple problems and interventions in stable contexts through to complex interventions in diverse and dynamic contexts*" – for deciding which type of managing-for-results approach is most appropriate. (Barder, 2012)

<i>Kind of:</i>	Elements	SIMPLE	COMPLICATED	COMPLEX
Intervention/ causal claim	Causality	Single linear causal strand, sufficient and necessary to produce the impact	Multiple linear and simultaneous causal strands, each sufficient to bring an effect but likely not necessary to bring about the intended impact	Multiple non-linear and recursively interacting causal strands, each unnecessary and insufficient for generating impact
	Context dependency	Cause-effect works in a particular context, and can be isolated from influences within that context	Cause-effects are context-dependent but certain configuration sets may work in and across certain contexts, and can be analysed and compared on consistency	Causal strands interact with context as an embedded system, with attenuating or amplifying feedback loops
	Implementing organisations	Single organisation	Multiple organisations in contractual relationships	Multiple agencies in developing and mostly non-contractual partner relationships and coalitions
Impact	Definition	Short-term effect attributable to a specific intervention = <ul style="list-style-type: none"> • ‘before-after’ difference in comparative ‘with’ and ‘without intervention’ cases • statistical inference of frequency and magnitude of cause-effect associations (statistical causality) 	Positive and negative, intended and unintended, short and longer term effects caused directly or indirectly by certain configuration sets of interventions in multiple cases/contexts (multiple causality)	Changes in patterns of interactive behaviour generated by certain mechanisms (generative causation or emergent change)
	Evidence	Statistical-scientific evidence of attributable effects	Convincing evidence of consistency in the effects from configuration sets	Convincing evidence of ‘mechanisms’ or interactive changes in a core set of indicators measuring complex systemic change, and estimations of plausible contributions to these changes
Question		What works? To what extent can the specific net-impact be attributed to the specific intervention?	What works for whom in what contexts? Has the configuration of intervention made a difference in the cases/contexts where it was implemented?	What works how and why in which contexts with which actors under which circumstances? How and why were the mechanisms created and activated in the particular context under investigation, and how could it possible work in different contexts?
Purpose	Intended uses	Accountability to/of international donors → funding decisions	Accountability to/of national governments and international donors → policy making and intervention planning by both donors and aid recipients	Accountability based on collective responsibility among the coalition partners (incl. donors, policy makers and implementing partners) at local, national and global levels → adaptive management; empowerment
	Polycymaking	Donor-driven	Expert-led	Facilitative - empowering

<i>Determining:</i>	Elements	SIMPLE	COMPLICATED	COMPLEX
impact assessment Design choices	Main approach	<ul style="list-style-type: none"> statistical approaches quasi-experimental designs and randomised control trials (RCTs) 	<ul style="list-style-type: none"> naturalistic-ethnographic approaches in case-based and network-based designs –e.g. QCA (Rihoux & Ragin, 2009); comparative and within-case studies (George & Bennett, 2005; Stake, 2005); systems dynamics/modelling (Midgley, 2007) theory-driven approaches involving process tracing and contribution analysis against a more linearly pre-planned Theory of Change (often Logframe-based) and using a mixed methods design that is mostly sequential and simultaneous (Coryn et al, 2011; White, 2009b) 	<ul style="list-style-type: none"> mixed methods designs (which include quantitative and qualitative methods, and may also incorporate quantified qualitative methods such as Sensemaker) nested in an overall realist approach, following a more generative process of Theory of Change articulation (Funnell & Rogers, 2011; http://www.sensemaker-suite.com/smsite/index.gsp) participatory approaches that build on ethical principles related to ‘voice’, ‘ownership’, ‘utility’ and ‘construct validity’ –incl.: Participatory Statistics (PS) (Holland, 2013); Comparative Constituent Voice (CCV) (Bonbright & Power, 2010); Most Significant Change (MSC) (Davies & Dart, 2007)
	Weaknesses / Strengths	<ul style="list-style-type: none"> + uncovering law through statistical analysis + eliminating bias through counterfactual - lack of explanation - weak construct or ‘concept’ validity - weak external validity or generalizability - weak truth validity or credibility and utility of findings for different stakeholders - weak understanding of context 	<ul style="list-style-type: none"> + discovering typologies or configuration sets applicable across contexts + understanding of more complicated development contexts and interventions - difficulty in interpreting highly complex interactions and combinations - doesn’t capture non-linear complex interactions and precludes unpredictable outcomes 	<ul style="list-style-type: none"> + in-depth understanding of complex development contexts + fine-grained explanation + brings power and politics back into the analysis, and therefore potentially empowering if using proper participatory processes - evidence based on estimations of contributions - risk of bias and loss of evaluator’s independence in the case of a participatory approach
	Challenges	<ul style="list-style-type: none"> matching of treatments and control groups in (quasi-)experimental designs isolation of causal relations from pre-empting causal processes ethical considerations regarding randomized sampling and exclusion of key stakeholders from knowledge-creation/use use and utility of evaluation findings for improving performance and impact intervention is heavily influenced by the specific method and data needs requiring scientific expertise → Cost 	<ul style="list-style-type: none"> attaining rigour in mixing of methods, processes and evidence while dealing with multiple partners reaching agreement among partners on evaluation design and conclusions enough data to look at all of the possible configuration sets and deal with ambiguity too many configuration sets or combinations in case of too many variables requiring technical-methodological expertise → Cost 	<ul style="list-style-type: none"> attaining rigour in methods, processes and evidence while dealing with power and politics in messy partnerships building competencies for understanding and explaining complexity among stakeholders (incl. donors) through creating a common language that most are unfamiliar with capturing the complex while keeping it relatively simple and manageable for creating ownership among key stakeholders (without reducing or ignoring complexity) requiring good facilitators and more time for good processes → Cost
	Requirements	<ul style="list-style-type: none"> 2 comparable cases for (quasi-)experimental design many diverse cases for statistical modelling and correlation sufficient large n samples for statistical analysis 	<ul style="list-style-type: none"> sufficient comparable cases of causal configurations ToC (preferably baselined) 	<ul style="list-style-type: none"> one case of an identifiable causal mechanism, with good access to multiple data sources

Annex II. PIALA Quality Assurance Framework

	RIGOUR = thoroughness, consistency and reliability of methods, processes and evidence	UTILITY = accessibility, credibility and value of methods, processes and evidence	FEASIBILITY = replicability, manageability and cost-effectiveness of methods and processes
<p>Defining impact</p> <p>How and to what extent does PIALA help adequately identify the changes and influences or contributions that the impact assessment should focus on?</p>	<p><i>Has the focus of the impact assessment been well enough defined against a well-articulated theory of change towards rural poverty impact?</i></p> <ul style="list-style-type: none"> • <i>Does the theory of change adequately visualise the change pathway that is presumed to lead to rural poverty impact?</i> • <i>Does it include most important causes and influences –both positive and negative, intended and non-intended?</i> • <i>Does it reveal most important change mechanisms and the intervention’s influences or contributions that the assessment should focus on?</i> 	<p><i>Does the theory of change reflect key (particularly primary) stakeholders’ views of rural poverty impact?</i></p> <p><i>Does the theory of change give room for both locally and globally relevant perspectives of the changes presumed necessary to realise rural poverty impact?</i></p> <p><i>Do stakeholders recognize and affirm the focus of the impact assessment on most pertinent changes and influences/contributions?</i></p>	<p><i>Is the process by which the theory of change has been articulated and the evaluation focus has been defined, replicable in a cost-effective manner?</i></p>
<p>Framing evaluation approach</p> <p>How and to what extent does PIALA help adequately determine the purposes, questions and standards that should guide the impact assessment?</p>	<p><i>Have the purposes and questions/indicators for conducting the impact assessment been clearly defined?</i></p> <p><i>Have criteria or standards for “valuing” the intervention’s influences on or contributions to rural poverty impact been clearly defined?</i></p>	<p><i>Do participants understand and affirm the evaluation purposes, questions/indicators and standards?</i></p>	<p><i>Is the way in which different paradigms (e.g. different definitions, purposes, questions and standards) are integrated into one approach manageable, replicable in a cost-effective manner?</i></p>
<p>Describing what happened</p> <p>How and to what extent does PIALA help rigorously collect, triangulate and collate the data needed to produce evidence for confidently answering the assessment questions?</p>	<p><i>Has an appropriate <u>sampling structure and size</u> been defined that permits adequate causal inference to serve the purposes or intended uses of the impact assessment?</i></p> <ul style="list-style-type: none"> • <i>Has it adequately taken into account contextual variability?</i> • <i>Has it adequately taken into account relative distribution and incubation of project contributions?</i> • <i>Have research and sample biases (e.g. the sort of bias that comes with limitations of research and sampling) adequately been avoided?</i> <p><i>Have appropriate methods been selected for <u>sufficient, reliable and coherent data collection</u> to describe observed changes and identify plausible explanations or causes?</i></p> <ul style="list-style-type: none"> • <i>Have possible causes of observed changes and impacts been exhaustively probed?</i> • <i>Have availability and perception biases (e.g. the sort of bias that comes with limitations of people’s knowledge and perceptions of possible explanations for impact or non-impact) adequately been avoided?</i> 	<p><i>Have methods and processes enabled key (particularly primary) stakeholders (both women and men) to equally and meaningfully engage in the collection of data and validation of findings?</i></p> <p><i>Have the methods been used and data been collected in a way that is transparent, gender/power-sensitive, and reflective upon participants’ feedback?</i></p> <p><i>Has data been collected in an ethical manner, respectful of participants’ right to prior consent, confidentiality and protection?</i></p> <p><i>Have data been <u>collated</u> in a way that is transparent, respectful of differences in views and perspectives, and relevant to the different key stakeholders?</i></p>	<p><i>Are the PIALA processes of sampling, data collection, causal analysis and validation replicable in a cost-effective manner?</i></p>

