

# Developing a Framework and Approach for IDRC Targeted Impact Evaluations

John Mayne

## Part 1: Overview report

1. Introduction	2
2. The challenges of assessing the impact of research for development	3
3. Useful concepts in assessing impact	6
4. Evaluability assessment for targeted impact evaluations	8
i. Impacts of interest	
ii. Pathways to impact	
iii. Likely contribution to impact	
iv. Measurement and data availability	
v. Evaluation designs	
vi. Usefulness, timing and cost of an impact evaluation	
5. The targeted impact evaluation examples	16
6. Concluding comments	18
References	20
Annex A: Glossary of terms	22
Annex B: Theories of change	24

## Part 2: Project Examples

E-procurement in Latin American and Caribbean countries	28
Chagas disease in Central America	40
School feeding programs in the Caribbean and the Philippines	49
Nutritional crops in Ethiopia	57

# Part 1: Overview Report

## 1. Introduction

“IDRC supports the generation of knowledge and innovation for positive change. The Centre will seek to enhance the well-being of larger numbers of people through these investments, accelerating development research on its way to big impact. “ (IDRC Strategic Plan 2015-2020, p. 7)

In the past ten years, most of IDRC’s evaluation had focused on assessing the outcomes of its funded research on influencing policy, practice and innovation of those using the research outputs—results within its sphere of influence. There is a renewed interest now in exploring the links between results in the sphere of influence and subsequent results in the sphere of interest, the effects on peoples’ lives or the environment. As part of that effort, IDRC is looking again at the feasibility and usefulness of undertaking impact evaluations of aspects of its programming.

This report aims to continue the discussion on impact evaluations at IDRC by suggesting possible ways to think about and move toward targeted impact evaluations. It sets out concepts that can be used to think about impact evaluations of IDRC’s funded projects and programmes, and discusses the types of issues that need to be considered in contemplating undertaking impact evaluations. It presents how impact evaluations could be framed, using four examples of past or current IDRC programming.

Impact evaluations seek to explore the contribution being made by interventions (such as IDRCs’ research projects and programs) and the longer-term effects of those interventions. In particular, IDRC is looking at ‘targeted’ impact evaluations, evaluations that focus on a set of related IDRC projects with a similar aim or aims. The intervention then is this package of IDRC projects. The idea here is to focus on more than a single research project and try to assess the impact of IDRC funding of the particular issue over a number of years and projects.

The study found that indeed, targeted impact evaluations using theory-based evaluation approaches were doable and could provide valuable insights on the impacts of IDRC funded research projects. Theory-based approaches are particularly well suited for complex interventions which is the typical case for IDRC. The examples illustrated numerous ways such impact evaluations could be carried out.

A key challenge in impact evaluations is that of causality, the causal link between the activities of the intervention and any observed impacts. Key in this is to realize that there are likely many causal factors at work and that the IDRC funded intervention is just one of several such factors. The issue then is the extent to which the IDRC intervention contributed to observed results and impacts, rather than trying to attribute impacts to IDRC.

Impact evaluations can be complex and costly. A structured evaluability assessment is needed in planning such impact evaluations, to ensure that meaningful and useful questions are addressed in a practical and timely manner. It would be important before undertaking any such evaluations to be quite clear on how the findings from the evaluation would be used. It would also be very useful to anticipate an impact evaluation as the intervention is implemented so that a useful theory of change could be developed and key baseline data tracked.

There are five more sections to this overview report. Section two outlines some of the main challenges in assessing the impact of research for development, both from IDRC's experience and from the evaluation literature. Section 3 offers some useful concepts to assessing impact; more detailed terms are included in a glossary in Annex A. Section 4 is the longest section, offering a structured approach to planning for targeted impact evaluations through an 'evaluability assessment'. Section 5 briefly summarizes the four examples of IDRC programming that were the basis for this study; each example is presented in full in Part 2. Finally, section 6 offers concluding comments.

## **2. The challenges of assessing the impact of research for development**

Assessing development impact is a challenge for most development interventions:

*There are multiple challenges to understanding causality*

There are many links in any chain that connects research to developmental improvements for a population. The relationships among the links will typically not be linear, they come about through different types of causal mechanisms, and contextual factors matter a lot. Actual results may not correspond to originally intended results, so initial results frameworks may or may not be helpful in framing causal connections.

*The contribution of others to impacts*

Once outside IDRC's sphere of control, IDRC can contribute to broader levels of change in our spheres of influence and interest, but the changes may never be wholly attributable to IDRC or the researchers supported. Many other actors and

factors also contribute to the observed changes. An impact evaluation may help articulate if there was a contribution, what the contribution was, and how significant it was to the change.

*The timeframe from research to development outcomes*

It can often take many years before the impacts of research are felt. So the results that are examined in an impact evaluation will probably be from research that was funded in previous programming periods. This raises logistical questions (are the individuals involved still around?), ownership questions (what program in IDRC would now assume ownership of the impact study?), and the type of learning that should be elicited (who can learn from the study?).

*Pathways to impact can be complex and uncertain*

The pathway from the actions of the intervention to impact can be long and complex. The research outputs may not be produced for beneficiaries, but rather intermediaries such as policy makers who then produce an ‘output’ intended for beneficiaries, resulting in a more complicated and longer pathway to impact. Delivering intended research outputs is not a given, nor is how they will be used, adding a real element of uncertainty as to how and if the research will have a development impact.

*There is more than just instrumental use of knowledge*

Impacts of research include specific innovations or specific policy or practice ideas, as an instrumental use of knowledge. However, the contribution of research can also include conceptual and other uses of knowledge as well. Lasting impacts of research might actually be in increased capacity to do and use research in the longer term, or build a lasting demand for evidence to inform practice or policy.

*Impact will not necessarily correspond to a funder’s or researchers’ organizational logic*

While IDRC organizes its work into programs and projects, “impacts” may not conform to those grant-making constructs. In fact, starting with IDRC grants and working forward from research findings to development impacts may be the wrong way to do it. It may be more useful to start with development impacts and work back to see if, when and how research fed into those impacts.

*Logistical and resource challenges*

This type of evaluation can be expensive, may require baseline data that are not available, and is both logistically and conceptually difficult. There may also be a need to meet the different needs of multiple key stakeholders.

Research impact has been studied and explored in a number of research fields. Recently there have been several reviews of this research impact literature (Milat, Bauman and Redman 2015; Greenhalgh, Raftery, Hanney and Glover 2016).

Greenhalgh et al. (2016: 4) note that:

Many approaches to assessing research impact combine a logic model (to depict input-activities-output- impact links) with a 'case study' description to capture the often complex processes and interactions through which knowledge is produced (perhaps collaboratively and/or with end-user input to study design), interpreted and shared (for example, through engagement activities, audience targeting and the use of champions, boundary spanners and knowledge brokers). A nuanced narrative may be essential to depict the non-linear links between upstream research and distal outcomes and/ or help explain why research findings were not taken up and implemented despite investment in knowledge translation efforts.

They conclude that:

(1) different approaches to impact assessment are appropriate in different circumstances; (2) the most robust and sophisticated approaches are labour-intensive and not always feasible or affordable; (3) whilst most metrics tend to capture direct and proximate impacts, more indirect and diffuse elements of the research-impact link can and should be measured; and (4) research on research impact is a rapidly developing field with new methodologies on the horizon

In their review, Milat et al. (2015: 6) conclude that:

Research impact assessment is a new field of scientific endeavour and typically impacts are assessed using mixed methodologies, including publication and citation analysis, interviews with principal investigators, peer assessment, case studies, and document analysis. The literature is characterised by an over reliance on bibliometric methods to assess research impact. Future impact assessment processes could be strengthened by routinely engaging the end-users of research in interviews and assessment processes. If multidimensional research impact assessment methods are to be widely used in practice by research funders and academic institutions, the right balance between comprehensiveness and feasibility must be determined.

And:

The vast majority of studies relied on principal investigator interviews and/or peer review to assess impacts, instead of interviewing policymakers and end-users of research.

On the basis of these reviews, a number of points can be noted:

- Many of the research impact models discussed in these reviews include forms of logic models depicting the pathway from research outputs to different types of impacts. The theory of change models discussed below have similar pathway models.
- The complexity of some models result in quite labour intensive and data demanding approaches if you want to explore each aspect of the models.<sup>1</sup> Assessing research impact is complex so there is a need for balance between the complexity of the models developed and the feasibility of data collection in approaches to be adopted.
- Cases studies of specific IDRC research impacts within a targeted impact evaluation will often likely be needed. But in exploring impact, there will be a need to go beyond dealing with the researchers funded by IDRC. Impact evaluations would need to gather perspectives of other actors involved in moving research to impact, and from the populations who have benefited (or not) from impacts.
- There is a proliferation of terms surrounding ‘research impact’ and some clarity for IDRC is likely needed.

The approaches discussed in this Overview Report are illustrated using four examples of possible targeted impact evaluations, namely projects working on:

- E-procurement in Latin American and Caribbean countries
- Chagas disease in Central America
- School feeding programs in the Caribbean and the Philippines
- Nutritional crops in Ethiopia

Each example is covered in Part II, with a brief summary provided in section 5 of this overview report.

### 3. Useful concepts in assessing impact

Impact evaluations aim to measure results along a pathway to impact and link them causally with an intervention. *Impact pathways* describe causal pathways showing the linkages between a sequence of steps in getting from activities to impact.

Issues of causality are central to impact evaluations, and credibly concluding on causal relations between an intervention and observed outcomes and impacts is the goal of evaluation designs for such evaluations. The causality issue is a

---

<sup>1</sup> IDRC has had a similar experience in trying to use the Canadian Academy of Health Sciences (CAHS) framework<sup>1</sup> with some of its health programmes (see Canadian Academy of Health Sciences 2009, Making an Impact: A preferred Framework and Indicators for Measuring Returns on Investment in Health Research).

significant challenge. It is realized that in the kinds of settings that IDRC works, there are usually many ‘causes’ behind an observed longer-term result. Mayne and Stern (2013) discuss causality in a research setting.

For some, impact evaluations are associated with experimental designs. However, this need not be the case, and for the most part, IDRC targeted impact evaluations would be looking at alternative approaches to assessing impact. There are many options for rigorous impact evaluations, as discussed by Stern, Stame, Mayne, Forss, Davies and Befani (2012).

The key concept needed is that of *contribution* rather than attribution. An IDRC intervention is expected to contribute to research and development outcomes and impacts, but is not the sole cause of those results. Rather it is an essential part of a package of causal factors that together bring about or contribute to the observed outcome. Neither in most cases, is it sensible to try to determine quantitatively how much of an observed change ‘is due to’ or can be attributed to the IDRC intervention per se. This attributing to an intervention is what experimental designs seek to do, and where such designs are feasible they can be used along with counterfactual causal reasoning. However, as noted, in most IDRC settings, this is not feasible, and the equally valid concept of a contributory cause is needed, i.e. the intervention contributes to observed impacts (Mayne 2012).

Related to these concepts, is the interpretation of ‘*making a difference*’. In counterfactual thinking, making a difference refers to what would have happened without the intervention. In contributory cause thinking, making a difference means contributing to a (positive) change on the impact of interest, such as people’s lives. This is how ‘making a difference’ is used here.

IDRC interventions are aimed in part at changing how people and institutions behave. In exploring impact evaluations at IDRC, it will be important to know if a pathway to impact can be developed, as the review of research impact models noted. But there is also a need to understand why the pathway is (or is not working; the assumptions behind the pathway model. This is where theories of change come in.

A *theory of change* (ToC) adds to an impact pathway by describing the salient causal assumptions behind the links in the pathway—what has to happen for the causal linkages to be realized. Theories of change are models of how change is expected to happen (*ex ante* case) or how change has happened (*ex post* case). *Causal link assumptions* are the salient events or conditions necessary or likely necessary for a particular causal link in a ToC to be realized; if the assumption doesn’t hold, then the expected effect from that link will not or will likely not occur.

Such impact pathways and the accompanying ToC will be needed if useful impact evaluations are to be undertaken. Annex B discusses models of theories of change.

Another pertinent concept is that of *scaling*. At one level, there may be improved wellbeing from the research outcomes for a targeted group of beneficiaries, such as farmers involved in the specific projects. But often impact is imagined at a much broader scale affecting a larger group of people, beyond original project sites. Scaling is an issue receiving considerable attention at IDRC, including in its Strategic Plan that talks about ‘large-scale impact’. IDRC has an ‘Scaling Science’ study which is uncovering pathways and principles for scaling research. Impact evaluations could be appropriate to evaluate not just the direct results of a project, but also the results of scaling processes supported within or beyond IDRC programming. Scaling, types of scaling and ToCs on scaling are discussed in Mayne and Johnson (2015), as well as in Annex B.

Numerous different and overlapping definitions of ‘results’, ‘outcomes’ and ‘impacts’ can be found in the literature and guidance on evaluation and monitoring. A glossary of terms is suggested in Annex A to be able to delineate and distinguish the effects of IDRC’s investments.

## 4. Evaluability Assessment for IDRC targeted Impact Evaluations

Evaluability assessments are studies to see the extent to which an intervention is ‘evaluable’ and how such an evaluation could be undertaken. They can serve a multiple of aims (Davies 2013; Milat et al. 2015; Greenhalgh et al. 2016):

1. to improve the intervention design
2. to inform the design of a monitoring system
3. to decide if a planned evaluation should take place, whether now or later
4. to inform the design of a planned evaluation.

For this study about targeted impact evaluations of research programming, it is the last two purposes that are relevant. IDRC would do evaluability assessments to inform:

- whether a targeted impact evaluation would be useful to conduct, and if so, when, and
- what design options are there for undertaking such IEs.

That is, the context is that over a number of years there has been IDRC investment in an set of research-based interventions. There is now, perhaps a number of years later, interest in looking past a focus on the extent to which the research has influenced policy or practice, to explore what impact has resulted



from the research outcomes in order to both learn about successful pathways to impact and to report on impact. The first step in to undertake an evaluability assessment.

To that end, Table 1 lists several issues that need to be explored, which are set out in more detail below. These can be considered as steps in doing the evaluation assessment. In most cases, however, there would be iterations among many of the steps. This framework was used in exploring designs for the four targeted impact evaluations identified earlier and reported on in Part 2.

**Table 1 Evaluability Assessment Issues**

1. **Impacts of interest.** What impact evaluation questions would be useful explore?
2. **Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?
3. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
4. **Measurement and data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?
5. **Evaluation designs.** What realistic evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
6. **Usefulness, timing and cost of an impact evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

1. **Impacts of interest.** What impact evaluation questions would be useful explore?
  - What is considered impact here? What level/extent of impact?
  - What specific interests does IDRC have about which impacts?
  - What level of evidence is seen as needed for these questions?

Clearly an early step would be to identify which impact questions were of interest to IDRC. This sounds straightforward, but may not be for several reasons:

- There may be a large number of impacts of interest, and, given the time and costs involved in gathering data, there may be a need to make decisions on which are of significant interest.
- Considerable care is needed in formulating meaningful impact questions.

