OTT Consulting

www.onthinktanks.org/ott-consulting

EVALUATION OF THE INTERNATIONAL DEVELOPMENT RESEARCH CENTRE'S STRATEGY TO SCALE RESEARCH RESULTS

—

2021

IDRC · CRDI

Canada

scaling science

Southern Hemisphere

SUPPORTING MEANINGFUL CHANGE

ABOUT THIS REPORT

This report presents the findings, analysis and insights from an evaluation of the International Development Research Centre's (IDRC) 2015–2020 strategic objective to invest in 'knowledge and innovation for large-scale positive change.' The evaluation was undertaken by OTT Consulting in partnership with Southern Hemisphere. It is accompanied by four case studies that explore a different thematic area related to scaling research results impact at IDRC.

ACKNOWLEDGEMENTS

The evaluation team leader was Simon Hearn, OTT Consulting Associate and the evaluation technical lead was Dena Lomofsky, OTT Consulting Associate and Partner and Senior Consultant at Southern Hemisphere. The accompanying evaluation case studies were produced by Cathy Chames (Southern Hemisphere), Ajoy Datta (OTT Consulting), Nana Davies and Tracey Phillips (Southern Hemisphere) and Dena Lomofsky. The evaluation team are grateful to Amy Etherington, Leslie Fierro, Robert McLean, John Gargani and Enrique Mendizabal for their support and feedback during the evaluation, and to all IDRC staff, grantees and representatives of other organisations who participated in the evaluation through interviews, surveys, focus group and sensemaking meetings. In particular to the IDRC evaluation advisory group members and the Policy and Evaluation division for their guidance, active participation in consultation meetings, and reviews of draft reports and case studies. The evaluation communications were led by Louise Ball (OTT Consulting) with editorial support from Sally Burt and Sophie Gillespie. Illustrations were designed by Magda Castría.

CONTENTS

TABLES

BOXES

FIGURES

ABBREVIATIONS AND ACRONYMS

AFS	Agriculture and Food Security
CAD	Canadian dollars
CARIAA	Collaborative Adaptation Research Initiative in Africa and Asia
CGIAR	Consultative Group on International Agricultural Research
CIFSRF	Canadian International Food Security Research Fund
FEH	Food Environment and Health
GrOW	Growth and Economic Opportunities for Women
HPRO	Health Policy Research Organisations
IDRC	International Development Research Centre
IMCHA	Innovating for Maternal and Child Health in Africa
IP	intellectual property
KIX	Knowledge and Innovation Exchange
LAC	Latin America and Caribbean
LVIF	Livestock Vaccine Innovation Fund
M&E	monitoring and evaluation
OD4D	Open Data for Development
PAD	project approval document
PCR	project completion report
PFAN	Private Finance Advisory Network
POEV	Policy and Evaluation Division
PRISE	Pathways to Resilience in Semi-arid Economies
ROSIE	Research on Scaling the Impact of Innovations in Education
SDG	Sustainable Development Goals
TPD@Scale	Teacher Professional Development at Scale

GLOSSARY OF TERMS

Scaling objective	The first strategic objective in IDRC's Strategic Plan 2015–2020 was to 'Invest in knowledge and innovation for large-scale positive change'.
Scaling Science	Scaling Science is the name of an initiative set up by IDRC's Policy and Evaluation Division during the strategic period. It intentionally has a double meaning: 'Scaling scientific research results to achieve impacts that matter' and the 'development of a systematic principle-based science of scaling' (McLean & Gargani, 2019, pp. 6–7).
Scale/ Scaling	Used without descriptors 'up', 'out' or 'deep' as any strategic intent to scale research results or an innovation.
Scaling impact	IDRC defines scaling impact as 'a coordinated effort to achieve a collection of impacts at optimal scale that occurs if it is both morally justified and warranted by the dynamic evaluation of evidence' (McLean & Gargani, 2019, p. 9). It is distinct from scaling 'up' or 'out' because the focus is on scaling the *impact*, rather than the action or activity itself. IDRC also note that 'scaling can supplement knowledge translation by pushing researchers to consider the benefits of knowledge and innovation beyond the uptake and application by immediate users' (McLean & Gargani, 2019, p. 15).
Guiding principles for scaling impact	IDRC has identified four guiding principles for scaling impact: justification, optimal scale, coordination and dynamic evaluation (McLean & Gargani, 2019). Justification: scaling is a choice that must be justified based on a balance of technical evidence and values and the agreement to scale should be shared by key stakeholders. Optimal scale: recognising that scaling produces a collection of impacts, and to determine optimal scale we must consider the trade-offs between those impacts. For example, how much impact, sustainability of impact, variety of impact and equity of impact. Coordination: scaling impact happens in complex systems. Multiple stakeholders are involved in the process of scaling. 'Their efforts may be cooperative, competitive, or complementary, and their roles may change over time. Consequently, coordination demands a high level of

planning, adaptation, and flexibility, along with a deep understanding of the system we place our innovations into' (McLean, Gargani & Lomofsky, 2020).

Dynamic evaluation: since scaling happens in a complex system, each action could initiate multiple intended and unintended actions requiring adjustment and adaptation. 'Dynamic evaluation is concerned with how, why, for whom, and under what conditions changing actions changes impact' (McLean, Gargani & Lomofsky, 2020).

Scaling pathway

A conceptual framework for scaling that emerged during this evaluation (see Figure 1). It comprises two interconnected pathways: a policy pathway and an innovation pathway, joined via an emerging third pathway related to system strengthening. The policy and innovation pathways relate directly to the corporate high-level scaling indicators adopted by IDRC and they reflect the way many programs reported progress against the scaling objective.

EXECUTIVE SUMMARY

The first objective in IDRC's 2015–2020 strategic plan was to 'invest in knowledge and innovation for large scale positive change'. Over the 5-year strategic period, senior leadership, programs and projects have grappled with the notions of 'scale' and 'scaling', conceptualising and approaching it in different ways and to different extents.

As the Centre enters a new 10-year strategic period, it remains committed to learning about scaling and sharing that learning with others. As part of this continued commitment, IDRC commissioned OTT Consulting, in partnership with Southern Hemisphere, to undertake an evaluation of IDRC's implementation of its strategic objective to scale and what was achieved by those efforts.

There was no standardised guidance or framework for implementation of the strategic objective across the organisation when the strategy was initiated, thus the evaluation is exploratory in nature, identifying the approaches, practices, systems and processes that supported or hindered scaling efforts. This evaluation took the form of a strategic review employing a mixed-methods design. Our approach to the evaluation was informed by utilisation-focused evaluation (Patton, 2008).

The evaluation included four interconnected components: (1) An organisational review of systems and processes, strengths and weaknesses, and outcomes achieved. (2) A study of grantee perceptions and experiences. (3) A study of IDRC's positioning within the wider research for development sector. (4) Four case studies targeting themes identified with IDRC staff: field building for scale, programming for scale, private sector engagement and scaling, and organisational learning about scaling.

In total, the evaluation included 88 semi-structured interviews with IDRC staff, grantees and informants from other organisations, plus five focus group discussions with 18 grantees from four regions. The evaluation team disseminated three online surveys and received responses from 43 IDRC program staff, 95 grantees and 16 funding organisations.

Through the course of the evaluation, a conceptual framework for scaling has emerged, which we refer to as the scaling pathway (see Figure 1). The scaling pathway comprises two interconnected pathways: one focused on policy and the other on innovation, joined via an emerging third pathway related to system strengthening. The policy and innovation pathways relate directly to the corporate scaling indicators adopted by IDRC and they reflect the way many programs reported progress against the scaling objective. The scaling pathway helps clarify the distinction between the supply and demand sides of scaling. The supply side (left hand side) refers to generation and mobilisation of

knowledge and innovation, while the demand side refers to use of the knowledge and innovation to support development outcomes at optimal scale. We also make a distinction between policy change or adoption of an innovation with primary intended users and policy change or adoption of an innovation beyond primary intended users at optimal scale (i.e., policy or innovation outcomes achieved through scaling the results achieved with primary intended users).

While similar to the result frameworks developed by IDRC programs, called 'program impact pathways', which define the expected research and development outcomes of activities and investments, the scaling pathway differs as it is not a logic model in the sense that projects are expected to traverse from left to right. Projects can have different entry points and exit points along the scaling pathway and programs can invest at different points in the scaling pathway simultaneously.