	<p><i>Have methods been <u>adequately sequenced</u> in a way that permits crosschecking and successive elimination of alternative explanations for the observed changes?</i></p> <p><i>Have methods been used <u>consistently across the sample</u> in a way that permits aggregated analysis?</i></p> <p><i>Have the methods been used and data been collected independently, thus free from political influence and organisational pressure?</i></p> <p><i>Did the researchers have full access to all stakeholders and information, and full autonomy in facilitating group discussions, conducting surveys, and processing data?</i></p>		
<p>Understanding causes and synthesising findings</p> <p>How and to what extent does PIALA help analyse the data and produce and present the evidence necessary for confidently answering the assessment questions?</p>	<p><i>Have all data been adequately analysed, and plausible links between observed changes, causes and impacts been identified and probed? Have weak links been eliminated and strong links identified?</i></p> <p><i>Have valid evidence been obtained from this causal analysis that confidently helps explain the intervention's influences/contributions to impact and answer the assessment questions?</i></p> <p><i>Have the intervention's influences or contributions to rural poverty impact been valued along agreed standards?</i></p> <p><i>Have useful projections about future impacts been made based on the explanation and valuation of intervention's influences/contributions?</i></p>	<p><i>Did key (particularly primary) stakeholders have the opportunity to meaningfully engage in the analysis of causes and the intervention's influences/contributions?</i></p> <p><i>Are the conclusions reached about the causes and contributions to the observed changes based on key stakeholders' critical reflection?</i></p> <p><i>Are these conclusions justified in the cultures and contexts where they have consequences?</i></p> <p><i>Has the process of analysing causes and valuing contributions enabled stakeholders to revise or deepen their understanding of the development processes and interventions in a way that helps them make useful projections about future impacts?</i></p> <p><i>Have they gained new knowledge and skills that they can use to influence future processes, decisions and relations affecting impact?</i></p>	<p><i>What conditions enabled or disabled a robust analysis of causes and contributions – and do these exist or can they be created in other contexts?</i></p> <p><i>What conditions enabled participation of stakeholders in sensemaking – and do these exist or can they be created in other contexts?</i></p>
<p>Assessing the outcomes of the pilot impact assessments</p> <p>To what extent did the pilot impact assessments produce gender-sensitive qualitative and quantitative evidence that is validated, transparent and robust enough to withstand scrutiny and enable key stakeholders to influence decisions towards sustainable impact?</p> <p>To what extent did the pilot impact assessments create and demonstrate an added-value-for-money for strengthening IFAD's self-evaluation system?</p>	<p><i>Gender-sensitive qualitative and quantitative evidence of impact on multidimensional poverty that incorporates perceptual data, and is supported by in-depth and gendered description.</i></p> <p><i>Sufficiently reliable reconstruction of before-after comparison (causal inference) without baseline data and building on multiple factual and counterfactual data and relevant and reliable secondary sources</i></p>	<p><i>Description of how different stakeholder groups were engaged, in the various processes of the impact assessment.</i></p> <p><i>Description of how data were presented back to different groups of key stakeholders, and how these groups were involved.</i></p> <p><i>Evidence from group reflections and surveys of stakeholders' increased awareness of systemic change affecting poverty and the potential use of evidence to influence decisions.</i></p>	<ul style="list-style-type: none"> • <i>PIALA's value-for-money compared to other approaches, or its cost-benefits for key stakeholders;</i> • <i>PIALA's relative user-friendliness and manageability according for country and project management;</i> • <i>PIALA's alignment with IFAD's self-evaluation system and complementarity to other M&E and impact assessment methodologies;</i> • <i>potential for standardisation that permits aggregated analysis.</i>

Annex III. PIALA Impact Analytical Framework

At the consultation workshop held in Rome on 17 October 2012, a discussion was held around the identification of key dimensions or indicators of impact that the PIALA pilots should focus on. Different frameworks and lists of impact-level indicators that are in use within IFAD were looked at⁷⁰, while seeking to agree on core converging and critical domains⁷¹.

Agreement was reached on the following:

- rural poverty should be characterized as multi-dimensional;
- rural poverty should be assessed not just by the RIMS indicators but also by poor women's and men's own perceptions;
- impact indicators and theories of change should be defined specific to the projects under evaluation;
- PIALA should be guided by an overall impact analytical framework that can link project-specific indicators/dimensions and theories of change to IFAD's Results & Impact Management System (RIMS);
- the PIALA impact analytical framework should include key dimensions related to rural poverty impact, enablers of impact, and processes/mechanisms triggering these enablers;
- PIALA should focus attention and resources foremost on those dimensions that are not covered well by other methods and processes in IFAD's self-evaluation system;
- PIALA should include a strong gender lens, showing if, how and why improvements in wellbeing and opportunities have occurred for women differently than for men.

The following analytical framework was provisionally agreed:

	Focus of PIALA	Analysis
Perceptions on rural poverty impact	<ol style="list-style-type: none"> 1. Nutrition and food security 2. Resilience 3. HH income and assets 4. Gender equality and women's empowerment (IFAD 9) 	<ol style="list-style-type: none"> 1. What has changed for whom and how? Benefitters; excluded; adverse effects. Gender lens/social lens.
Enablers (process outcomes)	<ul style="list-style-type: none"> • Institutions and policies • Voice / leadership for decision-making • Access to resources • Access to markets • Relationships • Capacity 	<ol style="list-style-type: none"> 2. How sustainable are these changes? 3. What are the consequences of these differentiated effects?

This was further refined in the PIALA research strategy⁷² into the following impact analytical framework that is indicative for the PIALA pilots:

⁷⁰ Including:

- IFAD's Strategic Framework 2011-2015 –in particular the goal, strategic objectives, outcomes and core principles of engagement of the Strategic Framework (IFAD, 2011a, p. 29)
- IFAD's Results Measurement Framework (RMF) that encompasses IFAD's self-evaluation system –in particular the impact dimensions to measure second level results or outcomes/impacts expected to be realized through IFAD-funded projects (IFAD, 2011b, pp. 5–6)
- IFAD's gender equality and women empowerment commitments in IFAD9 (IFAD, 2011a)
- the lists used by the initiatives of the M&E Harmonisation Group (cf. The M&E Harmonization Group of Food Security Partners, 2013)

⁷¹ The following shortlist was obtained from the discussion:

- food security and nutrition;
- income increase and resilience;
- natural resources and assets;
- rural productivity;
- access to services, value chains and markets;
- poor people's (cap)abilities, organizations and influence on policies and institutions;
- institutional capacities and policy frameworks enabling pro-poor agricultural and rural development;
- empowerment;
- gender equality and women's empowerment.

⁷² See: IFAD, 2013: p.14.

PIALA	IMPACT DIMENSIONS	WOMEN'S EMPOWERMENT
Rural poverty impact		
Changes in wellbeing	<ul style="list-style-type: none"> • HH income and assets • HH Food and nutrition security • # of people out of poverty 	<ul style="list-style-type: none"> • Cross-Cutting: Gender equality in impacts
Empowerment-centred impact enablers		
Asset-based agency	<ul style="list-style-type: none"> • Human capital • Physical capital • Natural capital • Financial capital • Social capital • Personal capital 	<ul style="list-style-type: none"> • Resources: Ownership of productive resources • Leadership: Membership in economic or social groups and comfort in speaking in public
Institutions	<ul style="list-style-type: none"> • Enabling institutions • Service delivery institutions 	<ul style="list-style-type: none"> • Rules and regulations responding to gender-specific needs and creating equal opportunities
Processes	<ul style="list-style-type: none"> • Social relations • Cultural norms and practices • Market transactions • Politics • Rights 	<ul style="list-style-type: none"> • access to and decision-making power over productive resources • sole or joint control over income and expenditures • leadership roles and influence • time allocation to productive and domestic tasks, and satisfaction with time for leisure activities
Vulnerability context	<ul style="list-style-type: none"> • Shocks • Trends • Cycles 	

Annex IV. PIA-DBRP Stakeholder Consultations

National level:

On Friday 14/6:

- 15:00 – 17:00 Interview with Nong Thu Hong Hanh, Head of Division Foreign Economic Relations Department, International Organizations and INGOs Division, Ministry of Planning and Investment (MPI)

On Tuesday 2/7:

- 15:00 – 17:00 Interview with Duong Quynh Le, Director Multilateral Division, Department of Debt Management and External Finance, Ministry of Finance (MOF)

On Wednesday 3/7:

- 9:00 – 11:00 Interview with Nguyen Viet Tuc, Deputy Director, Project Management Department, Vietnamese Bank for Agricultural and Rural Development (VBARD)
- 14:00 – 16:00 Interview with Vu Dang Toan, project coordinator, NTP for rural development, Ministry of Agricultural and Rural Development (MARD)

Provincial level:

On Tuesday 18/6:

- 09:00 – 12:00 Meeting with members of the PPSC, PPMU and DPMOs for sharing the headlines of PIALA and the DBRP project, and discussing a draft ToC (identifying gaps, causal links, areas of success, external influences)
- 13:30 – 17:00 Meeting with the DBRP M&E team⁷³ for selecting secondary data sources (incl. M&E data), identifying sampling criteria, and determining focus and scope of the impact assessment based on the outcomes of the morning session with the PPSC, PPMU and DPMOs.