Given the challenge of addressing causality in complex settings, it is important that the evaluation questions that an impact evaluation addresses are meaningful, well defined and realistic (Mayne forthcoming).

Some impact questions are too broad and general to adequately frame impact evaluations. (For instance, to what extent has the intervention met its objectives

and/or expected outcomes? What has happened as a result of the intervention? What would have happened without the intervention?) Table 2 suggests meaningful and useful impact questions.

The impact questions in Table 2 are discussed below. These are a menu of possible impact questions. In developing a design for a specific impact evaluation, a sub-set and/or more specific versions of these questions would be formulated.

**Q1. Impacts.** What impacts were realized or observed?

- To what extent were the expected impacts realized?
- What other related unintended impacts were observed?

#### **Table 2 Meaningful Impact Evaluation Questions**

**Q1. Impacts.** What impacts were realized?

**Q2. Contribution to impacts.** Did the intervention make a difference? How and why has the intervention made a difference?

**Q3. Other influences.** What other influences were at play?

**Q4. Sustainability.** Are the results sustainable?

**Q5. Generalizability.** Will the intervention work elsewhere?

**Q6. Likely impact.** Is it likely the intervention will make a difference in the future?

**Q2. Contribution to impact.** Did the intervention contribute to observed impacts; that is, did it **make a difference**, i.e.,

- Was the intervention a contributory cause to the observed impacts?
- What role did the intervention play?

How and why has the intervention made a difference?

- How and why have the observed impacts come about (or not)? What causal factors or mechanisms in what combination have resulted in the observed impacts? Why have the expected impacts not been realized?
- How has the intervention contributed to the various results along the pathway(s) to impact?
- For whom does the intervention work and not work, and why?<sup>2</sup>
- Where relevant, were the impacts not realized because of program design failure or implementation failure?
- How has implemented affected the success (or not) of the intervention?
- What were the contextual factors that contributed to the success (or not) of the intervention?

**Q3. Other influences.** What other influences outside the intervention made a difference in bring about the impacts?

- To what extent did other influences play a role in bringing about the impacts?

**Q4. Sustainability.** Are the results associated with the intervention sustainable?

---

<sup>2</sup> These are the realist evaluation questions (Westthorp 2014).

- If the intervention has ended, are the impacts still being realized? Why or why not? Are they likely to continue to be realized?
- If the intervention ends, will the conditions needed for impact continue?

**Q5. Generalization.** Will the intervention work elsewhere?

- In what contexts and circumstances will this intervention's theory of change work and not work, and why? Can this intervention be transferred elsewhere and scaled up?
- What generalisable lessons have we learned about how the intervention led to impact?

**Q6. Likely future impact.** Given current understanding about results to date, is it likely that the intervention will contribute to future, longer-term results?

- What are the apparent results-to-date?
- Was the intervention a contributory cause for those results?
- Is the theory of change still robust? That is, given the knowledge to date, is it likely that the future part of the theory of change will be realized?
- What can be done to enhance the likelihood of achieving future results?

These questions cover both an accountability perspective—in the sense of being accountable for knowing about the results of an intervention—and a learning perspective, but mainly the latter. Question 1 is the main accountability question, and the sub-question of role speaks to a learning priority. Question 5 could also be seen as having an accountability perspective but clearly has implications for learning. Question 2 is the main learning question, asking how and why and for whom the intervention is working. Questions 3 and 4 could have both accountability and learning implications, depending on the settings.

Theory-based approaches using theories of change for addressing these questions have been discussed extensively (Ling 2012; Stern et al. 2012; Institute of Development Studies 2013; Mayne, Stern and Douthwaite 2013; UNEG 2013; Copestake 2014; Ling 2014).

**2. Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?

- Has the pathway been discussed and agreed among stakeholders?
- Did the IDRC intervention have impact in mind when designed?
- Can pathways be identified that are clear, credible, coherent and feasible, in particular impact pathways in the sphere of interest?
- Did IDRC undertake any activities to help bring about the impact pathway in the sphere of interest?

In order to explore impact issues there needs to be plausible connections between the activities of the intervention and the impacts of interest. This is particularly the case for IDRC, where much of the focus has been on spheres of influence rather than spheres of interest—these farther out impacts may have been often

identified, but the pathways from influence to impact may not have been really spelled out in projects and evaluations.

In considering an impact evaluation, it would be useful to develop such pathways in a credible manner, get agreement on them and conclude if they are plausible, i.e., are they logical, are they consistent with prior knowledge and experience. Typically, impact pathways identify useful impact issues to explore, such as specific results that might need measuring and links in the pathway that are more contested and need attention in an evaluation. They can also suggest data that will be needed to conclude on the role the intervention had on the observed impacts.

3. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
  - Is it likely that IDRC funding made a difference, was a contributing factor?
  - With multiple partners working to achieve impact, what has been IDRC's relative role? A trigger to start change? A necessary component? A catalyst to speed up change?

The pathways from IDRC funded research to impact are usually quite long in terms of the causal steps involved and/or the time to impact. While the pathway may appear plausible, the influence of IDRC efforts may be at best quite marginal due to:

- the lengthy causal pathway,
- the modest IDRC involvement, and/or
- the numerous other actors involved.

A first step may be to try and understand just what sort of role IDRC played in the intervention, such as:

- a trigger to start a process of change,
- support, filling gaps where research is needed in a broader process of change, or
- a catalyst to bring about change in a more rapid and evidenced-based fashion than would otherwise have been the case.

Related to these considerations, is the need to ask if even showing that IDRC had an influence, perhaps with considerable effort, is that information something that IDRC could use in the future. This may be particularly the case, where IDRC has not engaged in any specific activities to influence the pathway from the sphere of influence to interest, and would not see undertaking any such efforts in the future. Nevertheless, even if no such activities are contemplated, it still might be useful to know if some features of aspects of the research products produced might have led to greater development impact.

4. **Measurement and data availability.** Are the impacts of interest measureable? What data is available on impact and on IDRC's contribution?
  - Is it feasible to assess or measure impact?

- Is there data already available on impact from prior studies and/or monitoring, perhaps from partners?
- Is the data of adequate quality?
- Are there prior evaluations and/or studies available with evidence on the influence the IDRC intervention had on policy or practice?
- Is there evidence available on the IDRC's contribution to impact?

Data are needed to address the impact issues of interest: (1) data measuring the impacts and (2) data for concluding on the causality issues. Initially, one would look to see if there is already available data that can be used, from:

- Existing monitoring and evaluation data that has been collected on the intervention and its effects, and/or
- Data that partners involved in the intervention have collected, and is available. Other partners may be involved in IDRC's sphere of interest and have been tracking effects. This is the case for the e-procurement project, where the InterAmerican Development Bank (IDB) was involved in funding the implementation of reformed policies and tracking effects in the different countries.

Of course, as will often be the case, available data may not be sufficient to credibly address the impact issues of interest, in which case the IDRC impact evaluation will have to collect new data. What data needs to be collected will depend on the evaluation design to be used.

5. **Evaluation designs.** What evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
- What new data on impact can be realistically gathered to strengthen an evaluation design? At what cost and in what timeframe?
  - Is more than one design needed to buttress weaknesses in an individual design?

Consideration of the above 4 elements (impacts of interest, pathways to impact, likely contribution to impact and data availability) are all input to deciding on evaluation designs. No attempt is made here to go through the various design options possible which are extensively discussed in the literature. In particular Stern et al. (2012) discuss non-experimental approaches to address impact.

One would normally expect that there would be multiple approaches used:

- Often comparisons among groups is possible, either comparison groups that have been part of the intervention or newly constructed groups. These can provide some level of evidence on the impacts of interest associated with the intervention.
- Theory of change approaches—such as quantitative comparative analysis, contribution analysis and process tracing—allow for a better

understanding of why impacts have been realized or not by setting out the more detailed pathway to impact, which can then be explored.

- Case studies explore in detail a specific case of impact within an intervention. Some of the examples explored during this study involve projects implemented in different locations, providing the opportunity for comparisons among the different cases/locations.

Criteria for evaluation designs could consider:

- If there is solid evidence from previous studies that IDRC funding did influence policy, practice, innovation, or capacity outcomes in some manner.
- If a ToC linking the outcomes to impact can be reconstructed. The 'robustness' along the pathway could indicate how far out the pathway to go.
- IDRC is seen to have played a role in getting from research outcomes to impact.
- Relevant data has been or can be gathered on the impacts of policy or practice change.
- Resources and time available.

6. **Usefulness, timing and cost of an impact evaluation.** Based on the above analysis, there is a need to conclude on whether or not it would be useful and worthwhile to undertake an impact evaluation, and if so, when.

- Are there significant impact issues of interest to IDRC?
- Is it likely that the pathway to impacts can be explored?
- Is it reasonable to assume that IDRC effort played a role in bringing about the impacts?
- Can the impacts of interest be measured?
- Are there data available?
- Are there practical evaluation designs that can be used?

If answers to these questions are generally positive, there still is the issue of whether it would be worthwhile to undertake an impact evaluation. As in all evaluations there is a trade-off between generating useful information and the cost and time of doing so. What is key here is to have good dialogue about how the results from an impact evaluation would be expected to be used within IDRC to improve the delivery of their programmes. The clearer the use and the more significant the use, the more one should be willing to pay for the evaluation.

There is no standard impact evaluation and hence no standard cost. Nevertheless, there are a number of ways that insights into the cost of an impact evaluation can be arrived at.

*Using a fixed percent of the intervention budget.* There has been an ongoing and long discussion in the evaluation community on whether the budget for an evaluation might be estimated using a fixed percent, such as 10%, of the intervention budget. But there is no consensus, with many thinking that 10% is unrealistically too high, with 2-3% perhaps a better figure. Thus, for example, 3% of a \$1M project would suggest an impact evaluation budget of around \$30,000. The problem with this thinking, of course, is that the cost of an impact evaluation will depend on the specific evaluation questions addressed, the methodologies used and the strength of the evidence needed.

*Estimating the number of days involved.* The typical IDRC targeted impact evaluation is likely to require field visits, so the days involved might be something like:

- Document review      5 days
- Field visits            14 days
- Analysis                5 days
- Reporting              7 days
- Total                    31 days @ \$1000 = \$30,000 + expenses (travel, etc.)

*Adjusting cost estimates based on the context:* Factors affecting the cost of a specific impact evaluation would include the following:

- *Time since project completion.* One may want to undertake an impact evaluation quite a few years after completion of the project(s) in order to allow enough time for impacts to have been realized. However, the project team may no longer be around making interviewing those involved difficult and setting up interviews with the current key stakeholders time consuming—this could be the case in the E-procurement example. If a related project is still ongoing, then this should be less of a problem, as in the Ethiopia example.
- *Other factors intervening.* The greater the time between the project completion and the impact evaluation, the greater the likelihood that other factors have likely influenced the current impacts. They will likely need identifying and assessing as to their contribution to the impacts measured, adding to costs.
- *Number of sites to visit.* If the project involved numerous sites, then the cost for site visits will be greater.
- *Prior data weak or not available.* There could be a challenge in finding prior data on completed projects and/or the data may not be that informative concerning current pathways to impact. Measurement costs for the impact evaluation would then likely be greater.
- *Amount of data to be gathered and assess.* Clearly the more data that needs to be collected and analysed, the greater the cost. IDRC would likely want to avoid focussing too much on ‘nice-to-know’ issues, and make sure

that it is well understood just how answers to impact issues explored will be used.

- *Time schedule for the impact evaluation.* If undertaking the impact evaluation is to be done with a short timetable, perhaps to be available for upcoming decisions, the cost of bringing together and managing a likely larger team, would be more costly than an impact evaluation with a less tight timetable.

In the end, any impact evaluation undertaken would be outsourced competitively, and hence accurate estimates of the cost would be obtained from the bids. Nevertheless, in setting out the terms of reference, the above considerations could provide IDRC with an idea of what to expect in terms of the cost of impact evaluations.

The bottom line is that there are significant costs associated with undertaking an impact evaluation. Low-balling the costs will likely lead to less credible findings.

## 6. Summary of the targeted impact evaluation examples

The approach for considering targeted impact evaluations outlined in this report was used to examine four examples of IDRC programming:

- E-procurement in Latin American and Caribbean countries
- Chagas disease in Central America
- School feeding programs in the Caribbean and the Philippines
- Child nutrition in Ethiopia

The reports for each are available. The examples demonstrate that impact evaluations are feasible with established methodologies, and they could explore a wide range of learning questions that could be useful to IDRC, its grantees, and other stakeholders. In summary:

*E-procurement in Latin American and Caribbean countries.* This targeted intervention represents long term investments by IDRC in e-procurement issues in Latin American and Caribbean countries which acted as a catalyst for advancing fair and more open e-procurement policies and practices in some 30 countries. Subsequent to IDRC's initial investments and support in a network of e-procurement officials, Inter-American Development Bank and the Organization of American States have come aboard as partners.

A targeted impact evaluation here could track impacts on small and medium enterprises, while exploring issues such as:



- To what extent and in what manner have the e-procurement policies and practices in countries been influenced by the IDRC investments?

The countries provide a basis for undertaking cases studies and a basis for comparison. Theories of change could be developed for the different countries sampled providing a basis for assessing the contributions made.

*Chagas disease in Central America.* IDRC invested in research on an ecohealth approach to combatting chagas disease. The success of this approach lead to its being scaled up in three Central American countries. Plans on now being discussed by key stakeholders for further adoption in Central America.

While the IDRC supported research already shows the efficacy of the prevention measures, impact evaluation issues of interest could include:

- To what extent have the changes in the houses, lifestyles and gender been sustained?
- To what extent has there been autonomous or independent scaling of the ecohealth approach beyond that supported through IDRC investments: other households, organizations and governments adopted and supported the approach?

An impact evaluation could undertake site visits where the approach was implemented to assess the extent and nature of sustainability. As well it could seek out cases where there has been induced scaling through discussion with officials, NGOs and observations in neighbouring sites. The links back to the IDRC investments would be strengthened by developing relevant theories of change, and using the site visits to confirm or revise the pathways to impact.

*School feeding programs in the Caribbean and the Philippines.* IDRC has invested in projects that involve school feeding programs in a number of countries. For this example, the investments of two unrelated projects in in two different regions were examined. The projects aimed at improving the nutritional status of children in schools through the introduction of more nutritious lunches coupled with nutritional education for parents. In the Philippines there is an ongoing follow on project scaling the approach to more regions.

Impact issues of interest were:

- To what extent and in what manner have the feeding programs, which rely on locally grown crops, been sustained?
- Given the impacts measured on nutritional status, what other factors in school feeding programs need to be addressed to ensure a more successful nutritional improvement?