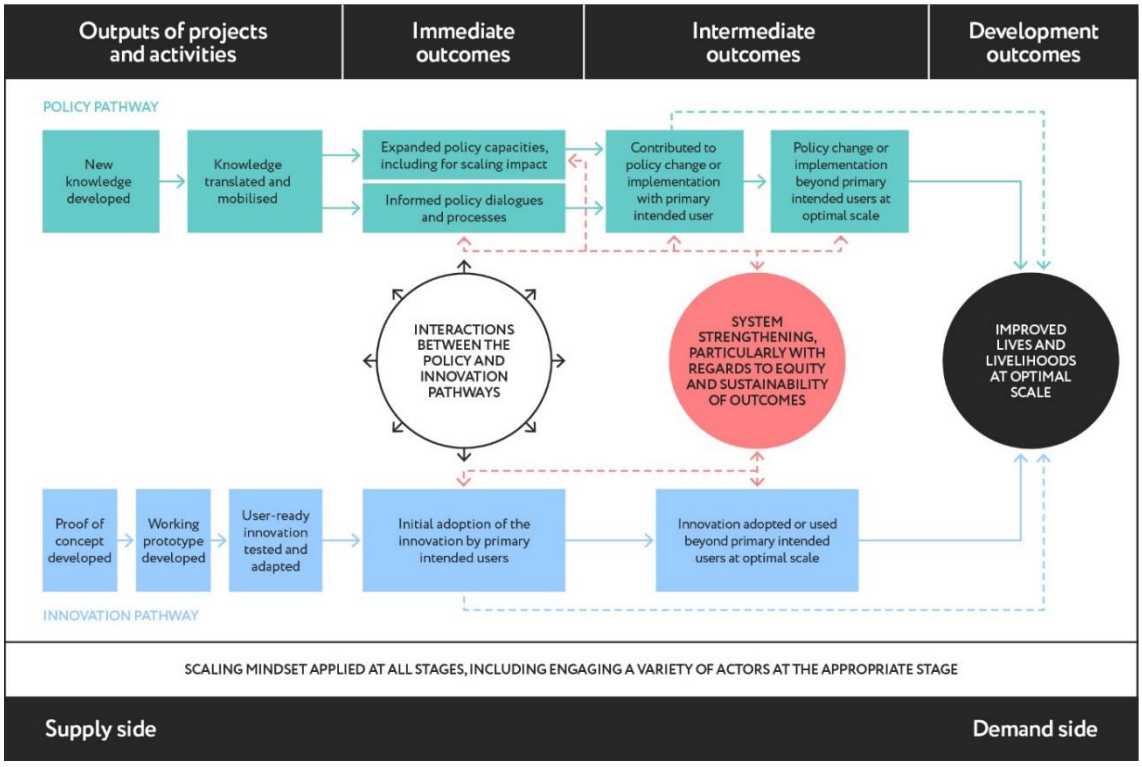

Figure 1: Scaling pathway (see accompanying text in section 1.3)

Outcomes achieved through the scaling objective

The evaluation included a systematic assessment of outcomes achieved through integrating scaling strategies into IDRC programs. The assessment identified outcomes using the scaling pathway as a framework, focusing on three types of policy outcomes and two types of innovation outcomes. Through an assessment of a diverse range of sources including program evaluations, program area progress reports, Trackify (a data

management tool for tracking evidence for program and corporate indicators) and surveys and interviews with staff and grantees, we identified 440 outcomes linked to scaling. This included 32 innovations being used beyond primary intended users and 170 contributions to policy change and implementation.

A notable example of an innovation being used beyond primary intended users is the Expanding Adoption of Nutritious, Disease-Resistant Potatoes in Colombia project which reached 6.5 million people by 2018 with three new varieties of potatoes. The potato varieties were developed with rural communities in an earlier phase of research and released to 650 potato producers for cropping. During this project, the focus was on the demand side of scaling – the project worked with seed entrepreneurs on sustainable business models, strengthening the capacity of mostly women smallholder farmers and improving nutrition awareness through direct household interventions and policy influence at village, municipal and national levels. The project's success was attributed to the combination of these diverse strategies.

An example of a contribution to policy change highlighted in this evaluation is The Role of the Private Sector in Reducing Corruption in Latin America project, which resulted in a new law on corporate anticorruption in Argentina. The project worked on the supply side of scaling in developing the evidence and drafting the initial bill, and on the demand side by working with a variety of actors in the system to ensure the law would have the intended effect. They worked with lawmakers in the lead up to the law's promulgation, and with the private sector to support their understanding of the new law and to strengthen the capacity of the prosecutors and defenders who would be responsible for implementing the legislation.

Overall, the outcome assessment showed a considerable result for IDRC given that the original target for the scaling objective was to support 'at least 20 initiatives that deliver solutions at scale'.[1] We expect there to be more outcomes than this given that the assessment undertaken for this evaluation was not exhaustive, did not cover all programs evenly and some of the sources used were 2–3 years old. While innovation outcomes could easily be identified as relating to scaling, this was not the case for policy outcomes. Few of the policy outcomes are reported in a way that makes it possible to determine the role of scaling – for instance, whether policy change is occurring beyond primary intended users as described in the scaling pathway.

1. The original target to support 'at least 20 initiatives that deliver solutions at scale' was tied to IDRC's strategic objective to invest in knowledge and innovation for 'large-scale positive change' (emphasis added). Over the strategic period, the Centre's understanding of scaling evolved and the corporate indicators used to monitor scaling and that were used as the basis for reporting outcomes in this evaluation were broader to track scaling outcomes more generally and not just large-scale impact, thus it is unsurprising that IDRC surpassed this original target. See section 4 on what outcomes were achieved for a more detailed analysis.

The evaluation found that the introduction of the scaling objective had two main positive internal effects on IDRC: a stronger shared purpose and collaboration within programs with fewer isolated projects, and a change in mindsets for many staff related to scaling such as considering scaling earlier to inform research design and process rather than only at the end of the research process.

Implementation of the strategic objective

IDRC took a flexible approach to implementing the strategic objective, whereby programs developed their own understanding, strategies and approaches to scaling. Given the diversity of programming and existing approaches to scaling, this flexible approach was appropriate, and programs used this opportunity well to adapt and develop tools for scaling. While some staff appreciated the flexibility, others found the lack of conceptual clarity around scaling terminology and approaches a challenge. Lack of conceptual clarity and a common approach to scaling also created difficulties for monitoring progress and evaluating results of the strategic objective across programs.

As programs experimented and innovated with approaches to scaling, they used and adapted various tools and processes to support their efforts. The design features and factors that have emerged that contributed to scaling confirm the importance of flexibility and adaptability given that most of the challenges that IDRC is trying to address are complex, and that scaling itself is an intervention with its own levels of complexity. The evaluation found that programs have developed ways to be flexible and adaptive but there needs to be stronger capacity for strategic leadership in programs to leverage these aspects for scaling.

According to staff and grantees, flexible funding mechanisms were one of the most helpful tools to support scaling efforts, although the evaluation did not find evidence of them being systematically documented or applied across programs. Programs have been experimenting with different approaches to allow more time for research results to scale, although more time is needed for research projects to engage with the demand side of scaling – which involves strengthening the knowledge and capacity of actors beyond the primary research users.

Programs have found that investing across the scaling pathway with different kinds of investments in their portfolios – specifically concurrent investments in the supply side of scaling (knowledge and innovation generation) and the demand side of scaling (mobilisation and use of knowledge and innovation beyond immediate users) – supports scaling. The evaluation evidence suggests that programs that have paid attention to the eco-system for scaling, bringing in the right partners at the right time, are been better able to support scaling.

Knowledge synthesis emerged in the evaluation as an important means for programs to identify opportunities for scaling research results. This is done in many ways in IDRC by bringing information together horizontally among research teams, across a portfolio, or vertically when communicating with policymakers. However, the evaluation found that knowledge synthesis is under-resourced in programs, and program staff told us they had insufficient time, capacity, skills and incentives to do it properly.

The role of responsible officers (i.e. IDRC program staff who manage projects) was crucial to scaling efforts and the evaluation found that they have been required to think and act more strategically and opportunistically. We found that many responsible officers are taking on this role, transitioning from funder and technical partner to knowledge broker, knowledge translator, coordinator and strategic thinker. However, there is not yet a formal recognition of this change in role and some staff feel that they have insufficient time, resources and incentives to carry it out effectively.

The 2015–2020 strategic plan recognised the importance of coordinating with actors who can support scaling. The evaluation found that IDRC programs are coordinating with a set of actors beyond the research community to support scaling research results. For instance, they have increasingly brought non-research actors – such as private sector partners or policymakers onto research project teams. The Centre is still most comfortable working with researchers and the public sector but is increasing its understanding of how to work with the private sector though many opportunities remain for learning more.

The evaluation highlights numerous opportunities both within and across programs for learning about scaling. We found that learning has occurred more within programs than across programs and it tended to be more conducive in less formal processes than in more formal processes. Externally funded programs were more likely to provide substantive learning opportunities than core funded projects, and learning tended to happen on a project-by-project basis rather than across a portfolio of projects. Staff reported that grantees were one of the most important sources for learning about scaling. The most pressing challenge for learning mentioned by staff was extreme time pressure.

IDRC's Scaling Science initiative was important for learning from and across programs as they experimented with incorporating scaling. The guiding principles for scaling impact that emerged from the Scaling Science work provide a common framework and terminology for scaling while allowing flexibility in scaling approaches and strategies. The principle-based approach has helped position IDRC as a key contributor in the scaling field – with a strong focus on scaling impact and attention to the equity of impact – yet more needs to be done to communicate these principles across the Centre. Programs are considering gender in their programming, and this is supporting the scaling of equitable gender outcomes. The Scaling Science work has provided a lens for

thinking about how to integrate gender equality into scaling strategies, however staff are not yet able to clearly articulate the link between gender and scaling processes.

Future considerations for strengthening scaling research results at IDRC

We have suggested the following considerations for IDRC to help build on the strengths and to address the challenges that emerged from the evaluation findings. We believe that many of these considerations will also be relevant and of interest to other organisations seeking to scale the impact of research results.

Corporate and cross-program considerations

Having a broad and diverse **conceptualisation of scaling** has made it difficult to collectively learn about scaling across the Centre and with grantees, while also enabling experimentation and adaptation in individual programs. Looking ahead to IDRC's next strategic period, could a more unified approach to understanding scaling of research results benefit IDRC? Should the Centre provide enhanced support for staff and grantees to better understand and use the concepts introduced in the Scaling Science work?

Learning about scaling within programs is stronger than learning across programs. In what ways can the Centre's upcoming learning agenda facilitate learning about key cross-cutting issues like scaling, and enable conversations about challenges and failures in scaling given that scaling involves higher levels of risk in programming and that not all projects could or should scale?