On Monday 08/7:

- 07:30 – 11:00 Group interview with Ho Vinh Sang, President, and Nguyen Van Dac, Vice-President of the Coconut Association (CA); Nguyen Van Chien, Vice Chairman of the Farmers' Association (FA) (*members of the PPSC*)
- 13:00 – 15:00 Group interview with XXXXX, President and Vice-President of the Women's Union (WU), in charge of the CIG Dev Fund (*members of the PPSC*)
- 15:30 – 17:30 Group interview with Tran Van Chinh, Vice Head of Business Registration, Nguyen thi Giao Chi, Vice Head of Planning and Consolidation, and Nguyen Xuan Tho, technical staff member of Planning and Consolidation at the Department of Planning and Investment (DPI)

On Tuesday 09/7:

- 07:30 – 11:00 Group interview with Tran Thanh Tam, director of the Seeds Centre, Huyah Thi Truc, officer of the NTP-NRD, and Le Van Nam, accountant of the NTP-NRD, at the Department of Agriculture and Rural Development (DARD)
- 13:00 – 15:00 Group interview with Nguyen Thi Khiet, Head of Labour & Salary Division; Phan Thanh Tung, Head of Vocational Training Division; and Le Thanh Phuc, responsible for Social Protection; at the Department of Labour, Invalids and Social Affairs (DOLISA)
- 15:30 – 17:30 Interview with Vo Van Minh, Head of Credit Department at the provincial branch of the Vietnamese Bank for Rural and Agricultural Development (VBARD)

On Wednesday 10/7:

- 18:00 – 22:00 Dinner meeting with Mr. Tran Anh Tuan, Vice Chairman of Provincial People's Committee, and Head of the PPSC, and Mr. Son, DBRP project manager

⁷³ Including, provincial and district-level officers and managers, members of the PPMU and the DPMOs responsible for M&E.

Annex V. DBRP Theory of Change

The DBRP project is based on a ‘making markets work for the rural poor’ paradigm of development which requires (1) a solid market analysis to ensure there are markets or potential markets; (2) capability to supply the market (ranging from individual entrepreneurial skills to infrastructure, services and policies); and (3) processes that ensure flow-through of pro-poor benefits from direct and multiplier effects. In the DBRP in B n Tre context, locally viable and pro-poor short value chains are the focus of such business development. Better access to these markets through these value chains is expected to create more local business and job opportunities for the poor, hence enabling poor rural households to generate a more resilient income base for meeting basic needs and re-investing in income-generating activities. Such access requires enabling conditions at three levels. At commune level, rural poor households need training in possibly lucrative activities (such as making cages for fighting cocks and industrial sewing for working in garment factories), and government must be able to respond to priorities of the poor emerging from village plans and the need for training, information and credit for starting up small businesses or investing in livestock (such as a cow or a goat). Furthermore, investments are needed in local infrastructure to boost local trade and investment (e.g. roads, bridges, markets and irrigation schemes). At the district and provincial levels, policies and programs, including a legislative framework for businesses, need to support a pro-poor market orientation.

The DBRP project supported the creation of these enabling conditions through:

- participatory and market-oriented Socio-Economic Development Planning (SEDP) focused on short local value chain development;
- a Commune Investment Fund (CIF) to be spent for 90% on infrastructure development and 10% on special market training and study tours for small producers and traders,
- a Common Interest Group (CIG) development fund that provides loans and training to CIGs involved in short value chains with minimum 70% poor and 50% female members; and
- capacity-building of provincial, district and commune line agencies (incl. DOLISA and DOIT) for creating an enabling businesses environment and enhancing access to credit, training and services for small start-up businesses and SMEs.

More participatory and market oriented Socio-Economic Development Planning (SEDP) is a critical mechanism in the theory of change of the DBRP. The annual SEDP sets out the priorities for CIF and CIG development investments in the communes and their villages.⁷⁴ By making the SEDP processes more market-oriented and participatory, development efforts are expected to be more responsive to the concrete market linkage and inclusion issues at grassroots level. By strengthening the management and service-delivery capacities of the various line agencies at the various levels, these processes are assumed to facilitate more effective implementation of all government projects/programs (incl. those funded by international agencies/institutions).

The theory of change is based on assumptions underpinning the ‘markets for the poor’ paradigm, about the roles and capabilities of rural citizens (in particular those related to gender and generational differences); about how poor/near-poor households can be reached by village-level groups, mass organisations, and line agencies; and about the influence of context. Their validity determine the degree and nature of project impact. Hence validation of these assumptions with stakeholders can help identify potentially useful indicators for impact monitoring. Evidence from the PIALA impact assessment shows that the following three key assumptions don’t hold true:

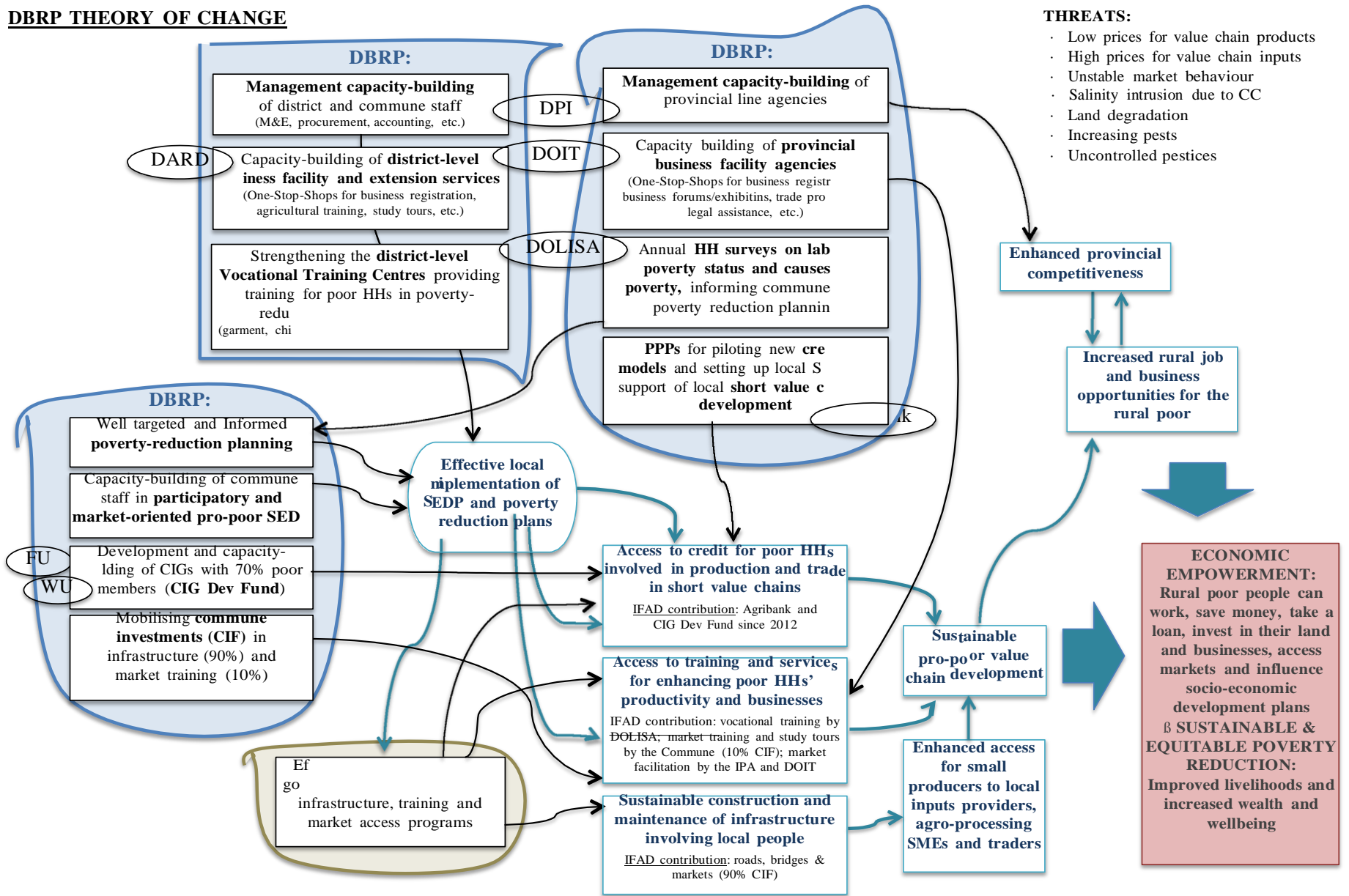
- A major assumption of the DBRP is that *if the SEDP processes are made more participatory and market-oriented*, while enabling conditions are put in place related to service provision at local levels and appropriate business facilitation and PPP mechanisms and policies at district and provincial levels, then local economies will grow and *poor and vulnerable groups (incl. women, youth and minorities) will be able to equally participate and benefit*.
- Another important assumption is that *poor households have equal access to services* and thus are equally reached, *if mass organisations are strengthened*. In principle everybody (including non-members) can enjoy these mass organisations’ support and services.