The evaluation design could include revisiting the sites of the previous school feeding programs to explore the sustainability issues, and help build and test a

theory of change for the pathway to improved nutritional status in each case. The theories of change could be used to under some form of contribution analysis, with a special focus in other influences and factors at play.

*Nutritional crops in Ethiopia.* The example consists of a series of four projects aimed at identifying and making available more nutritious and productive crops in Southern Ethiopia, such as varieties of chickpeas and haricot beans, to reinforce nutrition-specific interventions such as complementary feeding. The technical aspects of the two early projects were judged successful, and the focus in the just completed and current projects is on scaling up.

Impact issues included:

- To what extent was the system developed to produce the more nutritious pulse crops sustainable?
- To what extent has the enhanced nutritional status of the communities involved been sustained?
- What were the drivers and barriers to the scaling undertaken?

Again a key element of the evaluation design would be visits to the sites involved or a sample of them, interviewing those involved in implementing the project and the beneficiaries involved. Comparisons among the sites should be possible. A solid theory of change would again strengthen the interviews tools and provide a basis for linking the interventions with the enhanced nutritional status. Some of the sites might lend themselves to a more in depth case study to explore the various impacts, such as sites that were particularly successful (or the opposite) and/or sites where the implementation team is readily accessible.

## 6. Concluding Comments

With an increased focus on development outcomes, IDRC is considering undertaking targeted impact evaluations of selected completed projects. Important shifts in the literature around impact evaluation, and an examination of four examples from IDRC programming show that impact evaluations would be feasible. Such evaluations could add value to IDRC's evaluation work. The main intent of conducting targeted impact evaluations would be for learning purposes to apply the greater understanding and insights to future programming. As described in the section on evaluability assessment, and as highlighted throughout the examples that follow, different types of learning agendas could be supported through impact evaluation including understanding:

- why expected impacts were not realized and what could be done differently next time,
- if impacts are not yet realized, the extent to which they are likely to be realized in the future,

- how and in what way the IDRC efforts contributed to the impacts that were realized,
- which groups benefited from the intervention, why and under what conditions,
- the extent to which the impacts realized are sustainable,
- how scaling up might best be undertaken

A secondary focus could also be to independently confirm prior research findings from the projects. A convenient way to capture this focus can be through an examination of the sustainability of the project(s), looking at whether or not earlier claimed benefits continue to be realized, and understanding why or why not.

Given the nature of IDRC projects, for targeted IEs, there would be only limited scope for experimental designs:

- Randomized control trials would not likely be possible, unless set up at the outset of the projects.
- Where baseline data is available, before and after designs might be feasible.
- If there were control sites or groups in the original projects, there would be the possibility for comparisons with the treatment sites or groups.

What would be possible in all cases would be forms of theory-based designs. These approaches are particularly well suited to complex interventions with multiple actors and long time frames, as is the case for research programs. Here a theory of change could be revised or reconstructed to serve as the basis for exploring a range of impact evaluation questions through site visits and key informant interviews, surveys, and focus groups, confirming or further revising the theories of change. Where practical, specific cases studies could be undertaken to explore issues in more detail.

## References

- Canadian Academy of Health Sciences (2009). *Making an Impact: A Preferred Framework and Indicators to Measure Returns on Investment in Health Research*. Available at [http://www.cahs-acss.ca/wp-content/uploads/2011/09/ROI\\_FullReport.pdf](http://www.cahs-acss.ca/wp-content/uploads/2011/09/ROI_FullReport.pdf).
- Copestake, J. (2014). "Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches." *Evaluation* 20(4): 412-427.
- Davies, R. (2013). *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations*, Working Paper 40: DFID. Available at [https://http://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/248656/wp40-planning-eval-assessments.pdf](https://http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf).
- Greenhalgh, T., J. Raftery, S. Hanney and M. Glover (2016). "Research impact: a narrative review." *BMC Medicine* 14(78). Available at <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-016-0620-8>
- Institute of Development Studies (2013). *Impact Innovation and Learning: Towards a Research and Practice Agenda for the future*, Report of an international Workshop, Brighton 26-27 March 2013. Available at <http://www.ids.ac.uk/files/dmfile/WebsiteReport.pdf>.
- Lindquist, E. (2001). *Discerning Policy Influence: Framework for a Strategic Evaluation of IDRC-Supported Research*: University of Victoria.
- Ling, T. (2012). "Evaluating complex and unfolding interventions in real time." *Evaluation* 18(1): 79-91.
- Ling, T. (2014). *Putting Impact Evaluation in its place*. RAND Europe. Available at <https://http://www.uea.ac.uk/documents/6347571/7859678/Documents/UEA+Impact+Evaluation.pdf/4b33bbce-65ed-4c58-b77e-e38f155b16d5>.
- Mayne, J. (2012). "Contribution Analysis: Coming of Age?" *Evaluation* 18(3): 270-280.
- Mayne, J. (2015). "Useful Theory of Change Models." *Canadian Journal of Program Evaluation* 30(2): 119-142.
- Mayne, J. (2016). *The COM-B Theory of Change Model: Working Paper*. Available at [https://http://www.researchgate.net/publication/301701597\\_The\\_Capabilities\\_Opportunities\\_and\\_Motivation\\_Behaviour-Based\\_Theory\\_of\\_Change\\_Model](https://http://www.researchgate.net/publication/301701597_The_Capabilities_Opportunities_and_Motivation_Behaviour-Based_Theory_of_Change_Model).
- Mayne, J. (forthcoming). Realistic Commissioning of Impact Evaluations: Getting What You Ask For? In *Evaluation and the Pursuit of Impact*. A. Paulson and M. Palenberg, Eds, Transaction.
- Mayne, J. and N. Johnson (2015). "Using Theories of Change in the Agriculture for Nutrition and Health CGIAR Research Program." *Evaluation* 21(4): 407-428.
- Mayne, J. and E. Stern (2013). *Impact evaluation of natural resource research programs: a broader view*, ACIAR Impact Assessment Series Report No. 84. Canberra: Australian Centre for International Agriculture Research (ACIAR). Available at <http://aciarc.gov.au/publication/iaso84>.

- Mayne, J., E. Stern and B. Douthwaite (2013). *Evaluating Natural Resource Management Programs*, Practice Brief, AAS-2013-23. Penang, Malaysia: CGIAR Research Program on Aquatic Agricultural Systems. Available at <http://aas.cgiar.org/sites/default/files/publications/files/AAS-2013-23.pdf>.
- Michie, S., M. M. v. Stralen and R. West (2011). "The behaviour change wheel: A new method for characterising and designing behaviour change interventions." *Implementation Science* 6(42): 11 pages. Available at <http://www.implementationscience.com/content/pdf/1748-5908-6-42.pdf>
- Milat, A. J., A. E. Bauman and S. Redman (2015). "A narrative review of research impact assessment models and methods." *Health Research Policy and Systems* 13(18). Available at <https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-015-0003-1>
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (2012). *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper, 38. London: DFID. Available at <http://r4d.dfid.gov.uk/Output/189575/>.
- UNEG (2013). *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*, Guidance Document. Available at [http://www.uneval.org/papersandpubs/documentdetail.jsp?doc\\_id=1433](http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=1433)
- Westthorp, G. (2014). *Realist Impact Evaluation: An Introduction*: Methods Lab. Available at <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf>.

## **Annex A: Glossary of Terms**

There is no consistency in the literature in the terms used to identify the various types of results that could be associated with an intervention. This seems also the case for IDRC. Yet an organization needs to agree terms so that different individuals and groups in the organization are talking about the same thing.

IDRC uses a number of useful terms that can be built on. In particular, to distinguish levels along the pathway to impacts, IDRC often uses the concept of different spheres:

*Sphere of control* – results such as research outputs over which the funded activities normally have reasonable control

*Sphere of influence* – results such as influence on policy over which the funded activities normally have influence

*Sphere of interest* – results such as the effects of implemented policies over which the funded activities normally has limited if any influence.

Another useful concept that IDRC uses is that of boundary partners coming out of outcome mapping:

*Boundary partners* - those individuals, groups, and organizations with whom the program interacts directly and with whom the program anticipates opportunities for influence. These would be first users and perhaps secondary users and could involve beneficiaries and/or intermediaries (see below for definitions).

*Intervention* - a set of specific activities undertaken and their subsequent effects aimed at making a positive difference. It covers policies, programs and/or projects.

*Results* - the changes in events, conditions and lives associated with an intervention, typically described as the outputs, outcomes and impacts. A *result statement* is the exact text used to describe the result.

*Research outputs* – direct products from a research project: knowledge communication (publications, reports, frameworks, workshops, models) and/or technologies/innovations/networks.

*Immediate/first users* –those targeted by the research who are expected to make use of the specific research outputs, such as other researchers, policy makers and/or beneficiaries, all of which are boundary partners.

*Reach and Reaction* – the first users who actually receive the research outputs and their initial reaction

*Beneficiaries* - populations who are expected to benefit from the research process, outputs and subsequent results; ‘beneficiaries’ can also include organizations and the physical environment.

*Enabling environment* - the context, policies, markets, actions that are needed to bring about the expected outcomes and impacts. An enabling environment can include needed policies and regulations, research support, relevant markets and actions by the private sector, and actions to ensure that other key stakeholders, research users or participants support the interventions aims and efforts.

*Intermediaries* - other people, organizations, networks, companies or institutions who need to act in order to bring about the desired changes and impacts for beneficiaries. These may include the first users.

*Research outcomes* - represent adoption and further use of research outputs by intermediaries, immediate (first) users and/or beneficiaries . They are generated as a result of the research, capacity building and advocacy activities of the intervention. Research outcomes are of two types:

- *Capacity changes*, the changes in capabilities (knowledge, attitudes and/or skills) opportunities and motivation, of the first users, beneficiaries and intermediaries. Note that ‘capacity’ here is a broader concept than ‘capacity’ that IDRC may use, and is based on behaviour change models, discussed below. It often refers to the capacity to use research in developing policy.
- *Behavioural changes* are the changes in actual practices that occur in the first users, beneficiaries and intermediaries; that is, beneficiaries and intermediaries do things differently than they were doing before.

*Secondary users* – users of the immediate research outcomes such as those who react to a policy change. These may be companies, individuals or other actors such as farmers or health care workers.

*Direct (or immediate) benefits* are the subsequent improvements resulting from the changes in practices in the status of beneficiaries or of the enabling environment, such as improved productivity when farm practices change.

*Improved wellbeing* - the longer-term improvement in wellbeing/livelihoods of beneficiaries or sustainability of the physical environment associated with the direct benefits.

*Development outcomes (Impacts)* — Direct benefits and/or improved wellbeing changes.

## Annex B: Theories of Change

Mayne (2015) discusses behaviour-based theories of change (ToC). Figure B1 illustrates such a generic ToC model. The specific terms are defined in Annex A.

**Figure B1: The COM-B Based Generic Theory of Change**

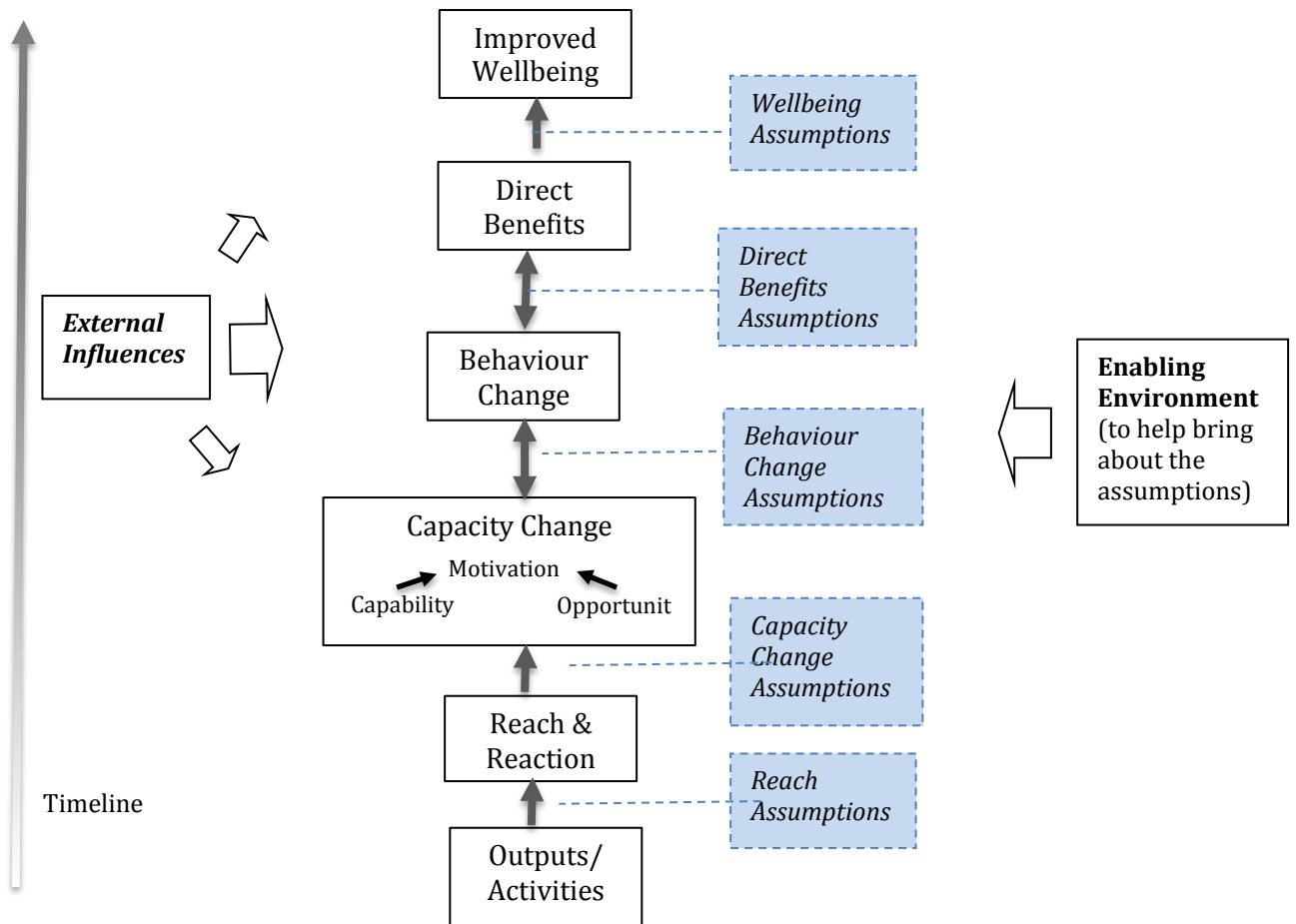


Figure 1 is based on the behaviour change model of Michie, Stralen and West (2011). This COM-B model argues that behaviour change comes about as the result of interaction between three *necessary* conditions, capabilities (C), opportunities (O) and motivation (M).