Monitoring, evaluation and reporting on scaling has presented some challenges for IDRC both in terms of what and how to monitor and report and assessing impact at scale within program timelines. How can IDRC's scaling outcomes and contribution to impact at scale be tracked and assessed in a more systematic and reliable way? To what extent is it feasible and appropriate to expand the scope of program monitoring and evaluation to better examine the significance of outcomes and incorporate more analysis of the potential benefits and risks of impact?

Considerations in programming

Scaling research results takes time which means it is important for programs to have **long time horizons for investments in scaling**. Across IDRC program portfolios, what is the right balance for supporting longer-term investments, multi-phase projects and strengthening strategic partnerships specifically with a view to scaling research results? Are there other ways that the Centre can support programs to 'position themselves' to achieve impact at optimal scale, even if scaling research results beyond primary intended users itself takes place after the end of the IDRC-funded project or program's lifetime?

Scaling is facilitated when there is a strong knowledge base, actors, infrastructure and resources in a field (a scaling eco-system) on both the supply and demand side of the scaling pathway. A **portfolio approach for scaling** can facilitate this. Could a portfolio approach be more strategically used within IDRC programs to build eco-systems for scaling by investing in projects that support knowledge and innovation supply *and* demand?

Flexible funding practices support programs to scale by building on existing work or responding to emerging opportunities. Should IDRC leverage flexible funding mechanisms more systematically across the organisation to scale research results? Could formal criteria and process be beneficial to promote flexible funding and support a more consistent understanding across the Centre of the flexible funding options that can be used to support scaling?

Coordination for scaling requires strategic thinking within programs as the logistics can be time consuming and resource intensive for programs and grantees. What can IDRC do to further support programs with the competencies, skills and resources required to coordinate multiple stakeholders across a scaling pathway?

Knowledge translation and synthesis are critical for scaling research results. Could IDRC provide additional support to provide more time and resources for knowledge translation and synthesis work with a focus on scaling? What is the most appropriate level of emphasis for knowledge translation and synthesis at the corporate, program and grantee levels, and how can these activities be effectively coordinated for enhanced influence and impact?

Grantees appreciate the support provided by program staff for scaling but **support to grantees** could be improved, particularly to better understand optimal scale and potential negative effects of scaling. These are important considerations for designing and implementing inclusive scaling processes that promote equitable outcomes. What additional support is required to facilitate more nuanced discussions with and among grantees about scaling throughout the research process in a way that encourages responsible scaling?

Considerations for external influence

IDRC has an important **position in the scaling field**, to encourage scaling of research results in a way that promotes equitable and sustainable outcomes, and to counterbalance the mainstream thinking of 'bigger is better'. To what extent and in what ways should IDRC continue to, or even strengthen its, influence on the evolving debates and dialogue in the scaling field, with the paradigm of principled scaling?

1. INTRODUCTION

The International Development Research Centre (IDRC) was established by an act of Canada's parliament in 1970 with a mandate 'to initiate, encourage, support, and conduct research into the problems of the developing regions of the world and into the means for applying and adapting scientific, technical, and other knowledge to the economic and social advancement of those regions'. In pursuit of this mandate, IDRC committed to 'invest in knowledge and innovation for large scale positive change' as one of the strategic objectives in its 2015–2020 Strategic Plan.

Under this strategic objective, IDRC committed to generate, identify and test scalable ideas and innovations. IDRC would further support the ideas and innovations that had implementation potential by working to connect solutions with actors who could help advance those solutions to achieve large-scale impact. IDRC also committed to 'examine early wins in scaling up to identify and share critical success factors' (IDRC, 2015, p. 5).

Much of IDRC's funded research relates to 'wicked problems' in 'complex systems', where there is a lot of uncertainty about what works, when, where, why and how. At the beginning of the strategy period, there was no blueprint for 'scaling up' or achieving 'large scale positive change', nor did a standard definition of scale or scaling exist at IDRC. The strategic objective reflected existing practice within the organisation, with several programs already experimenting with approaches to scaling. These early scaling efforts at the Centre (prior to 2015) focused on 'scaling up' innovations and research results. However, under the 2015–2020 strategy, IDRC programs were encouraged to experiment and innovate with respect to the strategic objective – they were to identify and define their own conceptualisations and approaches to scaling in a manner that was tailored to their unique programming context.

Thus, over the 5-year strategic period, senior leadership, programs and projects have grappled with the notions of 'scale' and 'scaling', conceptualising and approaching it in different ways and to different extents. Alongside the experimentation and reflection that programs were doing with respect to scaling, IDRC launched the Scaling Science initiative. This initiative was designed to advance the organisation's understanding of scaling and to learn from programs across the organisation. It began with a review of over 200 research projects, which resulted in the publication of a book *Scaling Impact: Innovation for Public Good* (referred to as the *Scaling Impact* book from here on) (McLean & Gargani, 2019) and the *Scaling Playbook: A Practical Guide for Researchers* (Price-Kelly, van Haeran & McLean, 2020).

As the Centre enters a new 10-year strategic period, it remains committed to learning about scaling and sharing that learning with others. As part of this continued

commitment, IDRC commissioned OTT Consulting, in partnership with Southern Hemisphere, to undertake an evaluation of IDRC's implementation of its strategic objective to scale and what was achieved by those efforts.

1.1 Evaluation purpose and users

1.1.1 Purpose

As stated in the terms of reference, the evaluation had a dual summative and learning purpose, with the following objectives:

- 'Assess results to scale the impact of research for development.
- Provide insights into the strengths and weaknesses of past and current programming to improve future IDRC corporate, program and project-level strategies.
- Consolidate learning from IDRC's experience to share with grantees and other research organisations who wish to strengthen the impact of their work at appropriate scales.' (IDRC, Terms of Reference for this Evaluation).

1.1.2 Evaluation questions

During the inception phase, the evaluation team identified the following evaluation questions based on an initial list provided in the evaluation terms of reference and in discussions with IDRC. They follow the summative and learning intentions.

Summative questions:

1. How well did IDRC meet its strategic objective to invest in knowledge and innovation for large-scale positive change (its 'scaling objective')?
- What did IDRC do to implement the strategy to scale at the corporate, program and project level?
- How well did it plan and implement, coordinate, and learn about scaling?
- What were the strengths and challenges for implementation?
- What outcomes were achieved by integrating scale into programming?
- How valuable and sustainable are those outcomes, and for whom?

Learning questions:

2. What can be learned from this experience to inform the future strategy about scale for IDRC, its grantees and the wider scaling community in terms of IDRC's

four guiding principles for scaling (justification, optimal scale, coordination, dynamic evaluation)?

- What can be learned from the evaluation of IDRC's scaling strategy, including the perspective of IDRC, its grantees, social actors and others?
- What can be learned from comparison with other research funding organisations?
- What considerations can be made to improve IDRC's scaling efforts in the next strategic period?

1.1.3 Users

We identified four main evaluation user-groups:

- IDRC senior management, who may use lessons and recommendations to inform implementation of the 2020–2030 Strategic Plan.
- IDRC program staff, who may use the results to adjust strategies and processes for scaling.
- IDRC grantees, who may use the results to inform their own scaling efforts.
- Other donors, researchers and innovators working to scale research results, who may use learning to inform their own strategies and practice.

1.2 Evaluation approach and methodology

This evaluation took the form of a strategic review that employed a mixed-methods design. Our approach to the evaluation was informed by utilisation-focused evaluation (Patton, 2008). The approach, and what informed it, is described in more detail below.

1.2.1 Approach

This was not a standard evaluation. The evaluand is not a clearly defined project or program but an organisation-wide strategic objective that cuts across all IDRC's programs and projects and is implemented in a variety of ways that have evolved over time. As described in the introduction, there was no standardised guidance or framework for implementation of the strategic objective across the organisation when the strategy was initiated. As such, we could not assess IDRC's effectiveness by looking at planned versus actual implementation of the strategic objective, nor could we evaluate how the organisation met or adapted its plans. Therefore, the evaluation was exploratory in nature – we identified the approaches, practices, systems and processes that supported or hindered scaling efforts by allowing them to surface through our analyses of data collected from IDRC staff, grantees and co-funders.

IDRC's approach to scaling was a dynamic and evolving practice, not a static uniform approach that can be evaluated easily. The Scaling Science work was a systematic effort to learn from programs and codify IDRC's approach to scaling. This resulted in the emergence of four guiding principles, first published in 2017 and developed further in IDRC's book *Scaling Impact* published in 2019, towards the end of the strategic period. As a consequence, the guiding principles and definition of scaling contained within were not available for use by IDRC programs for most of the 2015–2020 strategy timeframe. Therefore, using these concepts as a framework for a summative evaluation would have been inappropriate. Instead, we have built the principles into the questions driving the learning component of the evaluation.

In light of this context, we adopted a **strategic review approach** to answer the evaluation questions. This involved looking at IDRC's response to the strategic objective at the corporate and program level, grantees' perceptions and experiences of this response, and other similar donors' activities with respect to scaling. Considering the organisation-wide scope of the evaluation, the desire to bring in multiple-stakeholder perspectives and the need to first gain a deeper understanding of how the Centre responded to the strategic objective, we organised the evaluation activities into four interconnected **components**:

1. An organisational review of systems and processes, strengths and weaknesses, and outcomes achieved.
2. A study of grantee perceptions and experiences.
3. A study of IDRC's positioning within the wider research for development sector.
4. Four thematic case studies: field building for scale, programming for scale, private sector engagement and scaling, and organisational learning about scaling.

We engaged with IDRC throughout the evaluation, inviting staff to provide input and comments at critical stages and to join us in sensemaking of the data. An evaluation advisory group comprising representatives from across the organisation participated in inception meetings, an 'emerging-findings' workshop and a 'findings to recommendations' workshop.