⁷⁴ The CIF and SEDP processes are managed by the Commune Development Board (CDB). The CIG development fund is managed by the Women’s Union (WU) in collaboration with the Farmers’ Association (FA). This is expected to create greater ownership and responsibility.

- A third important assumption is that the DBRP's primary target groups or beneficiaries (i.e. the rural poor people, ethnic minorities and women) can best be targeted not so much directly but rather through a layered strategic approach that involves more wealthier farmers and successful business people and develop the relationships between the different groups and institutions in the villages/communes for risk sharing and enabling people to learn from more successful and richer people in developing businesses and playing the markets.⁷⁵

⁷⁵ From the interview with Mrs. Duong Quynh Le, Director of the Multilateral Division at the Department of Debt Management and External Finance, Ministry of Finance (MOF), Hanoi, 2/07/2013.

DBRP THEORY OF CHANGE



Annex VI. PIALA Generic Research Questions

The overarching evaluation questions for the PIALA pilots agreed at the consultation workshop in Rome in Oct 2012 and included in the PIALA research strategy, are⁷⁶:

- What has changed for whom and how?
- How sustainable are these changes likely to be?
- What has been the contribution of the IFAD-supported projects to these changes compared with other possible causes?
- What are the implications of this understanding for project management and planning?

The research strategy outlined further following generic research questions that would have to be tested and refined in the PIALA pilots⁷⁷:

Rural poverty-related questions:

Hypothesis: Extending relevant, appropriate and high quality services (e.g. financial and other extension services) and building the capacity of poor rural people to access those services will impact positively on poor women's and men's livelihoods and wellbeing

- What are the impact level trends in household income and assets, food security and nutrition, and number of people in poverty? How have these trends occurred in their particular contexts?
- How many people have benefited directly / indirectly from the project? How many people have participated in some sort of project activity?
- What groups of people have benefited most/more/less/least from the project? Rich, medium, poor people? Smallholder, landless? Centrally located communities, communities in remote areas? Men, women, youth? Ethnic groups?
- How or in what ways has the project benefited or disadvantaged these groups and in which ways (differentiated by social sub-group)? Has anyone suffered, or become worse off, as a result of the project? Who? How? Why?
- How do these different groups perceive and experience their poverty status? When do they no longer consider themselves as poor, or find themselves in a position where they can seize new opportunities to overcome their poverty?
- How many people no longer consider themselves as poor compared to before the project influenced their context/lives? To which of these groups do they belong?
- How sustainable are the project impacts likely to be? What are the key threats and their likelihood?

Questions related to key enablers:

Hypothesis: Effective and sustained impacts on wellbeing are best achieved through enhancing individual and group capability and addressing the institutions that govern behaviour and relations

- What has influenced rural poverty trends in particular contexts? (influencing factors/actors, enablers/disablers, key processes/mechanisms)
- What and how do (different groups of) poor people see as key enablers and disablers of change?
- How and to what extent have changes in individual and collective capability – their assets, attitudes, motivations, capacities and influence – contributed to these trends?
- How and to what extent have changes in policies, institutional structures and capacities, and regulatory frameworks contributed to these impact-level trends?
- What has changed in the relationships between different groups in the targeted rural communities due to the project? What has changed in the relationships within households?
- How has the project affected relationships between communities (or particular groups within these communities) and the local or district government?
- How has the project affected relationships between the communities (or particular groups within these communities) and private actors who can influence the project?

⁷⁶ See: IFAD, 2013: pp. 12-13.

⁷⁷ Ibid, p.15-17.

- How have these changes in relationships, institutions, policies and agency affected women's and men's conditions, positions and opportunities within and between the different groups benefitting less or more from the project?

Questions related to processes and risks:

Hypothesis: *The project's strategic interventions are key processes that will generate the envisioned impacts and enablers (or systemic changes). External risks will not undermine project outcomes and impacts. Internal risks will be manageable.*

- What were the crucial processes that have led to the combined and interactive changes in wellbeing, capability, institutions & policies, and relationships? How and to what extent has the project contributed to these?
- What were the mechanisms that have caused the relative in/exclusion of specific groups in specific areas of work? How and to what extent has the project influenced these?
- How has the anticipated project contribution to change been affected by external risks (e.g. declining market prices, climate change, migration, coordination amongst agencies and levels and incentives and capacities among agency staff)?
- How has the anticipated project contribution to change been affected by internal risks (e.g. resource constraints, social change)?
- How has the project's understanding and monitoring of risk enabled project managers to mitigate or manage these risks?

Annex VII. Intended PIALA Uses and Users

At the consultation workshop in Rome (17 Oct 2012), group discussions were facilitated around the question “who should participate and learn, and why?”

As a result of these discussions, it was agreed that:

- Global and national IFAD managers, national and provincial project managers, and the local research team should be involved in the design of the impact assessment.
- Local stakeholders such as rural poor people who are the intended project beneficiaries, their community leaders and organisations, and government agencies and other crucial “system changers”⁷⁸ should be involved in the data collection.
- Funders, managers, local stakeholders and local research team (as described above) should all be involved in the project contribution analysis.
- All stakeholders should ideally be considered as the potential users of the impact assessment outcomes, making it necessary to link analysis and feedback to learning and action at all levels.
- Priority however should be given to the national and local levels, since that’s where the greatest uptake and learning may be expected.
- Moreover, impact reflections around the ToC based on the evidence produced, should speak to the knowledge needs of key system changers and designers of next generation projects.
- IFAD country and global managers, government partners and other donors who are positioned to advocate or implement PIALA, should be considered as the main audience for the methodological reflections, while grassroots participants should be involved in the validation of the methodology.

In the PIALA research strategy, it is argued that utility or intended use/usefulness has several dimensions related to stakeholders’ different understandings of impact and their underlying change theories and assumptions. Broadly, utility is linked to learning and can be defined in ‘instrumental’ terms as the usefulness of evidence for adaptive management and performance improvement, as well as in ‘transformational’ or ‘empowering’ terms as an evaluation’s contribution to enhancing key stakeholders’ ability to engage with and influence decision-making, based on their increased understanding of change processes. The meaning or value of participation for the different stakeholders then depends on the extent to which their instrumental respectively empowering learning needs are fulfilled.

⁷⁸This was the name used by the participants at the global workshop in Rome on 17 Oct 2012.

Annex VIII. PIA-DBRP Specific Research Questions and Methods

KEY QUESTIONS	SUB QUESTIONS	FIELD LEVEL QUESTIONS	ASK WHOM?	METHODS
<p>What has changed in wealth/wellbeing at impact level? For whom?</p>	<p>1. Definition of wealth/wellbeing How do local people understand wealth and wellbeing? How consistent are people's definitions?</p> <p>2. Categories and characteristics: What different categories of wealth/wellbeing exist in people's understanding? What are the characteristics of each category?</p> <p>3. Distribution of households in each category: What is the proportion of households in each category? How has this proportion changed since 2008 – per category?</p> <p>4. Changes in proportion since 2008 – per category: Has there been a change in how people define poverty and wealth?</p>	<p>Could you please let us know:</p> <ul style="list-style-type: none"> - Your own definition of rich/poor - Groups (that not follow the state's definition) - Their characteristics(asset, source and amount of income, ways of spending money(health services, education, gai tri), business investment (hiring labor, loans (as capital); domestic labor resource. - Three most significant characteristics to distinguish the rich and the others. - The proportion of the rich/the poor in the village (percentage) - How has the perception of rich/poor changed in past 5 yrs? - How the group of the rich affects other ones (positive, negative, labour, voices in the community, business sharing experiences and labor tools) - How has the change in the percentage of each group in the past 5 years (resources and opportunities) 	<p>Mixed group:</p> <ul style="list-style-type: none"> • village leader • WU and FA and youth union leaders • 2 to 3 poor or near poor HHs (mix of women and men, all economically active) <p>Random HH sample</p> <p>M&E staff</p>	<ul style="list-style-type: none"> • Wealth-Wellbeing analysis • Mini-survey • Secondary data analysis (e.g. gov's poverty stats)
<p>What general changes have occurred in the villages? With what impact?</p>	<p>1. Changes in village What general changes have occurred in the villages? - concerning: livelihoods, institutions (governance, politics, rights), people's engagement in village planning and decision-making, infrastructure, culture and relationships, markets, and services (all sectors and extension, financial, legal, business).</p> <p>2. Who? Who has been affected by these changes? To what extent? Marginal or significant, quantified?</p> <p>3. Causes and effects</p> <p>4. Project contributions to this change</p>	<p>Please let us know:</p> <ol style="list-style-type: none"> 1. Changes (negative and positive) that affected your lives in the past 5 years) <ul style="list-style-type: none"> -Infrastructure (market, road, health unit, school) -Livelihoods (change in production, job placement, business opportunities) -Services (agricultural and aqua extension, veterinary, irrigation, bank, food safety, safe water) -Associations, groups, clubs, and community activities 2. What community groups (wealthy/poor, fame/female, old/young) have been affected by these changes (negative and positive)? The extent of change? <u>Note:</u> ask the number of male and female, new or upgraded roads, new business HHs, new jobs (+female), HHs getting out of poverty (female headed HHs, people who got vocational training (+female) 3. What are the reasons for the above changes? 4. Probe for project-related contribution to these changes 	<p>Mixed but women- or men-only groups, composed of:</p> <ul style="list-style-type: none"> • 2 CIG members, of which 1 poor • 2 BHH and SME members, of which 1 poor • 2-3 poor and relatively young HH members • max 1 elder <p>Key Informant Interviews</p>	<ul style="list-style-type: none"> • Generic Change and Causal Flow • Social and change mapping, with timeline
<p>What livelihood changes have occurred in the villages?</p>	<p>1. Livelihood changes for whom How have livelihoods changed and for whom? To what extent? Marginal or significant, quantified?</p> <p>2. Sustainability</p>	<p>Please let us know:</p> <ul style="list-style-type: none"> - List of occupations for women and man in the village - Percentage of women and men for each occupation in 2012/2008 - Rank the annual income of each occupation in 2012 (1-n) 	<p>3 mixed women- or men-only groups, e.g.:</p> <ul style="list-style-type: none"> • 1 with the male members of poor, active and (relatively) 	<ul style="list-style-type: none"> • Livelihood Matrix and Causal Flow • Secondary data analysis (e.g. gov's