*Capability* is defined as the individual's psychological and physical capacity to engage in the activity concerned. It includes having the necessary knowledge and skills. *Motivation* is defined as all those brain processes that energize and direct behaviour, not just goals and conscious decision-making. It includes habitual processes, emotional responding, as well as analytical decision-making. *Opportunity* is defined as all the



The diagram illustrates the IDRC Development Outcomes framework, structured around three concentric circles representing different spheres of influence, and a timeline axis.

**Spheres of Influence:**

- Sphere of Interest (Outermost):** Contains the final outcomes: **Improved Wellbeing** and **Beneficiary Direct Benefits**.
- Sphere of Influence (Middle):** Contains intermediate outcomes: **Beneficiary Behaviour Change** and **Beneficiary Capacity Change**.
- Sphere of Control (Innermost):** Contains the foundational activities: **Research + Outputs** (solid box) and **Research + Outputs** (dashed box, optional).

**Flow and Relationships:**

- Foundational Activities:** **IDRC Supported Research + Activities** leads to **Research + Outputs** (solid box) and **Research + Outputs** (dashed box, optional).
- Reach:** **Research + Outputs** (solid box) leads to **First Users Reach** and **Beneficiary Reach**.
- Capacity Change:** **First Users Reach** leads to **First Users Capacity Change**, which leads to **First Users Behaviour Change**. **Beneficiary Reach** leads to **Beneficiary Capacity Change**, which leads to **Beneficiary Behaviour Change**.
- Enabling Environment:** **First Users Behaviour Change** leads to **Enabling Environment**.
- Outcomes:** **Enabling Environment** leads to **Improved Wellbeing**. **Beneficiary Behaviour Change** leads to **Beneficiary Direct Benefits**.
- Groupings:** **Improved Wellbeing** and **Beneficiary Direct Benefits** are grouped as **IDRC Development Outcomes**. **Beneficiary Behaviour Change** and **Beneficiary Capacity Change** are grouped as **Research Outcomes**.

**Legend:** A dashed box indicates an optional component.

**Timeline:** The vertical axis on the left indicates the progression of time from bottom to top.

As noted, a key focus of IDRC evaluations has been on exploring the extent to which the research has influenced policy or practice. While the focus of this paper is on the subsequent effects, this earlier part of the pathway also may be of interest. Figure B3 illustrates a generic ToC for influencing policy in which there is a direct connection between research processes and policy makers.

**Figure B3: A Model of Research Influence on Policy**

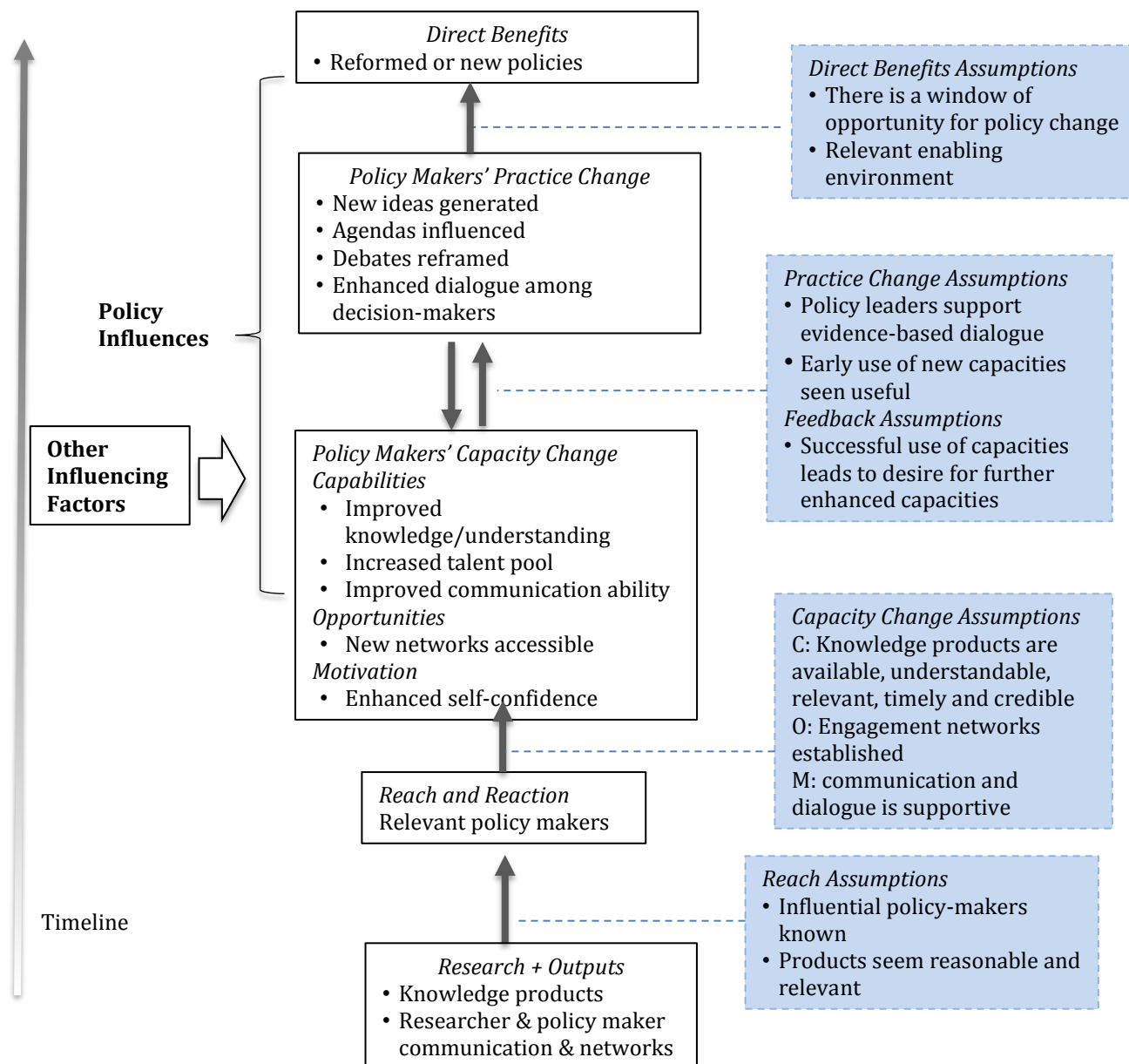
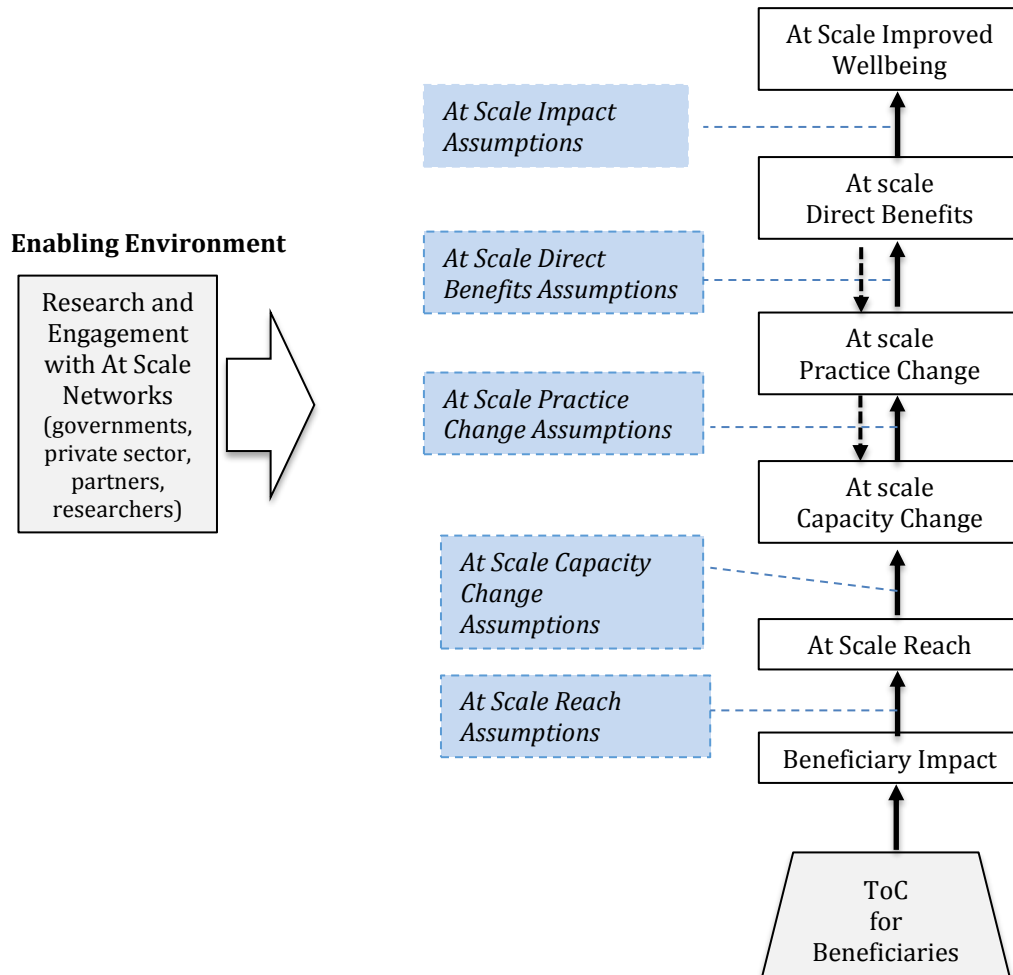


Figure B4 suggests one way to model scaling – a more detailed example of a ToC incorporating scaling is presented with the Chagas example.

**Figure B4: Generic IDRC Research Impact At Scale Theory of Change**



## Part 2: Project Examples

### Example 1: E-Procurement in Latin American and Caribbean Countries

#### Background

In the early 2000s, IDRC funded projects supporting the enhancement of e-government in the Latin American and Caribbean countries. During that work, e-procurement was identified as an issue worth looking at. IDRC help set up a network of e-procurement officials that grew over time from only a few countries to some 30 Latin American and Caribbean (LAC) countries. The network—Inter-American Network for Government Procurement (INGP)—identified numerous applied research issues which, with IDRC funding, were investigated.

The creation of the network was a forum for research priorities to emerge, for some research to happen, and considerable mutual learning and influence among network members to occur. Sometimes the research served more to keep an issue alive, on the agenda, and a matter of priority, rather than being an intervention to come up with specific policy remedies to adopt.

Initially, the Inter-American Development Bank was not involved, but over time they saw the value of what was happening and came on board. The Organization of American States supported a Secretariat for the INGP, while IDB supported national-level projects in which the good practices identified by the network were applied by national governments. In 2008, IDRC, with the Organization of American States (OAS) and IDB as partners, funded a project on “Strengthening procurement systems in Latin America and the Caribbean through the use of ICTs and the promotion of MSME participation” (105243; MSME means Micro, Small and Medium Enterprises). The rationale—the Narrative Theory of Change—for the project was:

Through the generation of relevant knowledge on ICT procurement, training of procurement officials and the development of networks and working groups of procurment officials for discussion and exchange of ideas, procurment in LAC countries would become more open and transparent, leading to greater access by SMEs. As a result, SMEs would become more competitive, win more contacts for procurement and become more successful.

Rather than the perhaps more typical IDRC-funded research projects, this IDRC programming was about agenda setting and field building, supported by user identified research projects. IDRC, as well, brought issues to the table, such pushing for opening internet-based transparent procurement schemes when previously procurement processes had been all closed office, paper-based.

## The example

There are 5 projects in this example. The first project is the main project around which we conceptualize an impact evaluation. The second is an ongoing project that follows on the main project. The last three were funded earlier and provide background and context to the main project.

- P1-105243: Strengthening procurement systems in Latin America and the Caribbean through the use of ICTs and the promotion of MSME participation. Start Date: March 2009, End Date: March 2013. [*The main project*]
- P2-107625: greening the economy through procurement. Approved Nov 2013, for 30 months. [*A follow-on project to the main project*]
- P3-103819 and P4-101929: broader network projects on e-government in Latin America. [*Prior broader context project showing that IDRC has invested in the broader question of e-governance for some time*]
- P5-104910: small 2-month project to bring in procurement officers to the research. [*Prior broader context project showing that more than just research was funded by IDRC*]

IDRC documents set out the basic elements of a theory of change (ToC):

### *Project outputs*

- The LAC e-procurement network — the INGP and working groups

### *Research outputs*

- Knowledge Communication – reports, papers, conference material, policy briefs, newspaper articles, participation in policy making meetings and events, appointments of young professionals, websites
- Training (for someone, some group or some institution)- procurement training, Ministries learning from young professionals, attendance at procurement events

### *Supporting actions (by the project, external actors, and contextual change)*

- involving officials in the project networks and working group
- IDB funded projects to implement reforms in procurement
- OAS funded the Technical Secretariat
- Ongoing government modernization

### *Reach*

- Heads of procurement units (policy makers)
- Procurement offices/professionals (policy makers)
- Young professionals/junior assistants in procurement

### *Research outcomes (Project outcomes)*

- Capacity changes
  - Policy makers: better understanding of potential role of ICTs
  - Policy makers: improved capacity of decision makers to appreciate and use research
- Practice changes
  - Universities: masters program and Ministries/Policy makers: junior assistants
  - Policy makers: (1) development of different policies for different countries, (2) use of greening criteria, (3) open contracting

### *Direct benefits*

- procurement policy reforms implemented
- enhanced SME access to public procurement

### *Impact*

- a more open and transparent procurement process
- more SMEs winning procurement contracts
- more sustainable procurement

This produces the draft ToC shown in Figure 1 below. Elements of this type of theory of change are discussed in the Main Report.