1.2.2 Methodology

In this section, we highlight key methodological aspects of the evaluation. A more detailed description of the evaluation methodology is provided in Appendix 2.

Inception phase

The inception phase began in October 2020, culminating in an inception report submitted to IDRC's Policy and Evaluation Division (POEV) in February 2020. For this phase we conducted the following:

- Two meetings with the evaluation advisory group established by POEV, comprising program leaders, program officers, program management officers, a regional director and POEV staff.
- In-person interviews with 14 members of staff and leadership at IDRC.[2] These individuals represented a cross-section of the organisation and had been involved in implementing scaling strategies. Interviewees included program officers, program leaders, directors and a vice-president.
- Interviews with 11 external stakeholders who were identified as leading experts in scaling within international development. We identified individuals through conversations with POEV and snowball sampling with initial interviewees.
- A scan of relevant documents and literature.

The inception report included details regarding the methodology for the proposed evaluation including information about sampling and an evaluation framework. The evaluation framework included detailed evaluation sub-questions developed through discussions with IDRC staff and a review of documents that described and evaluated other development organisations' responses to scaling. These documents included but were not limited to the World Health Organization's *Nine Steps for Developing a Scaling-Up Strategy* (2010) and an institutional review of IFAD's approach to scaling (Linn *et al.*, 2010). The evaluation framework was also informed by IDRC's book *Scaling Impact: Innovation for the Public Good* (2019) and the *Scaling Playbook* (2020). This framework became the basis for the coding sheets and data analysis plan used during the implementation phase.

Implementation phase

The implementation phase began in February 2020 with some adjustments in plans due to changes in the context and further insights regarding end-user needs. Specifically, the following changes were made:

- The Covid-19 pandemic started shortly after the implementation phase began, requiring the evaluation team to cancel travel plans and move all data collection online. We redesigned face-to-face focus groups with IDRC grantees and

2. Nine of the 14 interviews were with members of the evaluation advisory group.

workshops with the advisory group as online activities, and extended the data collection period to arrange online interviews.

- A set of programmatic case studies using an 'outcome harvesting' approach were reconceptualised based on conversations with POEV and the Evaluation Advisory Group. It was felt that an assessment of the results of scaling in a small number of programs would say little about the Centre as a whole. A broader assessment of a larger number of programs, focused on carefully selected themes, could produce more generalisable and useful data and learning, and would be more in line with a strategic review. The decision was made to conduct four 'thematic' case studies focused on specific areas of interest to IDRC: (1) field building for scale, (2) programming for scale, (3) private sector engagement and scaling, and (4) organisational learning about scaling. All cases would include a gender dimension.
- With the move to thematic case studies, we dropped the outcome harvesting approach from the evaluation methodology. To answer the evaluation question 'What outcomes were achieved by integrating scale into programming?' we instead drew upon IDRC's corporate monitoring system, program evaluation reports and respondent perceptions of outcomes.

1.2.3. Data sources and sampling

The data sources and sampling approaches used within each component of the evaluation are described below. We provide a breakdown of the sample sizes for each component at the end of this section in Table 1. Additional details are provided in Appendix 2.

Organisational review

The organisational review included data from an online staff survey, semi-structured interviews, a document review and IDRC's program results database, Trackify. The IDRC staff survey was circulated to all program staff via email and promoted on Twitter. We selected a purposeful sample (Patton, 2002) of senior members of staff for semi-structured interviews who could shed light on the organisation's response to implementing the strategic objective to 'invest in knowledge and innovation for large scale positive change'. We invited the following respondents to participate: IDRC president, vice-presidents of program and strategy branches, all three directors of program areas, the directors for POEV and information management and technology, and two regional directors based on years of service and availability.

The documents reviewed came from a comprehensive repository of internal and external sources that IDRC compiled and shared with the evaluation team. We identified additional sources through recommendations from interviewees, internet searches and

examination of documents available through IDRC's SharePoint system. Examples of program and project documentation the evaluation team reviewed included corporate, program and project plans, implementation plans, reports to the IDRC Board of Governors, program and project evaluations and learning reports.

The analysis of outcomes drew on 13 program evaluations, 9 program area reports to the Board of Governors, interviews and surveys of staff and grantees and 2 datasets extracted from Trackify, one for each of the corporate-level indicators for scaling. The detailed approach is provided in Appendix 1.

Study of grantee perceptions and experiences

We obtained data from IDRC grantees through an online survey (95 respondents) and online focus group discussions (5 discussions with a total of 18 grantees). The IDRC grantee survey was promoted via the IDRC website and Twitter. To increase take up, POEV asked program staff to encourage their grantees to participate in the survey. We selected participants in the focus group discussions through convenience sampling. An invitation to participate was sent to all participants of the IDRC scalingXchange, and survey respondents were given the option to participate via a separate web form to maintain anonymity of the survey results.

Positioning study

We used an online survey and semi-structured interviews to obtain input from funders. Potential survey and interview respondents were purposively selected to cover a range of research funders, including multi-lateral organisations, bi-lateral organisations, public-sector funded organisations, think tanks and philanthropic organisations. This process resulted in 38 potential respondents. We disseminated the online survey to the identified potential respondents with a personal request from the evaluation team to complete the survey. The team also promoted the survey via email and on Twitter to attract other relevant respondents. We used a criterion sampling approach to identify a shortlist of potential interviewees, comprising funders who had been involved in scaling initiatives and whom IDRC identified as good comparators for the organisation.

Thematic case studies

In collaboration with POEV, the evaluation team identified a sampling frame of 64 programs and projects for the case studies. Several sources provided input on which programs and projects to include in the sampling frame: we asked staff for suggestions during interviews in the inception period, we looked for good cases of scaling within documentation provided by IDRC and invited program leaders to suggest programs and projects for inclusion in the case study samples via an online questionnaire. The evaluation advisory group then reviewed an initial list and made recommendations about

programs to add and remove. POEV further identified a set of programs and projects as having scaling intentions based upon an earlier mapping analysis (Sanchez-Swaren, 2018) and projects relating to participants in the ScalingXchange were added.

Each case study focuses on a subset of this sample of 64, with some programs and projects included in more than one case study. Three of the case studies had a thematic focus (private sector and scaling; field building for scale; and programming for scale). These three case studies covered 10 programs[3] and 14 projects. Of these 10 programs, 6 were examined at the program level only; and 4 included program and project level. See tables 4 and 5 in Appendix 1 for a detailed breakdown of programs and projects per program area and per case study.

The evaluation team allocated programs (and some projects) to the field building, program design and private sector case studies based on intensity sampling – programs or projects likely to provide the richest information on the topic of interest to that case study were included. Interviewees were purposively sampled from within programs and projects to ensure a balance of gender, program areas and roles. We took a random sample from the original list of 64 programs and projects to select staff to interview for the organisational learning case study. Adjustments to this sample were made with the aim of achieving balance in gender of interviewees, program area representation and types of roles. We purposively selected regional directors not associated with specific programs or projects.

We aimed to ensure minimal overlap of respondents among the case studies. We selected interviewees for each case study based upon the respondents' availability and asked program officers in Ottawa to assist in identifying potential interviewees from regional offices, grantees and strategic partners. Some respondents were interviewed twice because of the relevance of their work for more than one case study and some elected to participate in interviews for a case study other than the one to which we had invited them.

Each case study draws on semi-structured interviews with IDRC staff, grantees and other funders as well as the document review (as described under the organisational review) and relevant literature. Each interview schedule included two cross-cutting questions about learning, outcomes and equity in addition to questions pertaining to the specific theme of interest. For each of the cases, we aimed to complete 10–20 semi-structured interviews (SSIs).

3. Programs include both IDRC programs and a sample of the co-funded programs they manage.

Table 1: Total number of respondents for all components (excluding inception phase)

Component	IDRC	Grantees	External
Organisational review	10 semi-structured interviews (SSIs) with senior staff 43 survey responses from program staff	-	-
IDRC's positioning	-	-	15 SSIs 16 survey responses
Case study 1: Field building for scale	7 SSIs	4 SSIs	1 SSI
Case study 2: Programming for scale	8 SSIs	8 SSIs	4 SSIs
Case study 3: Private sector engagement and scaling	11 SSIs with current staff, and 1 with former staff member	3 SSIs	-
Case study 4: Organisational learning about scale	18 SSIs	-	-
Grantee experiences and perceptions	-	5 focus group discussions (FGDs) with 18 grantees from 4 regions 95 survey responses	-

TOTAL	55 SSIs 43 survey responses	15 SSIs 5 FGDs with 18 grantees 95 survey responses	20 SSIs 16 survey responses

1.2.4. Limitations

The biggest challenge in designing this evaluation was the lack of a common definition of scaling or framework for implementation of the strategic objective. Without a clear set of criteria, we concluded that we would be unable to answer the 'how well did IDRC meet its strategic objective' question. Instead, we focused on identifying strengths and weaknesses, as well as identifying lessons about planning and implementing for scale, and coordinating and learning about scale during the strategic period.

The lack of a common definition of scale or scaling across the organisation made it difficult to ask questions in a way that would be understood by all respondents, and then to aggregate and compare responses. Anticipating this, we produced a blog and an animation to communicate to respondents the concept of scaling used in the evaluation. Despite this, not everyone we spoke to saw their work as relevant to scaling or could engage critically about their scaling work in interviews. This was particularly apparent when asking about scaling in relation to gender and inclusion, and private sector engagement.