<p>With what impact? How sustainable are these?</p>	<p>What are the risks to livelihoods? 3. Causes and effects What are the causes and effects of significant changes? 4. Project contribution How have project activities contributed to this change?</p>	<p>- List risks and score them for each occupation ((1-4 highest) - If you have the same condition of life, what job would you like to do ? Why - What groups (men/women, wealthy/poor, BHHs, CIGs, SMEs) have been affected (negatively and positively)? To what extent? - What are the reasons for the above changes? How the groups (wealthy/poor, women/men, CIGs, SMEs) have been affected by these changes? Did you have other supplementary jobs? - Probe for project-related contribution to these changes</p>	<p>young male-headed HHs, mixed with FA and male-headed CIG members; • 1 with the female members of poor, active and (relatively) young male-headed HHs, mixed with WU and female-headed CIG members; • 1 with the women of poor, active and (relatively) young female-headed households, mixed with WU and female-headed CIG members</p>	<p>poverty stats)</p>
<p>What institutional changes have occurred in the villages? With what impact? How sustainable are these?</p>	<p>1.Changes for whom How have institutions and power relations changed and for whom? To what extent? Marginal or significant, quantified? 2.Sustainability How sustainable are these institutional changes? 3.Causes and effects What are the causes and effects of significant institutional changes? 4. Project contribution How have project activities contributed to these changes?</p>	<p>Please let us know: - What are the organizations, agencies, associations, businesses, and individuals that affect (negatively and positively) your family life? Why? To what extent (<i>much, little</i>) - When you need advise, share important information, or need support, is it easy to contact them (directly, telephone, through the third person)? - Have these relationships been changed? Why? And to what extent? Any changes in the above relationship in the past 5 years? Reasons? - What is the most significant change? Why? For which groups? (wealthy/poor, men/women, CIGs, Associations, BHHs, SMEs) - Project contribution to these changes?</p>	<p>Mixed but women- and men-only groups, composed of 6-9 poor and nearly poor members of BHHs, CIGs, SMEs and poor HHs – mixed 50% women and 50% men. In-depth interviews on the CIGs and BHHs – how it works, problems, successes, impacts</p>	<p>• Institutional Analysis and Change Flow</p>
<p>3. Looking forward How good are the key changes, particularly for the poor, and what ‘good ideas’ exist to further enhance these changes? What does this mean for the project?</p>	<p>1. Field level What needs to happen to support change for the poor more effectively? Think about: livelihoods, institutions (governance, politics, rights), people’s engagement in village planning and decision-making, infrastructure, culture and relationships, markets, and services (all sectors and extension, financial, legal, business). 2. Analytical synthesis Has the project, on balance, contributed significantly to reducing rural poverty? If not, why not? If so, what seems to have made the difference? How can project contribute more effectively to that change process?</p>	<p>What has been done well, what has not in socio-economic development of the village, especially for the poor? If not, what needs to be done to ensure good practice, effectiveness and continuance of SEDP activities in the future?</p>	<p>Village head, poverty reduction officials, Participants in FGDs of In-depth interviews</p>	<p>• Wealth, wellbeing analysis (social mapping, timeline, ranking) • Generic change ranking + Causal Flow • Livelihoods matrix analysis + Causal Flow • Institutional analysis + Causal Flow • SEDP Group Interview (semi-structured)</p>

Annex IX. PIA-DBRP Sampling Structure and Research Schedule

Sampling structure

On the basis of the initial sampling framework in the PIALA research strategy and discussions held with experts in the PIALA external reference group, the following sampling strategy was developed together with the stats team of the local research organisation and the project M&E team.

- Of the 50 communes⁷⁹ in the 8 districts of Ben Tre province (Mekong Delta) where the project had been implemented, only the 26 communes⁸⁰ where the project was started in 2008, were considered for sampling. These 26 communes were divided into 3 agro-ecological zones, namely: salt, brackish and fresh water.
- From the 8 project districts, 3 were identified that fitted with these 3 zones, namely: Ba Tri District (salt water), Mo Cay District (brackish water) and Cho Lach (fresh water).
- For each of these 3 zones/districts, project communes were stratified from the 26 project communes by distance⁸¹ from the main inter-communal road, since this determined the relative accessibility of the communes/villages, which likely affected value chain market access and helped avoid researcher bias. This resulted into two clusters with a nearly 50-50 distribution of communes relatively close with a distance less than 2km, and communes relatively far with a distance greater than 2km.
- Project communes were then alternately and randomly selected from each of the two clusters until we had 6 communes in each zone/district, thus 18 project communes.
- In each sampled commune, one project or focus village⁸² was randomly selected for conducting a survey on the rural poverty impact indicators in the PIALA analytical framework (see Annex II). From these 18 focus villages, one relatively close and one relatively distant village in each of the three zones/districts were randomly selected for in-depth participatory research, thus 6 focus villages in total.
- In each of the sampled salt water and fresh water districts (2 of the 3 sampled districts), 3 non-focus communes were randomly sampled, and in each of these communes, 1 non-focus village was randomly selected, providing us with 6 non-focus villages for the mini-survey.⁸³ In 1 randomly selected village of the 3 in each district (thus in 2 non-focus villages in total), in-depth participatory research was conducted.
- Finally, in each of the sampled 18 project and 6 non-focus villages, 30 households were randomly sampled for the mini-survey (720 households in total), and approx. 55 villagers and 10 leaders were selected for the in-depth participatory research (approx. 520 participants in total).

Mini-survey: [(2+4 villages) * 3 districts] *factual* + [(1+2 villages) * 2 districts] *counterfactual*
= [18 * 30 HHs] *factual* + [6 * 30 HHs] *counterfactual*
= 540 HHs *factual* + 180 HHs *counterfactual* = **720 HHs in total.**

In-depth research: [2 villages * 3 districts] *factual* + [1 village * 2 districts] *counterfactual*
= [6 * (ca. 55 villagers⁸⁴ + 10 leaders)] *factual* + [2 * (ca. 55 villagers + 10 leaders)] *counterfactual*
= approx. 390 participants *factual* + 130 participants *counterfactual* = **approx. 520 participants in total.**

⁷⁹ In Vietnam, a commune is an geographical and administrative unit that is composed of a number hamlets and commune-level town or a village that is slightly bigger than the other hamlets.

⁸⁰ One of the criteria for selecting projects for piloting PIALA (agreed at the consultation workshop in Rome on 17 Oct 2012) was that they should have had been implemented for at least 5 years. The remaining 24 communes were added to the DBRP portfolio only after 2010. In these communes, the project had thus been implemented for only 3 years and thus presumably would be less mature in terms of impact.

⁸¹ We classified relatively close if distance was less than 2km, and relatively far if distance was greater than 2km. We ended having about 50-50 distribution of close and distant communes.

⁸² Focus villages are geographically targeted by DBRP village and commune-level interventions.

⁸³ Non-focus villages are not geographically targeted but exposed to the DBRP by its interventions at district and provincial levels.

⁸⁴ Population size of villages in Ben Tre province is average between 200 and 400 households.

Overall research schedule

		Mo Cay Bac and Nam Districts⁸⁵ 6 focus villages for survey, of which 2 in-depth		Ba Tri District 6 focus villages for survey, of which 2 in-depth	
		TEAM 1: Lien	TEAM 2: An	TEAM 3: Huong	TEAM 4: Tuy
Districts:	Day 1	SSIs at Mo Cay Bac District		SSIs at Ba Tri District	
Villages:		Mini-survey in villages 1.1 and 1.2	Mini-survey in villages 1.3 and 1.4	Mini-survey in villages 2.1 and 2.2	Mini-survey in villages 2.3 and 2.4
	Day 2-5	Mini-survey in village 1.5	Mini-survey in village 1.6	Mini-survey in village 2.5	Mini-survey in village 2.6
		In-depth research in village 1.5	In-depth research in village 1.6	In-depth research In village 2.5	In-depth research In village 2.6
Ben Tre:	Day 6	AM Village Feedback & PM in District Team reflections and synthesis			
	Day 7	REST			
		Cho Lach District 6 focus villages for survey, of which 2 in-depth		Ba Tri and Cho Lach Districts 6 NON-focus villages for survey, of which 2 in-depth	
		TEAM 3	TEAM 4	TEAM 1	TEAM 2
Districts:	Day 8	SSIs at Cho Lach District		SSIs at Mo Cay Bac District	
Villages:		Mini-survey in village 3.1 and 3.2	Mini-survey in village 3.3 and 3.4	Mini-survey in NON-focus village 2.7 and 2.8	Mini-survey in NON-focus village 1.7 and 1.8
	Day 9-13	Mini-survey in village 3.5	Mini-survey Village 3.6	Mini-survey in NON-focus village 2.9	Mini-survey in NON-focus village 1.9
		In-depth research in village 3.5	In-depth research in village 3.6	In-depth research in NON-focus village 2.9	In-depth research in NON-focus village 1.9
Ben Tre:	Day 13-14	TEAM REFLECTION AND SYNTHESIS			

⁸⁵ These two districts were before 2009 one district.