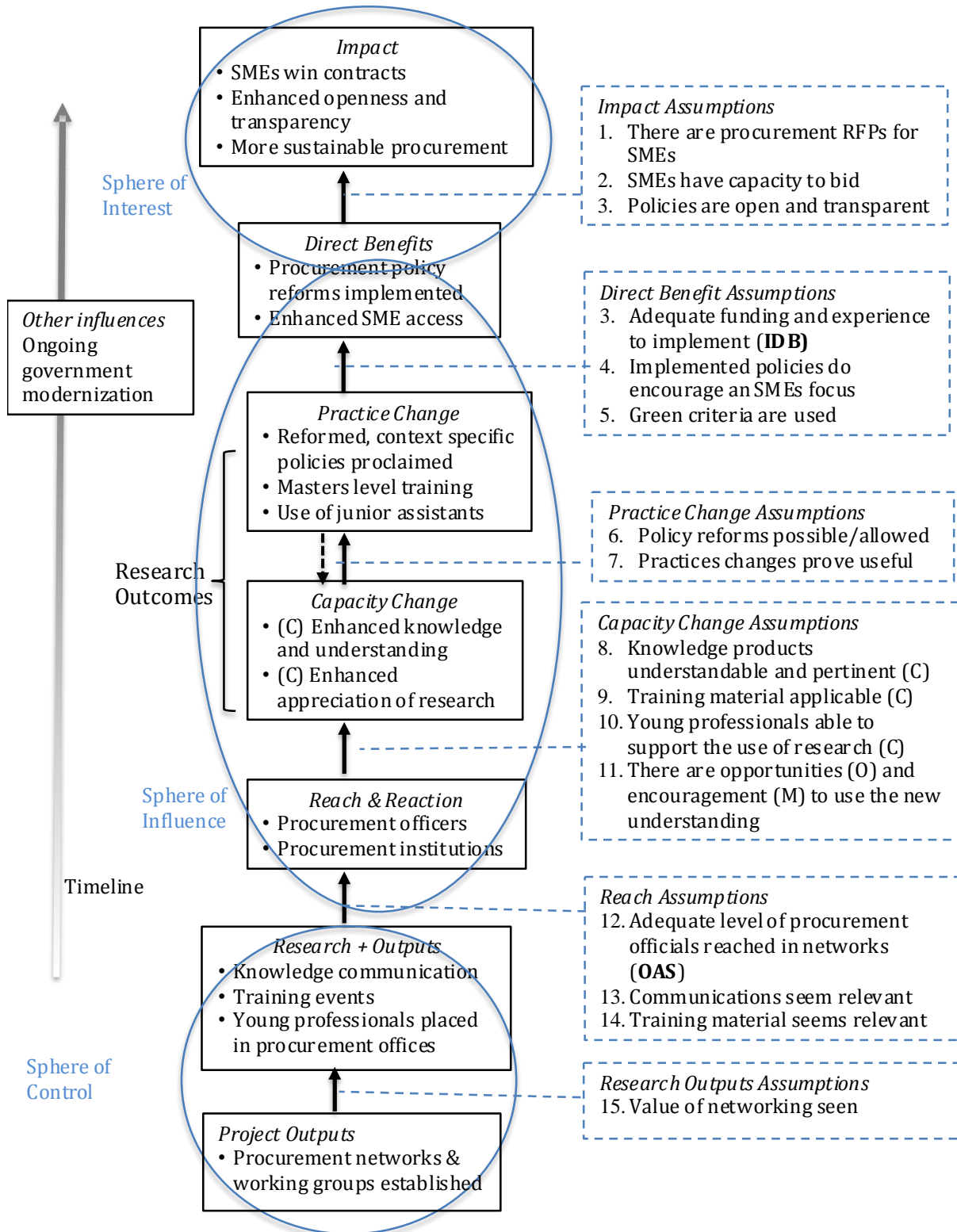
## **Evaluability Assessment**

Table 1 (from the Main Report) list the issues that need to be addressed in considering impact evaluation (IE).

**Table 1 Evaluability Assessment Issues**

7. **Impacts of interest.** What impact evaluation questions would be useful explore?
8. **Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?
9. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
10. **Measurement and data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?
11. **Evaluation designs.** What realistic evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
12. **Usefulness, timing and cost of an impact evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

**Figure 1: E-Procurement Project 1 Theory of Change**



## 1. Impacts of interest

Table 2 from the overview report lists the groups of useful impact questions to consider.

The overall impacts that the e-procurement projects aimed for are:

1. Better governance through more open and transparent public sector procurement process in LAC counties.

To evaluate this impact, indicators could be

developed for assessing the transparency of the procurement process in different countries. These might include such things as the percent of public procurement that was open to fair, evidence-based competition, the extent bids were restricted such as to national companies, and the extent to which the resulting contracts were made public. As well, of course, could be the impressions of those seeking procurement contracts as to the transparency and fairness of the process.

2. Greater involvement of SMEs in the procurement process, including SMEs receiving a higher portion on procurement contracts.

Indicators here could be such things as the percent of RFPs aimed specifically at SMEs, the percent of bids from SMEs and the percent of contracts won by SMEs. As well, SMEs could be interviewed on their involvement.

3. A 'greener' procurement process through the use of green' criteria in awarding contract that look beyond the lowest bid.

Indicators here might include the extent to which 'green' criteria were used in assessing bids and the percent of green bids that won contracts.

Sub-questions (from the Main IE report) are listed below, along with conclusions on the questions. For the purpose of illustrating the approach to evaluability assessment for impact evaluation, conclusions in *italics* assume what might be an IDRC position :

- What is considered impact here?
  - Impacts are listed above, with indicative measures.
- What level/extent of impact?

### Table 2 Meaningful Impact Evaluation Questions

- Q7. **Impacts.** What impacts were realized?
- Q8. **Contribution to impacts.** Did the intervention make a difference? How and why has the intervention made a difference?
- Q9. **Other influences.** What other influences were at play?
- Q10. **Sustainability.** Are the results sustainable?
- Q11. **Generalizability.** Will the intervention work elsewhere?
- Q12. **Likely impact.** Is it likely the intervention will make a difference in the future?



- *None of the impacts were set out as quantitative targets. Hence an IE could explore a variety of measures assessing the impacts. As there would likely not be a baseline, a challenge may be getting meaningful retrospective before measures.*
- What specific interests does IDRC have about which impact?
  - *There might be equal interest in each impact, recognizing that the green impact would be associated with a more recent project and hence might not be fully developed. The expectation would be that the e-procurement policy reforms in the LAC countries addressed the impact issues of interest.*
- What level of evidence is seen as needed for these questions?
  - *Acknowledging that these impacts are well into IDRC's sphere of interest, evidence that indicated IDRC's contributions, rather than completely robust evidence, would be likely adequate.*

**Impact evaluation questions** then could be:

*Impacts:*

- To what extent have these impacts been realized in the LAC countries, or in selected LAC countries? What were the key factors that brought about the impacts? What role did the INGP play? What were the key constraints?
- Why was there more achieved in some countries than in others?
- Have there been any other noticeable effects from the reformed e-procurement policies? These might be unintended negative effects such as increased costs due to increased involvement of SMEs, or greening or transparency enhancements spin-off effects on other e-services.

*Contribution to impacts*

- How did the IDRC intervention contribute to the impact?
- How far along the impact pathway can the influence of IDRC be shown?
- Where research knowledge clearly played a role in influencing policy and its implementation, what was it about the research knowledge and/or the context that led to its utilization?
- What roles did IDRC play in the realization of the impacts? What else might have been done to help bring about the impacts?
- What was it about the procurement networks and working groups that led to their success influencing the development of e-procurement in LAC?

*Other influences*

- Given that modernization was happening, what was the added value of the IDRC involvement?

*Given the context, it would not appear that there are particularly interesting impact issues in relation to sustainability, generalization, and the likely future impact. The e-procurement networks are set up to involve all LAC countries—so already is at scale—and continue to be active, with annual events and good*

practice awards—and hence it has proved to be sustainable. Future sustainability would seem to be quite far removed from the IDRC interventions, and dependent on political leadership and stability issues.

## 2. Pathways to impact

If the link between what IDRC funded and the impacts of interest is not really known, weak or involves many causal steps, this would suggest a careful reflection of whether an IE is useful at all. Setting out a plausible pathway to impact provides some assurance that further exploring IE options is worthwhile.

Figure 1 set out a possible expected pathway to impact. Here the sub-questions are:

- Has the pathway been discussed and agreed among stakeholders?
  - *Prior to an IE, the pathways would need to be discussed with stakeholders and likely revised.*
- Did the IDRC intervention have impact in mind when designed?
  - *Yes, at a general level. But the projects focused on IDRC's traditional sphere of control and to some extent the sphere influence, so the pathway needs some confirmation.*
- Can pathways be identified that are clear, credible, coherent and feasible, in particular impact pathways in the sphere of interest?
  - *The pathway set out seems reasonable, but it is recognized that there are numerous other influences at play that likely affected the development of e-procurement policies and their implementation, such as:*
    - i. In the LAC countries, modernization of government services generally have been and are going on, and would have gone on without IDRC involvement.
    - ii. Procurement by governments is often quite political, affected by the party in power and the historical context of procurement in each country.
    - iii. IDB has been funding aspects of the implementation of reformed procurement policies and hence likely has had some influence in the implementation processes.
  - *The implication is that the pathways in the sphere of interest in Figure 1 need to be further developed, identifying in more detail the risks and assumptions involved.*
- Did IDRC undertake any activities to help bring about the impact pathway in the sphere of interest?
  - *No, so defining the IDRC contribution may not be straightforward.*
- Does the pathway to impact involve scaling?
  - Scaling is not evident in Figure 4, but the pathway to impact involved from the outset a large number of LAC countries. Over

time the network grew from a few countries to 30. So scaling was part of the initial design of the intervention.

**3. Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the impacts?

This is perhaps the key impact issue for IDRC. An impact evaluation could assess whether IDRC indeed had an influence, show that and to better understand how any actual influence came about.

The sub-questions here are:

- Is it likely that IDRC funding made a difference, i.e., was a contributing factor?
- With multiple partners working to achieve impact, what has been IDRC's relative role? A trigger to start change? A necessary component? A catalyst to speed up change?

It is certainly plausible that IDRC funding influenced the impacts. The argument would be that IDRC at a minimum accelerated the modernizing of e-procurement and helped the modernizing of procurement policies be more evidence based. The impact benefits that have occurred, have done so earlier than they would otherwise have done. With a more solid evidence base, the reforms are also possibly more efficient and effective than they would otherwise have been.

Of course, while plausible, evidence supporting these contributions would have to come from an IE, such as through confirming pathways in the ToC for countries. Relevant officials could be asked about the influence the INGP had on policy and practice in their country in comparison to other drivers.

**4. Measurement and Data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?

- Is it feasible to assess or measure impact?
  - Yes, the impacts listed above could be measured. Some indicators were suggested.
- Is there data already available on impact from prior studies and/or monitoring, perhaps from partners?
  - While IDRC has not collected any impact data, there would appear to be some such data likely available. The 2008 Logical Framework Matrix suggests that some relevant impact data is being collected on the e-procurement policy reforms that have taken place in eight LAC countries. IDB looked at the impact of a project of SMEs may be collecting data on the implementation of reformed policies given that they are helping to fund such implementation. OAS may have

data on SME involvement and on the greening issue. Chile has some data from the 1900s, 2000s on the number and profile of SMEs that were providers for government contracts.

- *What impact data that has been or is being collected would have to be explored with IDB and OAS, the two partners, as well as national governments. It would be seem quite worthwhile for IDRC to explore data availability with partners prior to proceeding with an IE.*
- Is the data of adequate quality?
  - This would have to be explored with the partners.
- Are there prior evaluations and/or studies available with evidence on the influence the IDRC intervention had on policy or practice?
  - *Not sure but possibly.*
- Is there evidence available on the IDRC's contribution to impact?
  - *No*

**5. Evaluation designs.** What evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?

From the above analysis:

- A number of impact evaluations questions have been identified relating to measuring impact, the contribution to impact being made, and other key influences at work. The evaluation design would need to address these questions.
- A plausible pathway to impact has been identified, but there is a need to further develop the pathway in the sphere of interest.
- Existing impact data may be quite limited, but this needs to be explored with IDRC project partners.

This Example is one that involved a large number of LAC countries, and progress towards reformed e-procurement policies would differ in each country. On the one hand this provides possible **comparisons among countries** that could be useful in understanding how the drivers for change worked in different settings. But of course, collecting data in a number of different countries could be a challenge in terms of time, costs and compatibility. Nevertheless, comparisons between the more successful and less successful countries could be quite informative.

The 2013 *Technical Report* for Project 105243 written by OAS and reporting on the efforts of the INGP, identifies e-procurement policy changes in 8 LAC countries, noting that the changes are at least 'in part thanks to the contribution of the INGP'. By now there may be more such cases. This is an example of the need to explore in some detail with its partners just what has been done by others in tracking or examining impact. Answers could significantly affect the design of any IDRC impact evaluation.

*On the assumption that little has been done by others*, in order to address the impact questions of interest to IDRC, these **country cases**, or a sample of them, would need to be examined in some detail. Further examination of the country cases identified would suggest which cases to consider. And for some of the analysis suggested, it would be helpful to include in the sample, country cases where policy reform has been less successful.

One option then for an IE would be to undertake case studies on the selected country cases, using interviews with the stakeholders involved and observations to conclude on the impact evaluation questions, making use of comparisons among the cases as appropriate. Data collection here would involve one or more field visits to the countries in the sample. A limitation could be that some or many of those interviewed would not be aware of the role IDRC funding has played and hence not able to comment on if and how the funding has contributed to the impacts. This could limit the usefulness of the IE in terms of learning what aspects of the IDRC funding may have had a greater influence on the observed outcomes.

A stronger option would be to develop more robust theories of change for each country case identifying the key factors that played a role in bring about the impacts. If building ToCs were seen as part of the evaluation design, then the case studies could be used to test and refine the models. Indeed, if a reasonable number of country cases ToCs were built, there likely then would be a small number of types of ‘effective policy changes’ ToCs that could be identified, allowing for some generalization.

If the factors involved in such a ToC can be measured in a number of countries, then a possible design would be **Qualitative Comparative Analysis (QCA)**. QCA could be used to identify the key variables involved in bring about change, namely efficient and effective procurement policies.

Qualitative Comparative Analysis (QCA) is a social science research method that applies a systematic comparison to case study research. Its purpose is to refine and extend knowledge of the determinants of outcomes by looking at the similarities and differences of cases in terms of the causal factors and outcomes obtained ... . In the impact evaluation field, QCA helps to explore why some interventions were successful in achieving a particular outcome while others were not. It is suitable if an impact evaluation aims to use learning from existing cases to improve future interventions.

This would result in a stronger ToC that could then be the basis for exploring the causal links between IDRC’s efforts and the observed impacts, perhaps using then a combination of **contribution analysis and process tracing**. The end result would hopefully be a better understanding of IDRC’s contribution, such as:

- Which research or types of research were more useful in effectively influencing policy?
- Which types of communication and network building were more useful?

Since specific procurement policy reforms are not identified at the outset, the E-Procurement Example would seem to be a candidate for **outcome harvesting or outcome evidencing**, buttressed by the reconstructed ToCs. In either case, and somewhat similar to process tracing, the idea is for each of the country cases, to trace back from observed effects of the implemented policies to the intervention efforts. This should generate learnings about what worked and why.

The biggest challenge and cost here would likely be collecting data on the impacts of reformed procurement policies in different countries, and the views and observations of the various stakeholders involved in the changes.

At this point, it is not clear the extent to which the projects have influenced procurement policies in the various countries. Evaluations examining influence appear not to have been conducted. As a result, an IE here would likely have to include examination of changes in the sphere of influence as well as the sphere of interest. If IDRC funding has had little effect on the procurement policies that have occurred, then their influence on effects of the policy reforms is going to be small.

**6. Usefulness, Timing and Cost of an Impact Evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

The need for an IE here would need to be carefully considered. What would IDRC do with the findings of an IE in this case? Are there other similar type interventions—setting up networks to trigger improvements—that have been done, or are being contemplated?