Identifying the outcomes achieved through scaling and determining how valuable and sustainable those were and for whom was challenging because of trade-offs in data collection. Rather than a deep focus on a small number of programs or projects, as originally proposed, we took a broader, thematic look across a larger number of programs and relied on existing data sources including evaluations, program reports and monitoring data (Trackify). This meant that data was less detailed about outcomes than we would have liked, and the quality of data varied depending on the sources we drew on – making aggregation and analysis more difficult.

Survey respondents were self-selecting, and although the surveys were marketed at program staff and grantees (irrespective of whether they were actively engaged in scaling processes) the results show that most grantees were knowledgeable about scaling and implementing scaling strategies. This could indicate a bias – that those with more to say

about scaling completed the survey. The field building and private sector engagement case studies had limited external input and hence are biased towards IDRC perspectives.

1.2.5. Challenge

The Covid-19 pandemic started shortly after the implementation phase began. All planned field visits and face-to-face exchanges were cancelled, postponed and/or moved online. Given the difficult work circumstances, it was, at times, harder to convince people to participate in the evaluation, including IDRC staff. However, we were satisfied with the responses received for all the surveys. Even though the funder survey was a small sample (16), we had only identified 38 potential respondents and a response rate of almost 50% is acceptable. The evaluators were working under challenging circumstances, which contributed to lengthening the evaluation timeframe but did not affect the quality of the data.

It was also challenging to answer summative questions about effectiveness because the scaling work of IDRC is ongoing, even though the strategic period has come to an end. The decision to refocus the evaluation around a strategic review with a thematic focus helped with this.

1.3 Conceptual framework: The scaling pathway

The evaluation team developed a 'scaling pathway' through the course of the evaluation to make sense of the ways in which programs approached scaling and the outcomes they achieved from integrating scaling into their work (see Figure 1). We refer to the scaling pathway and the elements described in it throughout the report.

The scaling pathway comprises two interconnected pathways: a policy pathway and an innovation pathway. The policy and innovation pathways relate directly to the corporate high-level scaling indicators adopted by IDRC and they reflect the way many programs reported progress against the scaling objective.

The scaling pathway draws on two frameworks: policy outcome categories adapted by POEV from Carden (2009), and stages of innovation developed by an IDRC working group in 2018. We took these two frameworks and overlaid them on IDRC's generic results categories: outputs, immediate outcomes, intermediate outcomes and development outcomes. Although the scaling pathway has similarities with many program impact pathways, we intentionally avoided calling this an impact pathway to avoid confusion. The scaling pathway helps clarify the distinction between the supply and demand sides of scaling. The supply side (left hand side) refers to generation and mobilisation of knowledge and innovation, while the demand side refers to use of the knowledge and innovation to support development outcomes at optimal scale. A

BOX 1: High-level indicators used by IDRC to monitor scaling

1. # Innovations being widely used and adopted
2. # New policies implemented or changed

program or project that is intending to scale the impact of research needs to balance its investment to support both domains.

We integrated scaling language into the frameworks. For example, we added 'at optimal scale' to the intermediate and development outcomes, which is a concept developed by IDRC's Scaling Science work during the strategic period. We also proposed a distinction between policy change or implementation *with primary intended users* of the knowledge or innovation generated from research, and policy change or implementation *beyond primary intended users* at optimal scale. For example, an IDRC-supported research project had success in influencing policy at national level (achieving outcomes with primary intended users) and a follow up project was developed to target similar policies at regional level (to achieve outcomes at optimal scale beyond the primary intended users). Another example could be that IDRC-supported research informed the content of a new policy developed by primary intended users and was subsequently accepted and implemented by actors beyond those primary intended users.

The distinction between adoption by primary intended users and adoption at scale was clear in the innovation track but had not yet been articulated in the policy track. For example, the high-level indicator displayed in Box 1 counts the number of policies implemented or changed but does not distinguish who the decision-makers were – they could be primary users or others beyond this group. We proposed this distinction as a hypothesis – to prompt further reflection about the distinction between policy change in general and policy change at optimal scale.

The dotted green and blue lines in the diagram show that adoption of an innovation or policy change by primary intended users can contribute to development outcomes at optimal scale given that 'optimal' is a relative term defined differently in each case.

In the scaling outcome pathway, we also made explicit that scaling begins in the outputs and activities stage and continues throughout the pathway (as noted in in Figure 1). In the early stages, this is about positioning the program or project to achieve impact at optimal scale. A key activity to support scaling at all stages is engagement with other actors who have a role in achieving change at optimal scale. Many of these actors go beyond those who research projects would typically engage with (e.g., policymakers).

The two pathways are interconnected and not mutually exclusive (indicated by the white circle with arrows in the centre, intended to show multiple connections between the two pathways). For example, knowledge generated in the piloting and testing of an innovation is considered new knowledge and might itself be mobilised to influence policy. Likewise, policy change can affect and may be a precursor to an innovation being optimally used. A draft policy or policy instrument may even be the innovation a project is scaling, in which case progress can be tracked along both pathways simultaneously.

In addition to contributing to improved lives and livelihoods directly via the two pathways, an indirect route may be at play. This indirect route occurs through systems strengthening (see the red circle and associated double-ended arrows in Figure 1), which flows both ways between the policy and innovation pathways. For example, programs may contribute to policy capacity or dialogue without affecting actual policy change; this can still have a positive effect on people's lives through strengthening the system in which policy is developed and implemented. Likewise, an innovation may not affect people's lives directly but can improve the equity and sustainability of other initiatives to improve people's lives. As the Canadian International Food Security Research Fund (CIFSRF) program has demonstrated, programs can use their innovations as entry points to catalyse systems change (Shilomboleni *et al.*, 2019).

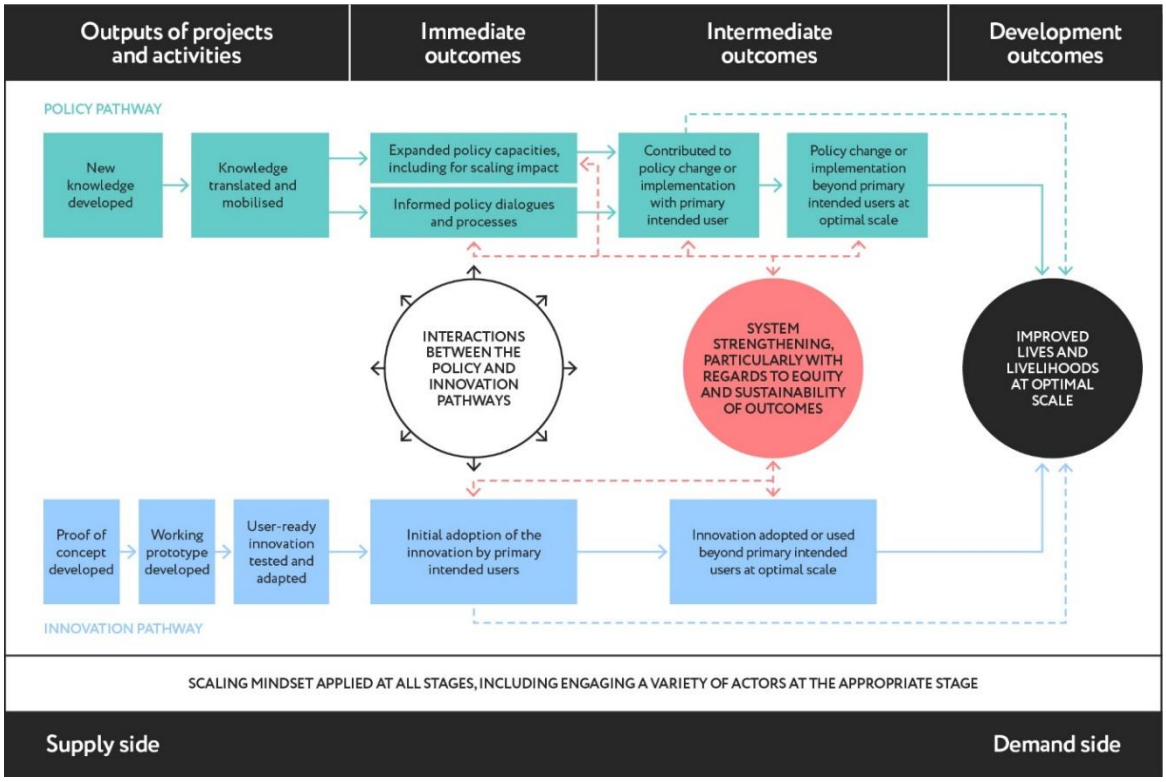

Figure 1: Scaling pathway (see accompanying text in section 1.3 for explanation)

2. IMPLEMENTATION OF THE STRATEGIC OBJECTIVE AT THE CORPROATE LEVEL

This section examines IDRC's approach to implementing the strategic objective to scale at the corporate level. It looks at the mechanisms put in place to socialise the strategic objective across the Centre, the approach taken to promote implementation by programs, and how IDRC learned about its progress and from the experiences of programs and projects during the strategic period. This section draws on interviews with 10 senior IDRC staff, interviews with program staff for the four case studies, the survey of 43 program staff, and a review of corporate documents.

2.1. IDRC's scaling journey

Since IDRC was established in 1970, the Centre has aimed to leverage the power of science to create meaningful impact in people's lives. In this way, IDRC's research has always focused on purpose-driven research that supports practical action and use (Gonsalves & Baranyi, 2003). The scaling objective in the 2015–2020 strategy was one of three directions pursued by IDRC in this period to fulfil its mandate. In the few years prior to the introduction of the strategic objective to scale, pockets of IDRC had begun incorporating ideas of scale and scaling up in their programming. Two examples from among the programs we looked at in this evaluation are the Canadian International Food Security Research Fund (CIFSRF) Phase 2, which began in 2013 and sought to scale up solutions that increase production, access and consumption of safe and nutritious food; and Innovating for Maternal and Child Health in Africa initiative (IMCHA), that was launched in 2014, and was designed to facilitate scaling of solutions through involvement of government decision-makers.