Village research schedule

WEEK 1 (15-21/7/2013):

Day 1 (15/7)

2 researchers are conducting mini-surveys with 20 poor HH members (primary beneficiaries), each in two villages (20 HHs / village).

Simultaneously 2 other researchers are conducting:

9:00-11:00 Group meeting with the members of the District Project Management Office (DPMO):

- *Head of Office*
- *Project M&E officer*
- *Project accountant*
- *Market facilitators*

11:30 -13:00 Lunch break

13:00 / 14:00 / 15:00 / 16:00

Individual interviews of 45 min with technical officers of:

- *Division of Economic Development and Infrastructure*
- *Division of Finance & Planning*
- *Division of Agriculture and Rural Development*
- *Division of Labour and Social Affairs*

17:00-19:00 Diner break

Day 2 (16/7)

One researcher finishes the 20 remaining mini-surveys in the villages of day 1 (10 in each village).

Meanwhile, the two other researchers move to a new village where the in-depth research will take place for the rest of the week. The schedule below (day 2-6) described the schedule for this in-depth research.

8:00-9:30 Livelihood change ranking and focus group discussion with the commune leaders (incl. FA, WA, commune head, commune planning staff and the accountant –all members of the CDB = Commune Development Board) at Commune People's Committee (2 researchers).

Ask about:

- The CIF and its link to new PPPs -- how CIGs are initiated and changes pre/post 2008 and pre/post 2010
- The CIG Development Fund and its role in credit and support for value chain development – which products are being focused on and what kinds of initiatives to improve production, processing and marketing

10:00-11:30 KII with commune poverty reduction official and commune planning staff⁸⁶ about SEDP process at Commune People's Committee-(2 researchers).

11:30-13:00 Lunch break

13:00-15:00 Social mapping, timeline and wealth/wellbeing ranking with a mixed group at the village (2 researchers).

Group composition total 6-8 people:

- *village leader*
- *WU and FA and youth union leaders*
- *2-3 poor or near poor HHs – preferably mix of women and men, all economically active*

15:30-16:30 8 Mini surveys (2 researchers)

17:00-19:00 Diner break

19:00-21:00 Reflection and data collation

⁸⁶ The commune planning staff are responsible for commune level planning processes, including for the SEDP, the commune annual planning and the DBRP annual planning.

Day 3 (17/7)

- 8:00-10:00 Generic Change Cause/Impact Flow with a mixed economic class women-only group (2 researchers)
Group composition: total 8-10 women
- max 2 female CIG members, of which 1 poor
 - max 2 female BHH members, of which 1 poor
 - max 2 commune female SME members
 - max 1 older woman
 - max 2 poor and relatively young female HH members
 - max 2 poor women from female-headed households
- In-parallel: 4 Mini-surveys (1 researcher)
- 10:00-11:00 6 Mini surveys (3 researchers)
- 11:30-13:00 Lunch break
- 13:00-15:00 Generic Change Cause/Impact Flow with a mixed economic class men-only group (3 researchers)
Group composition: total 8-10 men
- max 2 male CIG members⁸⁷, of which 1 poor
 - max 2 male BHH members, of which 1 poor
 - commune male SME members (max 2)
 - max 2 men from female-headed households
 - max 1 older man
 - max 1 male youth leader or economically active young man
- 15:00-17:00 12 Mini surveys (3 researchers)
- 17:30-19:00 Diner break
- 19:00-21:00 Reflection and data collation

Day 4 (18/5)

- 8:00-09:30 SEDP Process Flow and Assessment with SEDP participants about people's engagement in SEDP planning and decision-making (3 researchers)
Group composition:
Total 6-8 SEDP participants, including all those considered poor and mix of women and men (around 3-4)
- 10.00-11:30 Livelihood Matrix and Causal Flow with a mixed but men-only group (3 researchers)
Group composition: Total 8 men
- 2-3 male CIG members, of which 1 poor
 - 2-3 male BHH members, of which 1 poor
 - 2-3 poor and relatively young men head of household
- (Avoid attendance of the same participants of the generic change group on Day 3)
- 11:30-13:00 Lunch break
- 13:00-15:00 Livelihood Matrix and Causal Flow with a mixed but women-only group (3 researchers)
Group composition: Total 8 women
- 2 female CIG members, of which 1 poor
 - 2 female BHH members, of which 1 poor
 - 2-3 poor and relatively young female HH members
 - 2 head of female-headed households
- (Avoid attendance of the same participants of the generic change group on Day 3)
- 15.30-18:00 Data collation –identify gaps in data
- 18:00-20:00 Diner break
- 20:00-21:00 Reflection

⁸⁷ In villages where there are no CIGs, members of the CIG Development Fund will be invited instead.

Day 5 (19/7)

1 researcher continues with data collation, while the other two continue with the FGDs:

- 8.00-10:00 Institutional Analysis and Cause/Effect Flow with a mixed but women-only group
Group composition: Total 6-9 women
- 2 female BHH members, of which 1 poor
 - 2 CIG members/leaders
 - any women involved in SMEs
 - 2-3 poor and relatively young women from female headed households
- (Avoid attendance of the same participants from other discussions)*
- 10.30-11:30 Data collation
- 11.30-13:00 Lunch break
- 13.00-15:00 Institutional Analysis and Cause/Effect Flow with a mixed but men-only group
Group composition: Total 6-9 men
- 2 male BHH members, of which 1 poor
 - 2 CIG members/leaders
 - any men involved in SMEs
 - 2-3 poor and relatively young men head of household
- (Avoid attendance of the same participants from other discussions)*
- 15:30-19:00 Data collation, village feedback preparation
- 19:00-20:00 Diner break
- 20:00-22:00 Village feedback preparation

Day 6 (20/7)

- 7:00-9:00 Village feedback preparation.
- 9:00-12:00: Feedback and initial sensemaking with the village
Max. 21 invited participants, of which 30% economically active poor HH members (at least half must be women), including:
- 2 members of the District Project Management Office (DPMO)
 - 2 district technical officers from the line divisions
 - 1 Commune poverty reduction official
 - 4 village leaders (incl. village head, commune head, FA leader, WU leader)
 - 12 participants from each group discussion that took place in the past four days (max. 2 per group), incl.:
 - 2 near-poor CIG members (preferably 1 man and 1 woman)
 - 2 near-poor BHH members (1 man and 1 woman)
 - 2 economically active women from poor male-headed HH
 - 2 economically active women from poor female-headed HH
 - 2 economically active men from poor male-headed HH
 - 2 poor SEDP participants (1 man and 1 woman)
- 12:30-14:00 Lunch break
- 14:00-16:00 Processing of the results of the village feedback –finalise data collation
- 16:00-18:00 End-of-week reflection with the entire research team, and travel back to Ben Tre.

Day 7 (21/7): FREE

WEEK 2 (22-28/7):

Day 1-3 (22-24/7): remains the same as in Week 1.

Day 4 (25/7):

8:00-09:30 SEDP Process Flow and Assessment with SEDP participants about people's engagement in SEDP planning and decision-making (3 researchers)
10:00-11:30 Livelihood Matrix and Causal Flow with a mixed but men-only group (3 researchers)
11:30 -13:00 Lunch break
13:00-15:00 Livelihood Matrix and Causal Flow with a mixed but women-only group (3 researchers)
15.30-17:30 Institutional Analysis and Cause/Effect Flow⁸⁸ with a mixed but men-only group (2 researchers –one researcher continues with data collation)
18:00-20:00 Diner break
20:00-21:00 Reflection and travel back to Ben Tre

Day 5 (26/7):

09:00-11:30 Data collation and village feedback preparation
11:30-13:00 Lunch break
13:00-15:00 Overall methodological reflection with the entire team
15:00-19:00 Data collation and village feedback preparation
19:00-20:00 Diner break
20:00-22:00 Village feedback preparation

Day 6 (27/7): **FREE**

Day 7 (28/7):

9:00-11:30 Interim reporting
11.30-13:00 Lunch break
13:00-15:00 KIIs with available people around CIG support, access to training, credit and market services, SEDP process and other issues
15:30-19:00 Finalise data collation and preparation of village feedback

WEEK 3 (29/7):

Day 1 (29/7):

08:00-10:00 Institutional Analysis and Cause/Effect Flow with a mixed but women-only group (2 researchers –one researcher continues with data collation)
10:00-11:30 Village feedback preparation.
13:00-16:00 Feedback and initial sensemaking with the village
16:00-18:00 Processing of the results of the village feedback –finalise data collation
18:00 Travel back to Hanoi

⁸⁸ Lists of BHHs, CIGs, SMEs, HHs in the selected villages will be provided.