Given the large number of countries involved, the costs for even a light IE could be significant, involving visits to a significant number of countries, and given the time that has evolved, setting up relevant interviews could be a challenge.

Several options for IE were discussed above. A key issue would be the extent to which IDRC wants or needs evidence on the contribution made. The simple country case studies could provide a variety of opinions from officials on the contribution made by the INGP and IDRC's role. This might be adequate. More solid evidence would involve more detailed case studies and some form of contribution analysis or outcome harvesting, perhaps aided by QCA. All this would add significantly to the costs.

In terms of timing, the more time that has elapsed, the more challenging would be the IE in terms of reaching those involved over the years and getting useful data from them.

**Concluding comments**

It certainly appears possible to undertake an IE for this Example and there could be useful learnings for IDRC as a result. Those involved in the projects argue that the IDRC involvement in this area has been quite successful. This may be so, and the issue is whether an IE would be useful to undertake to confirm these views and to provide more detailed insight into which efforts were more successful and in what contexts.

Given the involvement of other partners in the overall efforts in e-procurement, it could be useful to explore with them whether there is interest in conducting. One would want to explore:

- Have some of these issue been assessed by IDRC's partners in the e-procurement field?
- Would some of the partners be interested in exploring some of the impact issues jointly?

## Example 2: Chagas Disease in Latin America and the Caribbean

### Background

IDRC has invested in research on Chagas disease in LAC countries for more than a decade. IDRC's scaling study on the various Chagas projects, sets out the problems addressed:

Chagas is a vector borne disease mainly endemic to regions in 21 Latin American countries. The disease is caused by the *Trypanosoma cruzi* (*T. cruzi*) parasite, with its effects being potentially life-threatening (World Health Organization World Health Organization, 2016). The disease is common among rural and poor communities in Latin America and the Caribbean (LAC) affecting more than 10 million people. In this region it is estimated that over 10,000 people die annually result from the disease (IDRC, 2011, p.1).

Typically, the disease presents in two phases. The first is an acute phase, during which large numbers of parasites are present in the bloodstream. This phase lasts for approximately two months after initial infection. Most individuals suffer from mild symptoms or no symptoms at all during this phase. The second is a chronic phase, where lower levels of parasites congregate in the heart and the muscles of the digestive tract. During the chronic phase, patients can suffer from heart disease, as well as digestive and neurological disorders. This can lead to death as a result of heart failure or degradation of the nervous system (World Health Organization World Health Organization, 2016).

In the past in Central America, the *T. Cruzi* parasite was transmitted to humans via two vectors. The first was the *Rhodnius prolixus* insect species. This species was not native to the regions in which Chagas disease was endemic in Central America, and as a result, was eradicated through a successful domestic pesticide program. The second vector, the *Triatoma dimidiata* insect, is a native species to these regions in Central America. It can live in homes, peri-domiciliary environments near these homes, as well as forested regions. Since the elimination of *Rhodnius prolixus*, *Triatoma dimidiata* has become the principal vector for the transmission of Chagas disease in Central America (IDRC, 2011, p.23).

Since the elimination of *Rhodnius prolixus*, long term management of Chagas disease through domestic spraying campaigns has failed. This is because of the fact that *Triatoma dimidiata* is native to the endemic regions. Although domestic spraying campaigns can eliminate infestations temporarily, reinfestation typically occurs within a few months. *Triatoma dimidiata* can easily survive in both peri-domiciliary and forested regions, and will migrate back to homes once pesticide levels have subsided (IDRC, 2011, p.23).



This targeted impact evaluation example consists of three projects:

**101812 An Ecosystem Approach to Chagas Control in Guatemala. 2004-2007**

This was a key earlier intervention aimed at plastering walls in the homes in two communities in Guatemala in order to reduce infestation of the parasites biting occupants in their homes leading to Chagas disease. In addition, there were two control communities. The result was a significant reduction in *T. dimidiata* infestation.

**103696-005 An Ecohealth Approach to Communicable Diseases in LAC. 2009-2010**

A follow-up project focussed on changing the floors in homes to a type of cement, and moving farm animals out of the house by building chicken coops and the like. The result was further success in reducing *T. dimidiata* infestation.

**106531 Ecohealth Interventions for Chagas Disease Prevention in Central America. 2011-2014.**

This is the main project in the example. The project involved scaling up through replication and adaption of the previous successes in three countries: El Salvador, Honduras and Guatemala. The project made a significant contribution to the controlling infectious diseases and also played a key role in developing the field of ecohealth in Central America.

As the projects involved changing cultural norms and habits, in addition to the more technical aspects, considerable attention was paid to the related social development and governance issues.

**Understanding the intervention: the Theory of Change**

The initial project in Guatemala (**101812**) sets out the components and expectations of the intervention:

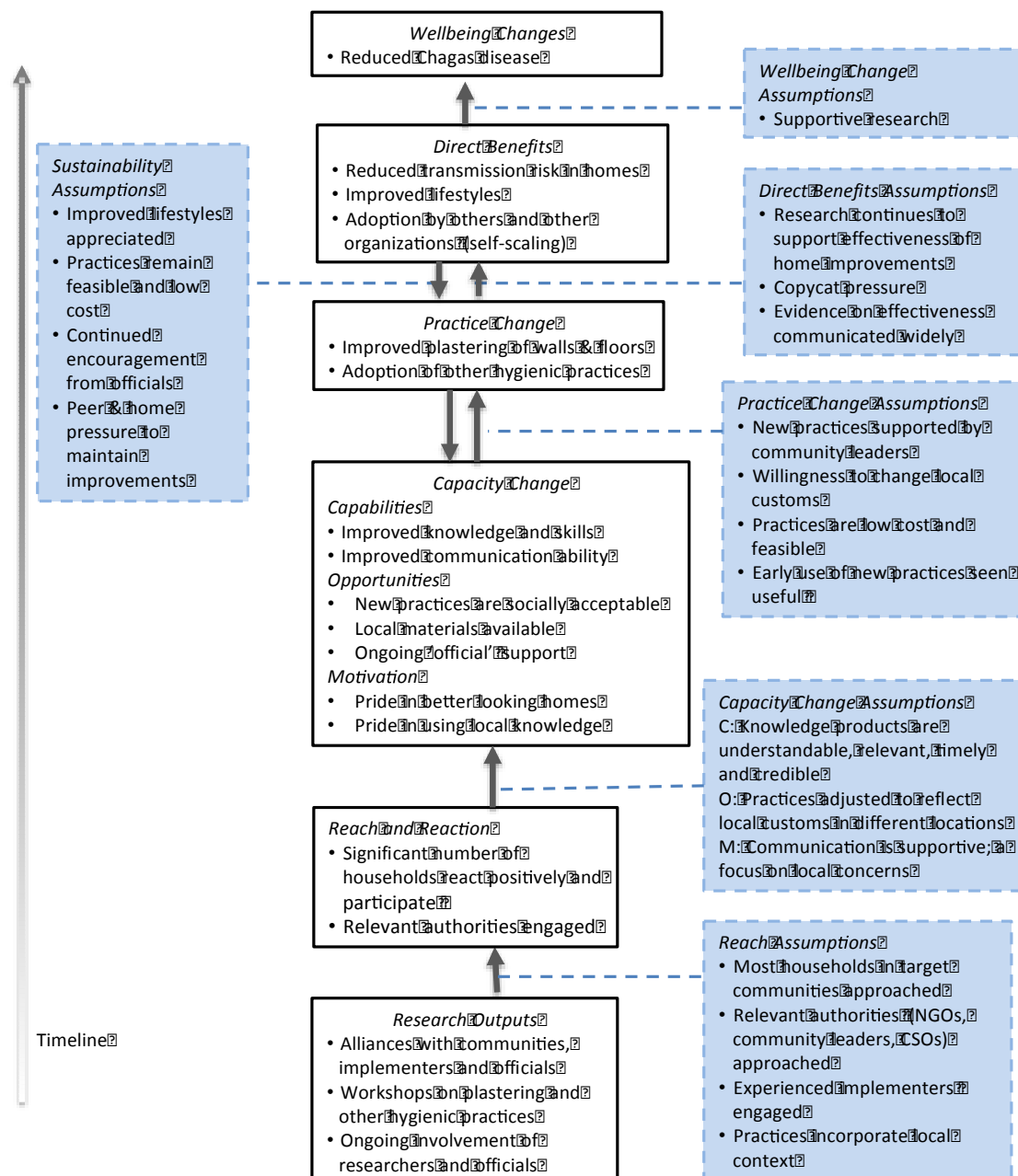
- Activities – workshops on plastering techniques, focus on other life-improvement activities to encourage involvement, ongoing advice and assistance by officials and researchers
- Reach – households in the treatment community
- Capacity change – Capability: knowledge and skills; Opportunity: local materials made available, local officials and trusted researchers provided support, practices are socially acceptable; Motivation: pride in better looking houses, local knowledge used, focus on local concerns
- Practice change – improved walls; some hygienic improvements
- Direct benefit – reduced risk within houses

- Wellbeing improvement – reduced infections (transmission) and improved lifestyles

This information along with similar information from the second ecohealth project (103696-005), can be used to build a theory of change for the basic intervention and the scaling-up version.

The theories of change (ToCs) used here are built on the COM-B behavioural change model (Michie, Stralen and West 2011), and discussed by Mayne (2016). The IE main report provides an overview. Figure 1 is an indicative ToC model for the scaling up Chagas project in the three countries.

Figure 1: A Model of the Chagas Prevention Intervention at Scale



The assumptions for each causal link aim to show all events and conditions that were seen as needed for the link to work. Also shown are what could be the assumptions for the sustainability of the improved direct benefits and wellbeing changes that were observed. They, of course, essentially repeat the assumptions on the right side of the figure. That is, the sustainability of the results depends on the sustainability over time of the collection of assumptions. However, a few additional assumptions are shown reflecting the scaling context.

If it were found that in fact the implementation was quite different in 1 or more of the new locations, then it might be worthwhile to build a more specific ToC to better reflect the local conditions dealt with.

In any case the ToC could provide a useful structured basis for examining the Chagas projects in more detail.

### Assessing Evaluability for Impact Evaluation

Table 1 (from the Main Report) list the issues that need to be addressed in considering impact evaluation (IE).

**Table 1 Evaluability Assessment Issues**

13. **Impacts of interest.** What impact evaluation questions would be useful explore?
14. **Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?
15. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
16. **Measurement and data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?
17. **Evaluation designs.** What realistic evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
18. **Usefulness, timing and cost of an impact evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

### Impacts of interest

Table 2, from the Main Report, lists several groups of potential impact questions.

While some impact evaluations of research programming might be needed because impact is unknown, in this case, the impacts of the research-based interventions were proven to reduce infestation rates in study communities, and that reduction is known to be a necessary conditions for reducing Chagas infection rates. So an impact evaluation isn't needed to demonstrate that the intervention is effective in real-life conditions To conclusively prove such a

reduction in chagas disease would require a 20-30 year cohort study, which does not seem like a worthwhile investment by IDRC, given the lengthy time period and the relatively large investment required compared to the investment made in the research projects.

The example also was testing an ecohealth approach to reducing infectious diseases. The approach proved successful and was able to be scaled up considerably through replication and local adaption. The approach was not without significant challenges involving building community engagement and trust, changing social customs and engaging local authorities for support. Numerous lessons have been learned, written up and published. So again this aspect of impact did not seem in need of much follow-up attention.

Other secondary expected benefits from the intervention in terms of improved lifestyles and enhanced gender equality have also been documented. But lifestyle and gender issues could be an area worth further investigation: What was the nature of the lifestyle and gender changes? How extensive were these changes and to what extent were they sustainable?

A perhaps unanticipated benefit noted in the reports is that in many cases, those seeing the improvements made in the houses of their neighbours have decided to improve their own homes. In addition, other NGOs have approached the research team in order to adopt the approach, and other municipalities have promoted the approach. However it is not clear from the reports the extent of this self-or induced-replication, but it would be part of the scaling up of the project. This could be an impact issue of some interest: induced-scaling.

#### **Table 2 Meaningful Impact Evaluation Questions**

- Q13. **Impacts.** What impacts were realized?
- Q14. **Contribution to impacts.** Did the intervention make a difference? How and why has the intervention made a difference?
- Q15. **Other influences.** What other influences were at play?
- Q16. **Sustainability.** Are the results sustainable?
- Q17. **Generalizability.** Will the intervention work elsewhere?
- Q18. **Likely impact.** Is it likely the intervention will make a difference in the future?

In discussions, there were, however, several other impact issues of interest:

**Implementation.** Implementing the scaled-up ecohealth project did meet some challenges in terms of the experience of those implementing the intervention in the different countries, and the need to adapt the intervention to the local contexts in each case. It might prove useful to document and assess the challenges and how they were addressed. Some of this has been done, but a more structured approach involving the three sites could be valuable.

The impact question would be:

- What lessons can be learned about implementing complex ecohealth intervention?

**Sustainability.** Of particular interest would be to know how sustainable the interventions have been.

- Have those who already have made the earlier home improvements aimed at preventing Chagas continued with the upkeep when walls, floors, and outside animal accommodation get old and deteriorate?
  - Have the cost for the improvements remained low and improvements quite feasible?
  - Has there been peer pressure to continue practices?
  - Has there been continued encouragement from officials?
  - To what extent have there been economic benefits from the animal management and empowerment of women?

These questions could be asked for both households in the treatment groups in all three projects, as well for those who on their own made similar household improvements.

One impact issue raised was the cost-effectiveness for households of the intervention. Initially, an annual cost was calculated, but what was the picture over time? This issue here is included in the sustainability issue, since cost would be a major factor in determining sustainability.

The earlier IDRC funded interventions had sustainability in mind, with the research focussing in part on local solutions to the technical challenges faced. This issue would address whether or not in practice over time, was the research adequate to ensure sustainability.

**Policy influence.** An additional issue mentioned related to the adoption of the home improvement practices into policy in the countries. There are discussions underway now to further scale up the approach to more countries. It could be useful to know what factors and conditions lead to the policy adoption in countries. Such questions as the following could be addressed: How were national level authorities engaged with? Who was involved? What was it about the intervention and its results that struck a chord? Etc.

So the impact-related questions of interest could be on:

1. Lifestyle and gender changes
2. Induced scaling
3. Implementation
4. Sustainability
5. Policy influence

## Pathways to impact/Likely contribution to impact

The pathways to impact (Figures 1) are reasonably clear. Furthermore, in this example, the link with and hence the contribution of the IDRC research is fairly straightforward, at least for the more direct impacts, namely induced scaling, implementation and sustainability.

Given the time lapse, consideration would have to be given to other influencing factors when look at the contribution made to lifestyle and gender issues, and policy influence.

## Measurement and Data availability

*Baseline data* should be available on which households were involved in the Chagas reduction efforts. It should also be possible to identify the self-scaling households through observation and/or discussion. Households involved in the control groups should also be known.