The new strategic objective put a spotlight on scaling and provided an impetus to unpack and use the concept specifically and intentionally. We heard in several senior staff interviews that the strategic objective to scale was intended to provide the necessary direction to support a shift in thinking within the organisation to raise IDRC's ambition to achieve more at higher levels. As one senior manager said, '*We had this ambition which required a culture change*' (SSI, IDRC staff).

Staff involved in designing the strategic objective told us that the starting point was to take the lessons learned from existing programs that had experimented with scaling approaches, most notably CIFSRF, and apply them to other programs.

When the 2015–2020 Strategic Plan was being developed, the context around international development funding was changing and accountability to achieve visible results with public money was intensifying. At the same time, IDRC's Board of Governors

was requesting greater focus on quantifiable results from senior managers and programs. Several interviewees suggested that this helps explain the focus on large-scale impact at that time.

There was also growing interest in scaling within the research for development sector in the years leading up to IDRC's 2015-2020 Strategic Plan. In 2010, the World Health Organization published *Nine Steps for Developing a Scaling-Up Strategy* and an institutional review of the International Fund for Agricultural Development's scaling up strategy was published the same year. In 2013, Wigboldus, SA and Leeuwis published an article on 'responsible scaling up and out in agricultural development' for the CGIAR Research Program. The Scaling Up Community of Practice (CoP) was founded in 2015 to provide a platform for knowledge exchange among practitioners on approaches to scaling up development interventions (IDRC joined the CoP in 2017).

The increased interest in scaling was therefore motivated both by an internal commitment to achieve greater impact with research and by increased external discussion around scaling up within the international development community and pressure to increase accountability for research funding.

Figure 2 presents a timeline of key moments in IDRC's scaling journey. It includes corporate activities and Scaling Science activities, and also shows a number of programs that we refer to in this evaluation – but not all of them. Since 2015, much has changed. What scaling means now in IDRC is not necessarily the same as how it was understood in 2015 when the Strategic Plan was launched. We explore this further in the rest of this section.

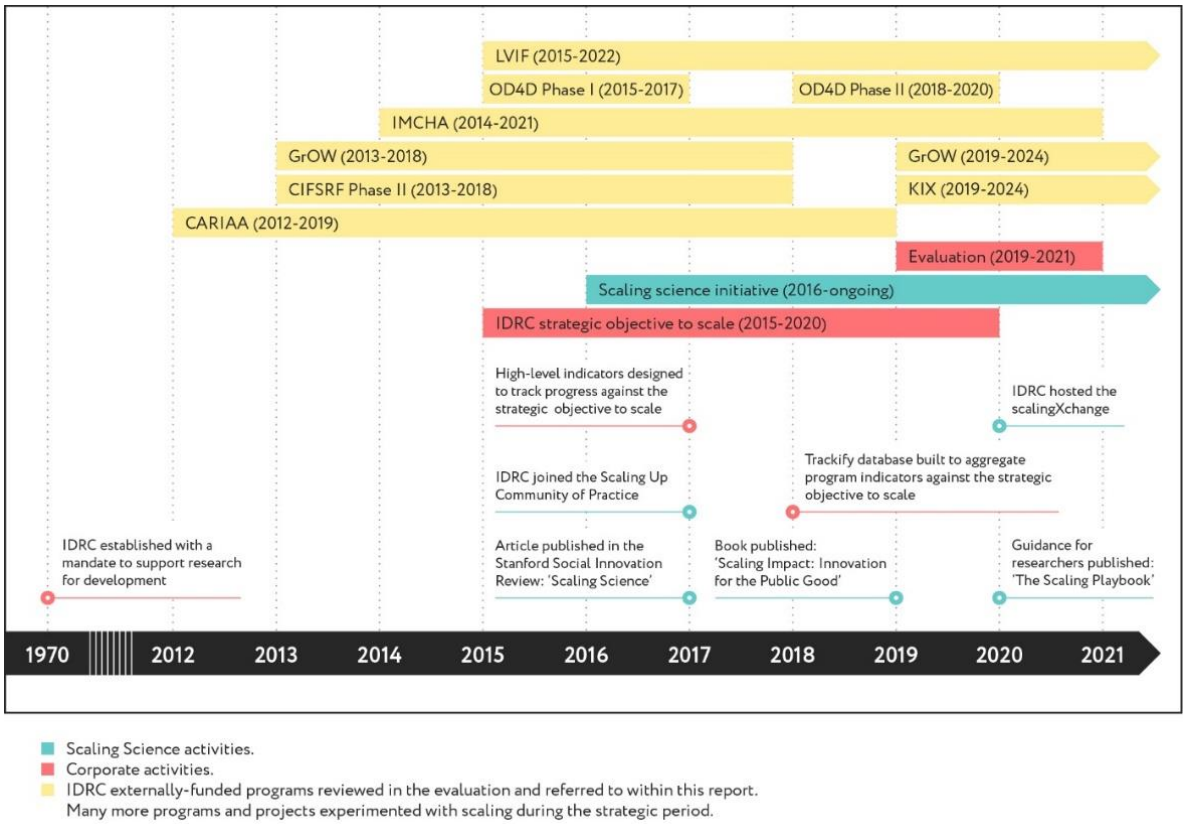

Figure 2: A timeline of key moments in IDRC's scaling journey

2.2. Corporate mechanisms

KEY TAKEAWAY: Following the introduction of the strategic objective, corporate mechanisms for program and project planning, monitoring and reporting were put in place to help staff consider scaling in IDRC's programming.

While some programs were already experimenting with approaches to scaling prior to the 2015–2020 strategic objective, the new strategic objective set in motion several corporate mechanisms to facilitate thinking about scaling among programs and staff across the Centre (see Figure 3). Program area implementation plans were required to outline how they would contribute to the strategic objectives, under which core programs would set indicators and targets; project approval documents and completion reports included a question on the strategic objectives; and programs were required to report on progress against the strategic objectives in program area progress reports and annual performance and learning reports to the Board of Governors.

These mechanisms broadened the number of programs and staff thinking about scaling in their work: out of 43 IDRC staff who responded to an evaluation survey, 51% said they

had been thinking about scaling prior to the strategic objective and 49% said they started thinking about after the introduction of the strategic objective.

While the strategic objective helped more staff and programs consider scaling, it was never intended for all IDRC programs or projects to scale. In the project approval documents it was clearly stated that not all projects were required to contribute to all strategic objectives.

Figure 3: Corporate mechanisms to implement IDRC's strategic objective to scale 2015–2020

2.3. A flexible approach

KEY TAKEAWAY: IDRC took a flexible approach to implementing the strategic objective. While some staff appreciated the flexibility, others found the lack of conceptual clarity around scaling terminology and approaches challenging.

IDRC took a flexible approach to implementing the strategic objective, allowing programs to develop their own approaches and strategies. The flexible approach to implementation was appreciated by some staff – among 43 IDRC staff members who participated in the survey, 60% identified 'flexibility to interpret and develop appropriate scaling strategies' as a strength of IDRC's response to the strategic objective. At the same time, 56% of staff respondents said that 'understanding the concepts of scale and scaling internally' was a challenge. This dual sense of appreciation for the flexibility and the challenges of not having a clear approach is captured in this IDRC staff survey comment:

I never felt pressure to adopt or implement a certain approach to scaling across projects. I think this was very positive for more experienced program officers with insight into how to interpret and implement this scaling objective across their portfolios. However, the flip side is that I'm not sure I benefited from sufficient guidance or [was] given examples of how this might look in practice. So, I'm not sure to what extent the work I've supported has contributed to this corporate objective or not, nor if the way I supported my projects was appropriate. (Survey, IDRC staff)

The diversity of approaches to scaling across the organisation during the strategic period is evidenced in the range of ways staff described scaling in interviews (see Box 2). At the time of the evaluation, some interviewees' understandings were informed by IDRC's Scaling Science study, others by the development of their own practice.

BOX 2: A summary of different ways IDRC staff interviewees interpret scaling

- Replicating the impact from intervention(s) or technology(ies) in one area/population in larger areas/populations.
- Replicating impact from intervention(s) or technology(ies) in one area/population and recreating it in a different area/population altogether.
- Distributing the learning from an intervention more widely – which some interviewees referred to as 'knowledge translation'; this includes promoting the uptake of research findings into policy.
- Catalysing change at a systemic level, including supporting more integration, collaboration, inclusivity and connection to context.
- Promoting the uptake of research findings in policy or policy change.
- Turning research into solutions for beneficiaries or target groups.
- Reducing the scale of an initiative (scaling down) if it is not achieving impact at the original scale (e.g., supporting greater numbers leads to a decrease in quality or there are other unintended consequences).
- Strengthening the conditions for research to have impact at a greater scale in the future.

A 2018 mapping study of 50 IDRC projects that expressed intentions to scale in project approval documents (Sanchez-Swaren, 2018) provided a preliminary understanding of

the geographic operating scale, stages of innovation and pathways to scale being adopted. The results from this snapshot again illustrated a diversity in scaling at IDRC (see Figure 4).