Annex X. Table of impact domains, indicators and target groups

Impact domains	Indicators/criteria of change	Target groups	Data sources
Institutional capacity	Commune capacity <i>(in participatory SEDP & poverty reduction planning and management)</i>	Commune officials <i>(particularly planning, marketing and poverty-reduction officials, and commune managers)</i>	<ul style="list-style-type: none"> • KIIs with commune, district, provincial and national officials; • Secondary data.
	District capacity <i>(in management and providing vocational training and business facility and extension services to poor and near-poor households)</i>	District officials <i>(particularly extension and technical support officials, vocational training centers, district managers and farmer and women union leaders)</i>	
	Provincial capacity <i>(in management, business facilitation, poverty reduction, and PPPs)</i>	Provincial officials <i>(particularly line agency officials from DOLISA, DPI and DARD, and farmer and women union leaders)</i>	
Empowerment	People's voice and influence <i>(e.g. in socio-economic and poverty-reduction planning and the implementation of these plans)</i>	<u>Primary:</u> Poor and near poor households <i>(particularly minorities and female-headed)</i>	<ul style="list-style-type: none"> • KIIs and focus group discussions on SEDP; • Secondary data. • Focus group discussions on institutional relations and causal flow analysis (gender-disaggregated).
	Changes in institutional relationships in the village and commune <i>(e.g. institutional relationships that create greater access to services, training and credit for poor and near-poor households)</i>	<u>Secondary:</u> All households <i>(incl. self-management groups)</i>	
Access	Access to credit for production and trade of short value chain products	Small local enterprises and small business households, producers and traders <i>(particularly poor and near-poor)</i>	<ul style="list-style-type: none"> • Mini-survey (gender-integrated) • Secondary data • Focus group discussions on livelihoods changes causal flow analysis (gender-disaggregated); • Focus group discussions on institutional relations and causal flow analysis (gender-disaggregated); • Focus group discussions on generic changes and causal flow analysis (gender-disaggregated);
	Access to training and services for production and trade of short value chain products		
	Access to markets for inputs and outputs of short value chains		
Jobs & Livelihoods	Changes in crops and livestock cultivation	<u>Primary:</u> Poor and near poor households <i>(particularly minorities and female-headed)</i>	<ul style="list-style-type: none"> • Focus group discussions on livelihoods changes causal flow analysis (gender-disaggregated); • Focus group discussions on generic
	Jobs and business opportunities	<u>Secondary:</u>	

		Business households and small enterprises	changes and causal flow analysis (gender-disaggregated); <ul style="list-style-type: none"> • Secondary data on value chains.
Wealth & Wellbeing	Food security (e.g. length and frequency of hunger periods)	Poor and near poor households (<i>particularly minorities and female-headed</i>)	<ul style="list-style-type: none"> • Mini-survey (gender-intergrated); • Secondary data.
	Income security (e.g. income increase, secure income sources, job/income-related risks, number of people moving out of poverty –from ‘poor’ to ‘near-poor’, and from ‘near-poor’ to better off)		Focus group discussions on wealth/wellbeing and social mapping

Annex XI. PIALA Report Outline⁸⁹—as applied in DBRP, VietNam

Executive Summary (2 pages)

Part A. PIALA (3-4 pages)

1. Intentions of PIALA, core principles/values
2. Criteria for assessing its success
3. Overview of methodology: what was done, successes and limitations

Part B. Assessing the impact of DBRP (20-25 pages)

1. DBRP description and theory of change:

This concerns an elaborated narrative plus visualization (2-3 pages max). Some comment on the difference between original logframe and current, updated version (elaborated through stakeholder interviews)

2. Analysis of key changes and their causes per theme:

The evaluation aims to be looking at IFAD-supported contributions to key mechanisms or processes of change leading to rural poverty impact. So in the DBRP's theory of change, this implies looking at the changes in "institutional capacity" and DBRP's part in this, and the effects of these changes on changes in "empowerment" (or "voice" and "relationships") as a key mechanism, "access" as another key mechanism, and "livelihoods/jobs" as the impact domain that directly affects rural poverty impact. So the document will need to reflect this cascading effect. Also, we need to focus on (a) most relevant changes (relevant to impact as shown in the theory of change, (b) causes of these changes, (c) DBRP's share in these causes, and (d) the relative strength of evidence of these changes, its causes and DBRP's contributions. Doing this for each impact domain/theme should enable us to eliminate doubtful and identify most plausible explanations of impact (or non-impact) and the DBRP's influences.

- Introduction, using the table with the impact themes/domains, criteria/indicators, and target group/s for each theme.
- Each of the five impact domains in the ToC (i.e. "institutional capacity" + "relationships and voice" → "access to training/services/credit" → "livelihoods and jobs" → "wealth and wellbeing") is analysed in a separate section. The collation tables are used to analyse and summarize in a short narrative the key changes with strong evidence, and the key expected changes with little or no evidence, the relevant causes and the extent of contribution compared to what was expected/intended:
- *Key changes*, including variation across villages and with comparison of non-focus villages. Focus on evidence of changes that are important to assess the DBRP's contribution claims that are indicated in the Theory of Change.⁹⁰ Weak and strong evidence of changes must be included, as well as a plausible explanation of why there is weak/strong evidence (e.g. it's too early to expect any significant change, which explains the availability of weak evidence). Data on food security and income from the mini-survey must be compared with project baseline data. Data from the mini-survey must be disaggregated per poverty category

⁸⁹ This overview is a synthesis of the initial outline that was sent to the local researchers and the additional guidance and coaching provided to them over email.

⁹⁰ All expected/intended and unexpected/unintended changes that are relevant, should be included in the analysis. Also "no change" may imply a finding that is relevant. Relevance means that it relates to one of the impact domains in the theory of change and helps test its causal link with its causing impact domain. For instance: "no food shortage/increase due to stable food provision for the whole family" is about "food security" and indicates no change in food security. It's a "no-change" that is relevant to the impact domain "wealth & wellbeing", of which a major cause may be for instance "more home gardening and raising chickens and pigs for family consumption". The latter is not an income-generating activity and thus not relevant to the impact domain "livelihoods & jobs" that is assumed to be an important cause of greater food security. Its a relevant cause of the "no-change in food security", however, because it may indicate that it's not so much jobs and livelihood opportunities that has created sustainable food security, but poor people's production for their own consumption. Combined with other findings, we might be able to conclude that, although changes in jobs and livelihoods have contributed to greater wealth, its effects on wellbeing in terms of food security and happy family life is less significant (or perhaps even negative in certain cases where families are apart because of labour migration).

and gender in order to be useful for saying something about changes in food security, income and credit for poor and near poor households, and men and women.

- *Relevant causes*: For each change, the related causes, inclusive project-related and non-project-related causes, must be described.
- *Extent of contribution*: A qualitative and quantitative estimation of the DBRP's actual share in these causes (e.g. which of these were due to the project and to what extent or how much?) compared to its contribution claims⁹¹. A rating and final comment need to be provided on the extent of contribution to each key change (see table below).

<i>Degree of contribution</i>	<i>Contribution descriptors</i>
Strong	Change is relatively clear, consistent, significant, and affecting target groups. Several clear links to DBRP interventions and supported activities.
Reasonable contribution	Change is more or less clear, consistent, significant, and affecting target groups. Few clear links to DBRP interventions and supported activities.
Minimal contribution	Change is rather unclear, inconsistent, insignificant and/or insufficiently reaching target groups. Limited evidence of DBRP contributions.
No contribution	Change is consistently absent, insignificant and not reaching targeted groups. Impact is unacceptably weak and does not meet minimum expectations. No evidence of DBRP contribution, or evidence of weak contribution.
Insufficient evidence	Evidence unavailable or of insufficient quality to determine impact and contribution.

3. Summary of DBRP contribution: (3 pages):

- to each change domain
- overall 'is DBRP meeting expectation of impact'? Why or why not?

ANNEXES

1. List of methods, per village/District, including number of people per category
2. Interim reports per District
3. References (secondary data)
4. Village assessment summary

⁹¹ For instance: DBRP has proportionally divided the total amount of CIF among project communes using certain criteria, which resulted in perhaps only 10% of all infrastructure development. The intention and thus contribution claim was not to fund and realize all infrastructure, but to help the communes realize their own plans, which then presumably would lead to greater market access.