As noted in other IDRC IE examples, it would very important to seek out studies that may have been done by others that relate to the intervention. For example, in the interview with the P.O., there was reference to a study on economic impact by an ecohealth network, which while not focussed specifically on IDRC funding, would be worth tacking.

## Evaluation designs

To assess ***lifestyle and gender changes***, interviews would have to be held with a range of households who used the ecohealth approach exploring the nature and extent of the changes and why they have occurred. It would also be useful to hold interviews with households who have not adopted the approach to see if similar changes have occurred, and if so why. Such comparisons would help strengthen any conclusions about the contribution made by the intervention to these changes. This analysis would be strengthened with the development and use of a theory of change linking the changes to the intervention.

The evaluation design to assess ***sustainability*** would be quite straightforward. One could return to the household which made improvements, some now 9-10 years ago, others more recent, and observe if the improvements have been maintained. Interviews with the households and local officials would be able to answer a range of questions, such as:

- Why were improvements maintained (or not)? Improved lifestyles? Peer pressure? Local officials? Improved health? Etc.
- What were the challenges in maintaining the improvements? Effort? Costs? Etc.
- What would have helped in the maintenance? More help in identifying local materials? More support from officials? Etc.
- Were resulting lifestyle improvements ongoing?

Finding of problems, issues and what worked well in supporting sustainability would be useful in planning future ecohealth interventions.

To examine the extent of ***induced scaling***, One could do a visible check of households in the neighbourhoods of those in the treatment group to see the extent to which there was spillover effects from the project—*self-scaling*. Where improvements are seen, the households could be asked what motivated them to undertake the improvements. The improvements might be related to seeing neighbours the improvements in the treatment groups, or might be due to general improvement in people's lives. If the link with the project were substantial, these spillover effects would be a credible addition to the scaling up of the project. This understanding of the self-scaling phenomena would allow its support and encouragement to be part of any future scaling.

Similarly, one could interview the NGOs and municipalities that were interested in the ecohealth approach to assess the extent to which they followed through and adopted or supported the approach. Did they receive training in the approach? Were they then able to implement the approach and with what success? Where implemented, were these Chagas 'hot spots'?

Visits to the households in the control communities would be helpful to see if household improvements similar to those in the treatment group have occurred in any event, perhaps reflecting general improvements in people's lives.

The design for the ***implementation*** lessons is similarly straightforward: case studies at each location, discussing with officials, implementer and households the challenges as implementation progressed. The IE could be a vehicle to explore these issues in more detail than was done during the intervention, when some information on implementation challenges was produced.

Interviews with the relevant authorities could be undertaken to assess the extent to which the projects have indeed been a factor in adopting related chagas prevention ***policies***. Challenges here could be that officials involved may have moved on and could be difficult to trace.

### **Usefulness, Timing and Cost of an Impact Evaluation**

Discussion are underway to adopt the ecohealth approach to chagas prevention in more countries. It would seem at this stage useful, through an IE, as soon as practical to

- Understand better the lifestyle and gender issues induced by the intervention
- Understand well the implementation challenges to date
- Confirm the sustainability of the approach
- To know the extent of induced-scaling and how it could be enhanced

- To understand better how the intervention contributed to affecting policy in the countries

The main cost of the IE would be the field visits to the treatment and control sites in the three countries, or a sample of them. As it is only a couple of years since the scaling project was completed, one would expect those involved to be accessible. If the induced-scaling were explored this would add to the costs since the households would not be pre-defined. For a similar reason, exploring the induced lifestyle and gender changes would add costs.

### **Concluding comments**

The success of the scaling project argues for effort to assess the impacts identified above, especially since the ecohealth approach to chagas prevention is expected to be adopted in still other LAC countries and regions. The IDRC scaling project provides some insights into scaling implementation issues. A good understanding of the sustainability of the projects, especially the earlier pilot projects, could provide valuable information for further implementation of the ecohealth approach, as could a better understanding of the induced lifestyle and gender changes.

### **References**

- Mayne, J. (2015). "Useful Theory of Change Models." *Canadian Journal of Program Evaluation* 30(2): 119-142.
- Mayne, J. (2016). *The COM-B Theory of Change Model: Working Paper*. Available at [https://http://www.researchgate.net/publication/301701597\\_The\\_Capabilities\\_Opportunities\\_and\\_Motivation\\_Behaviour-Based\\_Theory\\_of\\_Change\\_Model](https://http://www.researchgate.net/publication/301701597_The_Capabilities_Opportunities_and_Motivation_Behaviour-Based_Theory_of_Change_Model).
- Michie, S., M. M. v. Stralen and R. West (2011). "The behaviour change wheel: A new method for characterising and designing behaviour change interventions." *Implementation Science* 6(42): 11 pages. Available at <http://www.implementationscience.com/content/pdf/1748-5908-6-42.pdf>



## **Example 3: School feeding programs in the Caribbean and the Philippines**

### **Background**

This example consists of two independent projects involving school lunch programs in two quite different locations: the Caribbean and the Philippines. The projects involved using locally grown food crops to provide children with nutritious and low cost lunches along with nutrition education for teachers and parents. In the Caribbean the schools used locally grown crops produced by farmers, while in the Philippines the schools operated school gardens. The particular interest in this example is the school feeding program in each case and the extent to which it was effective. School feeding programs are a widely used intervention to address nutritional issues.

The two cases are:

#### **Case 1**

#### **106525: Improving the Nutrition and Health of CARICOM Populations (CIFSRF)    2011 - 2014**

This is the main project in this case, and it was a single phase, 3-year project. The productivity of local crops was enhanced through a variety of research supported approaches, and these crops were used to build nutritious lunch programs in schools in St. Kitts and Nevis, Trinidad and Tobago with the aim of enhancing child nutritional status and reducing obesity, as well as providing added income to local farmers.

#### **107803: Communicating Research for Policy Influence: Food Security in Latin America and the Caribbean    2014**

This is a small Research Support Project that supported the communication plans and capacities of a couple of different projects, including the CARICOM project. The program invested a little more time, attention and resources to ensuring the CARICOM project had good plans for policy influence and communicating to relevant stakeholders who could use their research.

#### **Case 2**

#### **106921: Integrated Approach to Address Food and Nutrition Security in the Philippines    2012 - 2014**

This 3-year project aimed at improving malnutrition in school children. It integrated school gardens, food fortification and nutrition education in 27 school districts. The research also focused on identifying sustainable bio-intensive garden practices using local knowledge and plant varieties.

**108156: Improving Food and Nutrition Security in the Philippines through School Interventions** 2016 - 2018

This is a new, ongoing, scaling up project that continues the first. The same grantees are involved.

**Understanding the intervention**

Although implemented independently, both projects had similar aims in terms of addressing several local problems:

- Poor nutritional status of children in schools: malnutrition in the Philippines and child obesity in the Caribbean.
- Limited use of local food crops in school lunches
- Poor local farmer communities

The interventions aimed to provide research assistance to local farmers and schools to grow productive and nutritious local crops for the schools, which use the food products to provide a much enhanced lunch food menu. As well, parents and teachers are provided with training on the benefits for children of a more nutritional diet both at school and at home. Benefits are expected for the children in terms of better nutritional status, for farmers who would have an additional and ongoing source of income and schools with lower costs for school food programmes.

The results framework for the two projects is something like:

**Research outputs**

- Productive, viable and low cost food products grown locally (eco-friendly gardens),
- communication products (policy briefs, videos, blogs, outcome stories, journal articles)
- teaching material on nutrition for schools
- partnerships among communities, schools, farmers and policy officials

**Research outcomes and impacts**

- Farm to Fork model of the value chain, from production to consumption
- policy influence in understanding the need for relevant ministries to work together in the Caribbean project and authorities issuing of guidelines in the Philippines.
- improved school menus
- improved nutritional status of children
- increased awareness on nutrition among schools, households,
- increased adoption of healthy eating habits

- improved income for farmers

At the outset, it would be useful to understand the differences in the two interventions: in terms of what was accomplished and understanding the differences, and the different implementation challenges. Perhaps not a normal impact evaluation issue, but an issue that could provide valuable lessons for other similar school feeding interventions.

## Assessing evaluability for impact evaluation

Table 1 (from the Overview Report) list the issues that need to be addressed in considering impact evaluation (IE). Each is explored below.

**Table 1 Evaluability Assessment Issues**

19. **Impacts of interest.** What impact evaluation questions would be useful explore?
20. **Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?
21. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
22. **Measurement and Data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?
23. **Evaluation designs.** What realistic evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
24. **Usefulness, timing and cost of an impact evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

### 1. Impacts of interest

Table 2 (from the Overview report) lists the types of impact issues to consider in undertaking an IE.

The expected impacts from the interventions can be grouped as follows:

**Nutritional status and obesity.** Have the nutritional status of school children improved? Have malnutrition and obesity rates improved?

- Have children consumed the more nutritional lunches? How and why (or why not) has this happened? What conditions needed to be in place?

- Have households been providing more healthy meals in keeping with the nutrition training they received? How and why has this happened? What conditions needed to be in place?
- Have malnutrition and obesity rates dropped? How and why (or why not) has this happened? What conditions needed to be in place?

### **Table 2 Meaningful Impact Evaluation Questions**

- Q1. **Impacts.** What impacts were realized? Is more evidence needed?
- Q2. **Contribution to impacts.** Did the intervention make a difference? How and why has the intervention made a difference?
- Q3. **Other influences.** What other influences were at play?
- Q4. **Sustainability.** Are the results sustainable?
- Q5. **Generalizability.** Will the intervention work elsewhere?
- Q6. **Likely future impact.** Is it likely the intervention will make a difference in the future?

### **Local productive**

**nutritional crops.** Have farmers/schools been able to provide locally grown nutritious food products to the schools? How and why has this happened? What conditions needed to be in place?

- As a result, has the income of the farmers increased? How and why has this happened? What conditions needed to be in place?
- Are the school gardens in the Philippines sustainable? How and why has this happened? What conditions needed to be in place?

**Policy influence and research uptake.** Have the results from the interventions been able to influence policy with respect to school feeding programmes in the area and nationally? How and why has this happened? What conditions needed to be in place? What did this contribute to the results? Have the research findings been picked up by others? How and why has this happened? Did the communication strategy help?

Except for the issue of policy influence, it would appear from the FTRs of the pilot projects that there has been evidence gathered on these impact questions that could be compared between the cases. Extensive data was collected on the nutritional and obesity issues as well as on the productivity of the local crops grown. Another comparison that could be made on the nutrition issue could be with the literature on improving child nutrition. Are the case findings consistent with the literature? Are there other obvious things that could have been done to enhance nutrition status?

### *Contribution to impacts/Other influences*

There is a fairly clear link between these pilot interventions and most of the expected impacts. That is, the local crops were produced and used in the school lunches which were consumed. Positive impacts would most likely be the result of the interventions, unless there were other related interventions going on at the same time. Lack of positive results such as no reduction in obesity would suggest

that other factors such as availability of sugary drinks were contributing to the obesity levels that were not adequately influenced by the school lunches and nutritional education provided.

The intervention's contribution to any observed changes in farmers' income and related policy is plausible, but would need to be confirmed. There may be changes in markets, other events or other interventions going on.

And there likely would be other influences at play, such as peer pressure among the children, influence from TV, lack of adequate exercise, radio and other advertisements on food products, price rises in farmers' products and unrelated policy initiatives. Some of these influences might explain observed results or less than expected results in some locations.

One or more detailed ToC could be quite useful here to help understand the context and the causal mechanisms at work, and conclude on how and where the intervention made a difference, and what might have been done to strengthen the observed results.

### *Sustainability*

This would seem to be a key issue, and the sustainability challenge was noted in the Philippines FTR. Issues would be:

What happened to the programmes after the initial funding stopped?

- Did local farms continue to supply schools?
- Did the school gardens continue to use the new technologies?
- Did the school lunch programs continue?
- Did nutritional training in schools continue? Did home dietary practices improve?
- Did nutritional status of children continue to improve?
- Did obesity rates eventually drop?

And for each of these questions, the follow up questions would be:

- Why and how did the result continue?
- What were the drivers and barriers—the other contributing factors?

This is where the benefit from an IE could be realized. Understanding what else might be needed to improve nutritional status, over and above the school feeding programs, in a particular setting. This could include looking at the availability of affordable nutritional foods in local markets, exercise, non-healthy food available in schools, whether or not those getting the nutritional training decide what food to buy and who gets what in the household, health status and health services availability, etc.

It is not clear the extent to which sustainability was addressed at the outset; that is, what features of the intervention were put in place to maximize the chance that benefits from the intervention would continue.

### *Generalizability*

In the Philippines, the follow-on project which has just started is a scaling up project, seeking to expand the use of the approaches elsewhere, based on the lesson learned from the prior project. Thus it is directly testing the extent to which the intervention can work elsewhere.

### *Likely future impact*

Results from these interventions were expected to be realized over the 3 years of the projects. However, in the Caribbean project it was noted that the time period was too short for obesity rates to fall. Hence the issue of likely impact becomes in essence the sustainability issue: will the results be sustained or be realized later?

Summing up, the *impact questions of interest* for an impact evaluation would seem to be:

1. To what extent and in what manner did the projects influence policy?
2. Did the interventions make a difference, especially the school feeding programs? How and in what manner?
3. What other influences also played a role?
4. Were the observed impacts sustained?
5. To what extent will the follow on project in the Philippines be able to scale up the intervention and impacts to a larger region, i.e., different locations?

## **2/3. Pathway to impact/Likely contribution to impact**

As noted, for the most part causality is not a main issue: local farmers and schools are assisted in growing local crops which are used in the schools to provide more nutritious and low cost lunches. Nutritional training is provided to mothers and teachers. The intended results are then expected to be realized. It sounds like a logical and reasonable pathway, and could be modelled in the form of one or more ToCs. There could be nested ToCs for the actual lunch programs and one for the nutritional training component.

It seems quite likely that improvements in nutritional status, and increased income for farmers selling to schools, would be due to the interventions.

## **4. Measurement and Data Availability**

Nutritional status for malnutrition and obesity of children have standard metrics that are used, as have productivity and income. There are also metrics for assessing policy influence that IDRC has used.

In both locations, extensive baseline, monitoring and end of project measurements have been taken for a number of key variables relating to child diets and farmers' crops. However, the measures used were not the same for the two cases, in part since the nutritional problems were different, undernutrition in the Philippines and obesity in the Caribbean.

## 5. Evaluation design

The questions of interest discussed earlier were identified as:

1. What implementation lessons can be learned when the two cases are compared?
2. To what extent did the projects influence policy?
3. Did the interventions make a difference in nutritional status of the school children? How in what manner? What other influences also played a role?
4. Were the observed impacts sustained?
5. To what extent will the follow on project in the Philippines be able to scale up the impacts?

**1. Lessons.** Despite being independent the two cases had similar aims and similar intervention theories of change. But they had different settings, used somewhat different approaches, and used different measures for key variables of interest. Comparing successes and challenges would be useful, as well as comparing with the large literature on enhancing child nutrition.