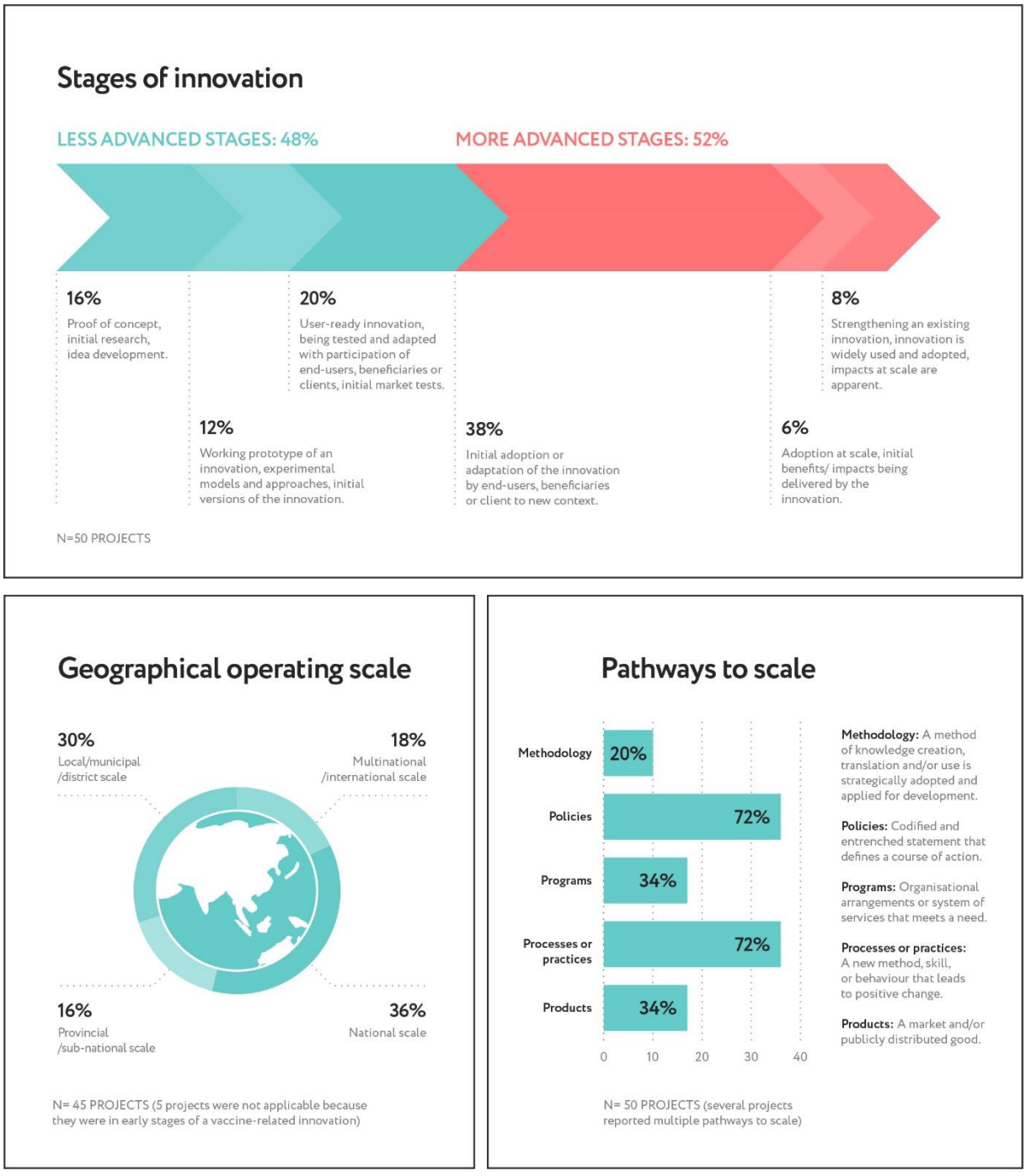

Figure 4: Results of a mapping study of 50 IDRC projects that expressed intentions to scale, 2018

2.4. Monitoring progress

KEY TAKEAWAY: In 2017, IDCR set up a program performance monitoring system to monitor and report on the strategic plan, including two corporate-level indicators designed to track progress against the scaling objective (see Box 1) and a database called Trackify to link program indicators with corporate indicators. Lack of conceptual clarity and/or a common approach created difficulties for monitoring progress and reduced the quality and usefulness of the data. While the Trackify database does contain useful data for monitoring progress against the scaling objective, the data was not sufficiently reliable to be used in corporate reporting without cleaning and analysis.

In 2018, the bespoke 'Trackify' database was built to track program indicators across their impact pathways and, where relevant, aggregate them against the corporate-level indicators to monitor the strategic objectives. The intention was to allow programs to pursue diverse outcome pathways while enabling some standardisation for cross-program aggregation of results (IDRC, 2018b). The data from Trackify was used in annual performance and learning reports to the Board of Governors. However, not all programs used Trackify to report against indicators; for example, the CIFSRF and Collaborative Adaptation Research Initiative in Africa and Asia (CARIAA) monitoring systems were already well developed when Trackify was launched and they had their own monitoring processes in place to meet the particular needs of the funding partners.

Programs were responsible for managing the submission of their data to Trackify, including matching up their own indicators with the corporate indicators. Some programs submitted data after writing program completion reports so they could use extracts from the report as evidence. While this was appropriate given the variation in program impact pathways and approaches to scaling, it led to inconsistencies and errors in the data, some of which had not been detected prior to this evaluation. The issues stemmed from programs having different standards on what they considered to be a scaling outcome and there being no centralised system for identifying and resolving the inconsistencies and errors. Interviewees also told us that the program performance monitoring system emphasised quantity of results over quality and so provided data that was less useful for programs to learn about scaling.

The evaluation team conducted an analysis of the Trackify data corresponding to the two corporate-level indicators for the scaling objective for the purpose of identifying outcomes relating to scaling. The analysis found that Trackify is a useful source of data for results relating to scaling and contains data that is not reported in other program-level sources, for example, program evaluations and program area progress reports. However, the data was not sufficiently reliable to be used in corporate reporting. The analysis by evaluation team required extensive review and cleaning of the data. For

example, only 1 in 5 submissions to Trackify against the indicator relating to contributions to policy change could be confidently classified as such by the evaluation team (97/482). The vast majority of these did not contain sufficient evidence or were more appropriately classified as informing policy dialogue rather than policy change. Further, less than 1 in 10 submissions against the innovation indicator were classified by the evaluation team as innovations being used at optimal scale (16 / 170). Further details of this analysis are in Appendix 2.

2.5. Learning at the corporate level

KEY TAKEAWAY: IDRC's Scaling Science initiative was important for learning from and across the diversity of approaches. The resulting guiding principles for scaling impact provide a common framework and terminology for scaling while allowing flexibility in scaling approaches and strategies. The Scaling Science work has been recognised beyond IDRC by others interested in scaling research results, but more work needs to be done to communicate this work across the Centre.

The most prominent corporate-level learning mechanism was the Scaling Science initiative led by POEV but with several programs substantively engaged. The strategic objective set IDRC's ambition 'to be recognised for sharing its learning in scaling up solutions, helping position Canada as a leader in innovative approaches to development' (IDRC, 2015). Thus, Scaling Science was set up to help develop and share IDRC's learning about scaling within and beyond the Centre. Given the flexible and experimental approach to implementing the strategic objective at the program level, having a mechanism in place to learn from and across that diversity of experience was important.

The initiative began with a series of case studies examining how IDRC's southern research community is advancing research to impact at scale (IDRC, 2018b). The study resulted in an article 'Scaling Science', published in 2017 in the *Stanford Social Innovation Review*, a book *Scaling Impact: Innovation for the Public Good* published in 2019 and a practical guide *The Scaling Playbook* published in 2020. Reflecting the diversity of contexts, programming and approaches to scaling across the Centre, the Scaling Science study proposed four guiding principles to be considered in scaling impact: justification, optimal scale, coordination and dynamic evaluation.

The Scaling Science work was presented at various internal events and POEV initiated a working group, known as the 'Scaling Science Critical Friends' to select and review case studies, which provided space for cross-organisation learning. In 2020, IDRC hosted a

scalingXchange, bringing together a group of southern 'Scaling Advisors' to exchange learning and experience in efforts to scale for the public good.[4]

Multiple staff interviewees and survey respondents recognised the Scaling Science work (and POEV in particular) for helping them to develop their understanding of scaling. Other interviewees had different understandings of scaling, were not familiar with the Scaling Science work or did not fully understand the key concepts of scaling impact versus other types of scaling.

In interviews with external stakeholders engaged in scaling, we heard that IDRC is considered a key contributor to thinking about scaling. Lessons from the Scaling Science initiative have been used by other organisations, including the International Maize and Wheat Improvement Centre, to inform the development of practical tools for project selection, monitoring and development of scaling potential (IDRC, 2019a). According to a staff interviewee, the ideas from the Scaling Science initiative were also commended by a former Deputy Chief Scientific Adviser to the UK Department of International Development (now called the Foreign, Commonwealth and Development Office). See section 5 for more detail on positioning IDRC's approach to scaling and section 3.3. for more on the role of Scaling Science in supporting learning across programs.

2.6. Conclusion

Following the introduction of the strategic objective, corporate mechanisms were put in place to facilitate staff thinking about scaling in IDRC's programming. It was never intended that all programs and projects would scale and IDRC did not mandate how programs and projects should approach scaling. Programs developed their own understanding, strategies and approaches to scaling. We believe that this flexible approach to implementation was appropriate given the diversity of IDRC programming and existing approaches to scaling at the time the strategic objective was developed. However, it also created challenges.

While some staff appreciated the flexibility, others found the lack of conceptual clarity around scaling terminology and approaches at IDRC a challenge in responding to the strategic objective. Lack of conceptual clarity or a common approach also created difficulties for monitoring progress, evaluating the results of the strategic objective and learning across programs.