Bibliography

- Ackerman, J. (2011). Co-governance for accountability: beyond “exit” and “voice.” In *The Participation Reader*. Zed Books.
- Bamberger, M. (2005). Institutionalizing impact evaluation. A key element in strengthening. In *From Policy to Results: Developing national capacities for country monitoring and evaluation systems*. UNICEF.
- Barder, O. (2012). Complexity, Adaptation, and Results. *Center for Global Development*. Retrieved from <http://blogs.cgdev.org/globaldevelopment/2012/09/complexity-and-results.php>
- Befani, B. (2012). *Models of Causality and Causal Inference* (Review prepared as part of the DFID study: Broadening the range of designs and methods for impact evaluation). DFID (Department for International Development).
- Bonbright, D., & Power, J. (2010). Private Sector Metrics Contributions to Social Change: Customer Satisfaction Meets Agriculture Development. *IDS Bulletin*, 41(6).
- Bourguignon, F., Gertler, P., Kress, D., Manning, R., & Gottlieb, J. (2006). *When Will We ever learn? Improving Lives through Impact evaluation*. Centre for Global Development.
- Chambers, R. (2008). *Revolutions in Development Inquiry*. Earthscan.
- Chen, H. T., Donaldson, S. I., & Mark, M. M. (2011). Validity frameworks for outcome evaluation. *New Directions for Evaluation*, 2011(130).
- Copestake, J. (2013). Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches (Draft 18 Oct 2013). Centre for Development Studies, University of Bath.
- Cornwall, A. (2002). *Making spaces, changing places: situating participation in development*. Institute of Development Studies (IDS).
- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation*, 32(2).
- Davies, R. (2005). Scale, Complexity and the Representation of Theories of Change Part II. *Evaluation*, 11(2).
- Davies, R., & Dart, J. (2007). *The “most significant change” (MSC) technique: a guide to its use*. Rick Davies.
- Forss, K., Marra, M., & Schwartz, R. (2011). *Evaluating the Complex: Attribution, Contribution, and Beyond*. Transaction Publishers.
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. John Wiley & Sons.
- Gaventa, J., & Barrett, G. (2010). *So what Difference Does it Make?: Mapping the Outcomes of Citizen Engagement*. Institute of Development Studies (IDS).
- George, A. L., & Bennett, A. (2005). *Case Studies And Theory Development In The Social Sciences*. MIT Press.
- Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions and emerging confluences. In *The SAGE handbook of qualitative research*. SAGE.
- Holland, J. (2013). *Who Counts? The Power of Participatory Statistics*. Practical Action.
- Holma, K., & Kontinen, T. (2011). Realist evaluation as an avenue to learning for development NGOs. *Evaluation*, 17(2).
- IFAD. (2007). IFAD Innovation Strategy. International Fund for Agricultural Development (IFAD).
- IFAD. (2011a). *Draft Report of the Consultation on the Ninth Replenishment of IFAD’s Resources*. International Fund for Agricultural Development (IFAD).

- IFAD. (2011b). IFAD Strategic Framework 2011-2015. Enabling poor rural people to improve their food security and nutrition, raise their incomes and strengthen their resilience. International Fund for Agricultural Development (IFAD).
- IFAD. (2011c). Learning for improved performance. An integrated knowledge management system. IFAD.
- IFAD. (2011d). Results Measurement Framework 2013-2015. International Fund for Agricultural Development (IFAD).
- IFAD. (2012). *Methodologies for Impact Assessments from IFAD IX (Note to IFAD's Executive Board Representatives)* (Note to IFAD's Executive Board Representatives). International Fund for Agricultural Development (IFAD).
- IFAD (2013). PIALA Research Strategy. Improved Learning Initiative (Internal Document). International Fund for Agricultural Development (IFAD).
- Jones, H. (2009). Improving impact evaluation production and use. *Overseas Development Institute (ODI)*.
- Julnes, G. (2011). Reframing validity in research and evaluation: A multidimensional, systematic model of valid inference. *New Directions for Evaluation*, 2011(130).
- Kauffman, S. A. (1995). *At Home in the Universe: The Search for Laws of Self-organization and Complexity*. Oxford University Press.
- Khagram, S., & Thomas, C. (2009). Evidence for development effectiveness. *The Journal of Development Effectiveness*, 1(3).
- Leeuw, F., & Vaessen, J. (2009). *Impact Evaluations and Development: NONIE guidance on impact evaluation*.
- Levy, S., & Barahona, C. (2002). *How to generate statistics and influence policy using participatory methods in research* (Working Paper). Statistical Services Centre, University of Reading.
- Marchal, B., Belle, S. van, Olmen, J. van, Hoérée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2).
- McGee, R. (2013). *Improving the Evaluability of INGO Empowerment and Accountability Programmes* (Practice Paper No. 1). Centre for Development Impact (CDI).
- McGee, R., & Gaventa, J. (2011). *Shifting power: Assessing the impact of transparency and accountability initiatives* (IDS Working Paper No. 383). IDS.
- Merrifield, J. (2002). *Learning citizenship* (IDS Working Paper No. 158). Institute of Development Studies (IDS).
- Mertens, D. M. (2009). *Transformative Research and Evaluation*. Guilford Press.
- Mertens, D. M. (2010). Philosophy in mixed methods teaching: The transformative paradigm as illustration. *International Journal of Multiple Research Approaches*, 4(1).
- Mertens, D. M., Bledsoe, K. L., Sullivan, M., & Wilson, A. (2010). *Utilization of mixed methods for transformative purposes*.
- Midgley, G. (2007). Systems thinking for evaluation. In *Systems concepts in evaluation. An expert anthology*. EdgePress of Inverness.
- Mohan, G., & Hickey, S. (2004). Relocating participation within a radical politics of development: critical modernism and citizenship. In *Participation: from Tyranny to Transformation?: Exploring New Approaches to Participation in Development* (pp. 59–74). Zed Books.
- Mowles, Stacey, R., & Griffin, D. (2008). What contribution can insights from the complexity sciences make to the theory and practice of development management? *Journal of International Development*, 820(6).

- Nobuko, F. (2010). *Beyond Logframe; Using Systems Concepts in Evaluation*. Foundation for Advanced Studies on International Development (FASID).
- OECD-DAC. (2002). *Glossary of Key Terms in Evaluation and Results Based Management*. OECD-DAC.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. SAGE.
- Patton, M. Q. (2010). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use* (1st ed.). Guilford Press.
- Patton, M. Q. (2012). A utilization-focused approach to contribution analysis. *Evaluation*, 18(3).
- Pawson, R. A. Y. (2008). Causality for beginners.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. SAGE.
- Ravallion, M. (2012). Fighting poverty one experiment at a time. A review of Abhijit Banerjee and Esther Duflo “Poor economics: A radical rethinking of the way to fight global poverty.” *Journal of Economic Literature*, Forthcoming.
- Reeler, D. (2007). *A theory of social change and implications for practice, planning, monitoring and evaluation*. Centre for Developmental Practice.
- Rihoux, B. (2006). Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods. *International Sociology*, 21(5).
- Rihoux, B., & Ragin, C. C. (2009). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE.
- Rogers, P. (2009). Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. *The Journal of Development Effectiveness*, 1(3).
- Russell, J. (2012). A Philosophical Framework for an Open and Critical Transdisciplinary Inquiry. In *Tackling Wicked Problems: Through the Transdisciplinary Imagination*. Earthscan.
- Saini, M., & Shlonsky, A. (2012). *Systematic Synthesis of Qualitative Research*. Oxford University Press.
- Skuse, A., Fildes, J., Tacchi, J. A., Martin, K., & Baulch, E. (2007). *Poverty and digital inclusion: Preliminary findings of finding a voice project*. UNESCO.
- Snowden, D. (2002). Complex acts of knowing: paradox and descriptive self-awareness. *Journal of Knowledge Management*, 6(2).
- Snowden, D. (2007). The origins of Cynefin (part 1-6). *cognitive-edge.com*. Retrieved from <http://cognitive-edge.com/blog/entry/3505/the-origins-of-cynefin-part-1/>
- Stake, R. E. (2005). Qualitative Case Studies. In *The Sage handbook of qualitative research (3rd ed.)*. SAGE.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. (Working Paper No. 38). Department for International Development (DFID).
- Sweetman, D., Badiee, M., & Creswell, J. W. (2010). Use of the Transformative Framework in Mixed Methods Studies. *Qualitative Inquiry*, 16(6).
- Taylor, D. B. (2013). *Evidence Based Policy and Systemic Change: Conflicting Trends?* (Working Paper No. 1). The Springfield Centre for Business in Development.
- The M&E Harmonization Group of Food Security Partners. (2013, July). The Food Security Learning Framework. BMGF, USAid, DFID, FAO, IFPRI, IFAD, OECD, MCC, WB.
- Van Hemelrijck, A. (2014). *Understanding ‘Rigor’: Challenges in Impact Evaluation of Transformational Development*. PhD Research Outline Paper (revised version). Institute for Development Studies (IDS).
- Vogel, I. (2012). *Review of the use of “Theory of Change” in International Development* (Draft review report). Department for International Development (DFID).

- Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation, 18*(4).
- White, H. (2009a). *Some Reflections On Current Debates In Impact Evaluation* (Working paper No. 2009-1). International Initiative for Impact Evaluation (3ie).
- White, H. (2009b). *Theory-based impact evaluation: principles and practice* (Working Paper No. 3). International Initiative for Impact Evaluation (3ie).
- White, H., & Bamberger, M. (2007). Using Strong Evaluation Designs in Developing Countries: Experience and Challenges. *Journal of MultiDisciplinary Evaluation, 4*(8).
- White, H., & Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework* (Working Paper No. 15). International Initiative for Impact Evaluation (3ie).
- Williams, G. (2004). Towards a repoliticization of participatory development: political capabilities and spaces of empowerment. In *Participation: from Tyranny to Transformation?: Exploring New Approaches to Participation in Development*. Zed Books.
- Wood, B. (2011). *The evaluation of the Paris Declaration. Phase 2: Final Report*. Danish Institute for International Studies.
- Woolcock, M. (2009). Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. *The Journal of Development Effectiveness, 1*(1).
- Zelik, D., Patterson, E., & Woods, D. (2007). Understanding Rigor in Information Analysis. In *Proceedings of the Eighth International NDM Conference*.