**2. Policy influence.** To address the policy influence question, the pathway to policy influence ToC could be developed and explored with those officials involved using key informant interviews.

**3. Explaining how.** To better understand just how and why the results were realized (or not), it could be useful to undertake some form of contribution analysis based on a ToC developed for each intervention. The focus would be on the context and assumptions in place that explain what happened, and other factors of influence, with data gathered from the various stakeholders.

**4. Sustainability.** This is an issue that seems worthy of being explored, by revisiting the sites to observe what is now taking place and seek explanations from those involved for why. While one hopes for long-term sustainability, looking even a few years after the project termination would give useful insight into the sustainability issues.

**5. Scaling up.** The evaluation of the follow on project in the Philippines should be able to address this issue.

Evaluation questions #2 and #3 could be addressed at the same time through field visits with the data collection being informed by and used to verify and/or

revise ToCs developed for each project. The policy influence question would also require field visits but would involve those country officials responsible for health matters.

Sorting out the relative impacts of the feeding program and the training programs would likely require a quite elaborate design, perhaps comparing with schools with no such programs or only the lunch program (if there are any) and/or looking in detail at the food-lives of quite a few kids over a lengthy time period, hoping there is variation in the adoption of the training so some comparisons could be made. The value of doing this, however, may be limited; both components are part of the nutritional causal package and are needed to make a difference.

It would also appear useful to check the results obtained against the literature on improving the nutritional status of children, including through school feeding programs.

In addition, the evaluation could explore the project design and implementation issues for each project, looking for lessons on designing and implementing these types of projects.

## **6. Usefulness, timing and cost of an impact Evaluation**

Sustainability is an important issue and it would be quite important to know the challenges that have arisen in this regard. If an IE were undertaken soon, insights on sustainability—what might have been done differently to enhance sustainability—could be used to strengthen the currently underway Phase 2 project in the Philippines.

The costs involved would be site visits to the locations in the CARICOM countries and the Philippines interviewing the schools administrators, local farmers and local officials.

## **Concluding comments**

A useful impact evaluation of these projects would not appear to require an elaborate design. This is largely due to the fact that the key impacts of the projects were addressed in the FTRs, although checking and comparing results would be valuable. This issue of interest are understanding better just how and why the observed results came about and the sustainability of those results.



## Example 4: Nutritional Crops in Ethiopia

### Background

This final example consists of a series of four projects with the same grantees, three of which are completed and the fourth is underway. The projects were aimed at identifying and making available more nutritious and productive crops in Southern Ethiopia, such as varieties of chickpeas and haricot beans. The projects are very close together in time, with one phase picking up where the previous one left off. The technical aspects of the two early projects were judged successful, and the focus in the just completed and current projects is on scaling up.

The four projects are:

**#1. 10798 Scaling Up Pulse Innovations for Food and Nutrition Security in Southern Ethiopia (CIFSRF Phase 2) 2015 - 2018**

Ethiopia has one of the highest prevalence of protein-calorie malnutrition and micronutrient deficiencies in the world. This project will scale up new pulse food products (chickpeas and beans, for example) in southern Ethiopia to address food insecurity and malnutrition.

**#2. 107540 Promoting Adoption of Chickpea Technologies in Southern Ethiopia (CIFSRF) 2013 - 2014**

This project's design was to improve the understanding of the key drivers and processes that will facilitate and accelerate large scale adoption of chickpea production technology in the southern highlands of Ethiopia. It was designed as a pilot learning and action research on best practices and approaches to scale up, i.e. how to bring more benefit to large number of farmers from successful research results.

**#3. 106927 Improving Nutrition in Ethiopia through Plant Breeding and Soil Management (CIFSRF) 2012-2014**

This project illustrated how a nutrition-sensitive agriculture interventions such as pulse production can be a platform to reinforce nutrition-specific interventions such as complementary feeding.

**#4. 106305 Improving Food Security in the Highlands of Ethiopia through Improved and Sustainable Agricultural Productivity and Human Nutrition (CIFSRF) 2010-2013**

## Understanding the intervention

The key project in this example is #2, the 1-year just completed (107540) project and its predecessor # 3 also just completed (106927). Project #2 tested scaling up of the new technologies, with Project #1 now implementing the approaches on a broader scale over three years.

*Project #2 General Objective:* To understand and promote the key drivers and processes that will facilitate and accelerate large-scale adoption of chickpea production technology in the southern highlands of Ethiopia.

Projects #3 and #4 provided the scientific evidence for proceeding with the initiative.

From the IDRC documentation, the components and identified results of the Project # 2 intervention were:

### *Activities*

- Participatory seed selection among the farmers, followed by
- Seed planting, harvesting, and distribution

### *Research Outputs*

- Chickpea production with the identified selected varietal seeds
- A Guide for Chickpea production
- Success stories communicated
- Geographic Information Systems platform and maps

### *Technical research findings*

- Preferred varietal of chickpeas in terms of productivity
- Community-based production of seeds
- Optimal fertilizer trials
- Training
- Need for research on scaling

### *Expected Outcomes/Impacts*

- Significant seed production
- Increased income
- Partnerships among the actors involved in production and distribution of the seeds and chickpeas
- Increased consumption of chickpeas
- New skills of farmers
- Increased involvement of women in pulse production
- Increased understanding of the factors affecting scaling
- Policy awareness by officials of potential of chickpea production
- Significant increase in number of farmers using the varieties

- A foundation laid for sustainable impact.
- Increased understanding of impacts at scale

The technical aspects of the intervention were well documented in terms of production, productivity, and use. On scaling, a review of the relevant literature on scaling agrifood products in the region was undertaken.

Project #3 had established the gains made in nutritional status through the consumption of the pulse crops.

## Assessing Evaluability for Impact Evaluation

Table 1 (from the Overview Report) lists the issues that need to be addressed in considering impact evaluation (IE). Each is explored below.

**Table 1 Evaluability Assessment Issues**

25. **Impacts of interest.** What impact evaluation questions would be useful explore?
26. **Pathways to impact.** For the intervention, were there or are there plausible pathways to impact?
27. **Likely contribution to impact.** Is it reasonable to assume that IDRC has influenced the observed impacts?
28. **Measurement and data availability.** Are there reasonable measures for the impacts of interest? What data is available on impact and on IDRC's contribution?
29. **Evaluation designs.** What realistic evaluation designs are available for addressing the questions of interest and for concluding on the intervention's contribution to impact?
30. **Usefulness, timing and cost of an impact evaluation.** Would it be worthwhile for IDRC to undertake an impact evaluation, and if so, when?

## *Impacts of interest*

Table 2 (from the Overview report) lists the types of impact issues to consider in undertaking an IE.

### *Impacts*

The expected impacts from the project #2 can be grouped as follows:

**Table 2 Meaningful Impact Evaluation Questions**

1. **Impacts.** What impacts were realized?
2. **Contribution to impacts.** Did the intervention make a difference? How and why has the intervention made a difference?
3. **Other influences.** What other influences were at play?
4. **Sustainability.** Are the results sustainable?
5. **Generalizability.** Will the intervention work elsewhere?
6. **Likely impact.** Is it likely the intervention will make a difference in the future?

1. Commercially viable **production of nutritious pulses**. This includes issues of
  - adoption by farmers
  - productivity of the varieties,
  - production, distribution and sale of the seeds, and
  - increased income for farmers.
2. **Increased consumption** of the nutritious pulse crops, and possibly increased nutrition status.
3. **Enhanced empowerment** of women working on the pulse crops
4. Possible **unintended impacts** from the project
5. Increased understanding of **scaling** the technology and approach
6. **Increased awareness by authorities** of the potential and value of these pulse technologies

The project technical reports provide some information on many of these impacts along the pulse crop value chain. It is not clear if unintended impacts and increased awareness by authorities were explored and are probably worth addressing.

While the focus of Project #2 was on increasing consumption of the pulse, Project # 3 had established the link to enhanced nutritional levels.

The current project that is underway involves a number of sites in different locals. It could be useful to compare the contexts of these different sites, as well as compare with the previous sites. This information could later help explain differences in impacts that emerge from the different sites.

#### *Contribution to impacts/Other influences*

The project was in essence a pilot directly involved in delivering many of the impacts, starting with the two prior projects exploring the agronomics of the pulse crops and hence in these cases, the intervention clearly made a difference. An issue that might explored is to see if there any other influences at play in bringing about the results.

#### *Sustainability*

This would seem to be an important issue to address in the intervention. Now, several years later, what has happened in the pilot sites in terms of continuing production and distribution? Have the multi-partner platforms that were established been sustained? Have the income and nutritional benefits continued? And why have things continued or not? Further, one could explore if the conditions for sustainability are likely to continue.

A confounding factor, however, could be that the pilot site(s) were also sites in the current expanded project. However, there appears to be some scope for comparison among communities. In some areas, the project only provided

agricultural interventions about pulses, including maybe providing some pulse seeds, but there was no further ongoing engagement. Those communities could be compared with others, in which the project had ongoing connection.

If so, it might not be possible to credibly assess their sustainability as part of the just completed project. If the sites are different, then returning to the pilot sites to see what is happening would be quite valuable.

### *Generalizability*

The main aim of the project was to “To understand and promote the key drivers and processes that will facilitate and accelerate large-scale adoption of chickpea production technology in the southern highlands of Ethiopia”. Scaling lessons were drawn from previous projects in the region and incorporated in the project design, but evidence how the scaling actually worked is being addressed in the follow-up scaling project.

This is a clear issue that could be addressed in an impact evaluation, going back and eliciting the lessons learned from the attempt at scaling. It would also be useful to determine just why the scoping study was not completed.

Addressing this question could be challenging if the current project is using the same sites.

### *Likely impact*

The current project of further scaling up of the pulse technologies is underway. Impacts from the project are 3-4 years away. An impact evaluation now could act as an interim evaluation, reviewing the implementation to date and taking into account the lessons learned on scaling and impact from the previous project. It could then conclude on the likelihood of achieving the intended results from this project.

### ***Pathways to impact/Likely contribution to impact***

As noted above, causality seems not to be a key concern for many of the benefits from these projects. The pathways to impact are reasonably clear, and the results noted directly linked to the activities of the pilot project. The FTR pointed to three pathways to impact (theories of change) which could be developed: market oriented production, food security and market orientation, and food and nutrition security. The likelihood that the project contributed to the observed results is high.

### ***Measurement and Data Availability***

In general, data availability here would seem to be adequate. Baseline and final measures were taken on many variables of interest, and visits to the sites and

those involved quite feasible to undertake. Pulse crop production was measured as well as consumption. The increase in the numbers of women involved in pulse production was measured, as well as their role in decision-making. Studies were done on the challenges faced by women in this sector.

### ***Evaluation designs***

The key relevant evaluation questions would seem to be:

1. Sustainability
  - Has the introduction and production of pulse crops continued?
  - Have livelihoods improved, including income?
  - Has the increased roles for women continued?
  - Has nutritional status improvement continued?
  - Will the conditions for sustainability continue?
2. Unintended impacts and other influences at play
3. The awareness of authorities
4. Scaling
  - What lessons were learned about the scaling?
  - What were the drivers and processes facilitating the scaling?

Questions 1, 2 and 4 could be addressed through field visits to the project sites, and addressed at the same time. The evidence could be strengthened through comparisons among sites where the implementation was different. Question 3 would require interviews with the relevant authorities. In all cases, the questions would be explored as to why things have continued or why they have not, identifying drivers and barriers.

Thus, given the solid evidence-based findings on impacts and their direct causal links to the project, the impact evaluation design here need not be very elaborate, and fairly straightforward, involving

- Review of project documents on impact
- Visits to the various project implementation sites, observing and interviewing the actors involved along the value chain: farmers, households, extension workers, private sector players, implementers, researchers, and officials. Data would be gathered on the identified impacts of interest, especially sustainability, such as:
  - What has happened and is happening along the pulse value chain in terms of impact and why? Are any unintended impacts noted?
  - To what extent are previous changes being sustained, and why or why not?
  - What gaps or constraints, if any, were or are there in the value chain for the pulse crops? Were they considered in the project design?

- Has the involvement of women increased? What changes, if any, have or are occurring in relation to the women involved?
- What other influences were or are at work that might explain at least in part the observed results?
- The lessons from the scaling in the just completed project could also be explored through interviews with those who implemented the project.
- Focus groups across and among actors could be a good way to elicit a discussion on changes that have occurred and an explanation for the changes.
- It would be useful to explore if there might be reasonably accessible comparison communities or households that could be consulted on key issues. While causality would not appear to be a key concern, an exploration of such comparisons would provide supporting evidence on the efficacy of the intervention. Issues explored in interviews and/or focus groups could ask about adoption of the improved pulse varieties, consumption of the more nutritious pulse crops, the role of women in pulse production, etc.

Some of the sites might lend themselves to a more in depth case study to explore the various impacts, such as sites that were particularly successful (or the opposite) and/or sites where the implementation team is readily accessible.

This design could be strengthened by developing a solid theory of change (ToC) for the intervention, or more likely several ToCs, and using the ToCs as the basis for data collection and analysis. One ToC could be for the generic intervention: producing the nutritious and productive pulse seeds, which are adopted by farmers, leading to increased income, consumption of more nutritious food and greater empowerment of women in the process. A second ToC could be on the scaling efforts, building the partnerships for making the pulse crops commercially viable in more districts in Southern Ethiopia. The ToCs could serve as a basis for the interview tools, with the interviews eliciting the underlying ToCs of the stakeholders, including how the different elements of the ToCs interact. The interviews could also gauge the importance of the contextual factors and assumptions needed for the intervention to work.

Confirming and/or revising these ToCs based on this evidence would provide a solid basis for understanding why the approaches continue to work (or not).

The one complication would be if the follow-on project (**107984**) was using the same sites as the just completed project. Then disentangling the effects of the two projects might be difficult.

### ***Usefulness, timing and cost of an impact evaluation***

An impact evaluation undertaken in the near future on this example is quite feasible and could serve two purposes:

1. Confirm impacts from the just completed project (**107540**) and especially sustainability of the project results, as well as identify lessons in the scaling up efforts.
2. Act as an interim evaluation for the currently underway project (**107984**), checking on progress to date and comparing contexts among the new and previous sites.

The cost of such an evaluation should be quite reasonable, involving a field-visit to the pilot project sites, interviewing a sample of those involved, and observing the value chain in operation.

### **Concluding comments**

This example involved a pilot project that was well monitored with regard to many of the impacts of interest, which were undoubtedly the result of the intervention. Lacking was the expected study on scaling. As a result, the key impact evaluation questions relate to the scaling question as well as to sustainability of the results that were observed. Both of these issues can be credibly addressed through field visits to the project sites and interviews with the stakeholders involved.

The advantage of doing such an impact evaluation soon would be the lessons learned for use in the current larger scaling up project.