Given the flexible approach taken to implement the strategic objective, the decision to create the Scaling Science initiative to learn from and across the diversity of approaches

4. Read more about the scalingXchange on the OTT website: https://onthinktanks.org/series/the-2020-scaling-x-change-building-a-field-from-the-global-south/.

was important. This work has been used outside of the Centre by other funders and researchers, contributing to realising IDRC's stated ambition 'to be recognised for sharing learning in scaling up solutions, helping position Canada as a leader in innovative approaches to development'. By promoting a principled approach to scaling, IDRC has allowed for continued diversity in approaches to scaling, while providing a common framework and terminology. More work needs to be done to communicate the results of this work within IDRC moving forward.

3. IMPLEMENTATION OF THE STRATEGIC OBJECTIVE AT THE PROGRAM LEVEL

3.1. Program design and implementation

Given IDRC's flexible approach at the corporate level, programs implemented scaling to varying extents and using different approaches and strategies. This section looks across the diversity of approaches to identify design features, practices and systems that supported or hindered scaling efforts within IDRC programs during the strategic period 2015–2020.[5] We identified six areas that emerged from interviews with program staff and grantees and in program evaluation reports and articles as being important considerations for implementing scaling ambitions: having a longer-term perspective on scaling; investing across the scaling pathway at a program or portfolio level; flexible funding; knowledge translation, and synthesis in particular; coordination structures that support collaboration and learning across portfolios; and support to grantees. We also identified challenges for implementation, particularly around the role of program staff. We also explored how programs addressed gender equity and scaling in their program design as an area of interest to IDRC. For each area, we discuss what worked well and what were the challenges.

3.1.1. Longer-term investment and thinking

KEY TAKEAWAY: Scaling impact requires time and IDRC has been experimenting with different strategies to allow more time for research results to scale: longer program timeframes, a phased approach to programming and strategic partnerships. We have been told that this has supported scaling efforts, but in many cases, the research phases of programs are not long enough for research results to scale and realise development outcomes.

A recent report by the Scaling Up Community of Practice (Kohl, 2021) considers it good practice to design projects with the understanding that scaling is a 10–15-year process. This was echoed by several interviewees – across IDRC staff, grantees and external stakeholders – who told us that scaling takes time and thus short funding cycles are a barrier to achieving and demonstrating impact at scale. The same was reflected in IDRC program evaluations. For example, an analysis of the contribution of CIFSRF Phase 2 projects to food security noted that, 'while pilots can generate interesting and valuable

5. This section draws on interview and documentation data from all programs looked at within this evaluation (total number of programs), which includes data from all four evaluation case studies. However, it draws more heavily on findings from two case studies: field building for scale and programming for scale. All programs were purposefully selected for the case studies as best case/good examples to explore the case study questions.

results within 3 or 4 years of operation, refining these into working models that can operate at scale usually takes longer: a decade may be necessary' (Wiggins *et al.*, 2018). Similarly, IMCHA's summative evaluation highlights the long timeframes needed to achieve policy influence, concluding that the final year of the program was too early to assess the policy impact of most projects at scale (Decosas, Deville & Medina, 2020).

CARIAA also recognised this time challenge and asked evaluators to identify projects that could benefit from further investment. The evaluation identified a number of opportunities for additional support to take outcomes and impacts further with more time, resources, scaling up or replication. The evaluators invited the management of the program and the consortia to develop a transparent process to extend CARIAA and to consider the identified opportunities, which included extending pilots, providing technical support to government to scale up innovations and expanding data sets. One example identified was to replicate '… research conducted by [Pathways to Resilience in Semi-arid Economies] PRISE in Kenya to update climate data and scenario for arid and semi-arid counties (21 so far), which could be expanded to the growing number of countries now considered in this expanding zone (46 in total)' (Lafontaine *et al.*, 2018, p. 68).

Similarly, respondents in the field building for scale case study suggested that it takes 10–15 years of investment and support for a field to mature. The case study further highlighted that scaling is more likely to happen in a more mature field, where the eco-system to support scaling is more developed. For example, the introduction of a tax on sugar-sweetened beverages in South Africa was informed by IDRC-funded research but that research built upon IDRC's support to an emerging field of work on non-communicable diseases (NCDs), primarily in Latin America. The IDRC East African regional office is now helping build the field for research into NCDs and sugar tax in other parts of Africa. This example shows how field building can start in one region and slowly influence or develop in others. But it takes time.

Another example, from the field building for scale case study, that illustrates the importance of a long-term perspective is in the field of digital education, mentioned by one respondent. IDRC-funded research in the early 2000s contributed to the adoption of the '1-laptop per child' policy in some countries in Latin America, but it was soon realised that the devices did not help the children if there was no material for learners and teachers to use on the devices. Learning from failure, IDRC then invested in developing electronic games for mathematics education. In 2020, the context changed and there was a greater push for digital education because of Covid-19. This created an opportunity to scale the methodologies developed in Latin America; the respondent spoke about eight countries that were interested in the methodology. These methods emerged out of 18 years of work and investment in digital education in Latin America, and they are now scaling. This example illustrates the value of early investment in a field

and being able to take advantage of opportunities for scaling as they emerge. It also emphasises that when scaling grows out of a field building effort, it could even extend the timeframe beyond 15 years, as time is needed to strengthen the ecosystem or enabling environment for scaling research results. This is necessary when scaling research results in an emerging field, or when building the field in a new location. The field might need strengthening at different points in the scaling pathway to support scaling research results, which takes time. We expand upon this in section 3.1.2.

Three characteristics of or approaches to program design and implementation emerged from interviews and document reviews as responding to this challenge: longer programs to allow more time for implementation; a phased approach to program design to continue projects that show promise; and strategic partnerships with other funders to continue scaling research results beyond IDRC's investment.

Longer program periods

IDRC staff from across various programs told us that longer time-frames are an important feature for programs that intend to support scaling. Several of the programs we looked at in this evaluation were 7+ year programs, and interviewees suggested that this was intentional because of the programs' complexity and the problems they were trying to address. In the five programs we looked at in the programming for scale case study, the length ranged from 4–9 years, with 3 out of the 5 programs being 7+ years.

However, even with the longer programs, we heard from grantees and IDRC staff that the time it takes to set up large partnerships means that *project* implementation time is still relatively short. Thus, in the cases of the CARIAA and IMCHA (7-year programs), project implementation periods were 4-5 years. Similarly, the CIFSRF final evaluation noted that for the complex scaling up that CIFSRF aimed for, at least 1 year was needed for partnership building to develop trust, forge mutual understanding and clarify goals, roles and responsibilities (O'Neill & Manchur, 2018). Furthermore, while many of the programs we looked at were 7+ years, allowing for 4–5-year project implementation periods, this was not the case across all projects with intentions to scale. In the 2018 mapping study of 50 IDRC projects that expressed intentions to scale in project approval documents was 2.6 years (Sanchez-Swaren, 2018).

A phased approach

Some programs and projects used a phased approach to continue support over a longer period where research showed promise for scaling. A few illustrative examples include:

- The Digital Learning for Development (DL4D) project Phase 1 (2015–2018) sought to improve educational systems in developing countries in Asia by testing

digital learning innovations. DL4D Phase 2 (2018–2021) was introduced to scale the innovations that worked.

- CIFSRF took a two-phased approach, in which a subset of grantees in the first phase received grants in the second phase, allowing IDRC to fund the more promising projects (Cathexis, 2019). In total, 11 of the 18 projects in Phase 2 built on innovation[6] concepts tested in Phase 1 (O'Neill & Manchur, 2018). One example of a successful Phase 2 project is that of double fortified salt in India (Diosady, Mannar & Menon, 2018; Wiggins *et al.*, 2018).
- The Growth and Economic Opportunities for Women program (GrOW) Phase 1 funded 14 research projects in 50 countries (primarily South Asia and sub-Saharan Africa). Building off the insights from the first phase that emerged from synthesis work, the focus in Phase 2 is on a smaller set of projects with a narrower research agenda on gender segregation, unpaid care and women's collective agency in just five countries in East Africa.
- In the Livestock Vaccine Innovation Fund (LVIF), projects received 18 months of funding and then projects that show promise for scaling were given a further 18 months.

Some programs and projects used a phased approach to be able to continue support over a longer period where research showed promise for scaling.

Strategic partnerships with other funders

In interviews, we heard that there are limits to what the Centre can support in terms of scaling impact within its research for development mandate. As one IDRC staff respondent explained, 'We are expanding out and thinking of partnerships, but we are aware of our mandate – we don't have funding to continue at scale' (SSI, IDRC staff). In addition, one grantee respondent said, 'They [IDRC] have funding restrictions – they cannot fund implementation and scale up is about implementation which needs more resources' (SSI, grantee).

While IDRC's mandate includes a wide range of activities from early new knowledge creation to knowledge translation and implementation research, achieving impact at scale requires collective action with actors who have complementary mandates and interests. Programs are recognising this and using partnerships with other funders strategically to support scaling beyond what IDRC's investment alone can achieve. This

6. CIFSRF defines innovations as technologies, products or models that address a specified need and are either new inventions or old ones that are applied in new ways.

can be done through funding partnerships, leveraging partner relationships and partnering with other actors.

Funding partnerships: The five IDRC programs examined in the programming for scale evaluation case study included both co-funded and parallel-funded partnerships.[7] We heard from case study interviewees from a number of programs that co- and parallel-funding models supported scaling in several ways, including being able to fund larger programs in which approaches could be tested and learning shared, and looking across a portfolio to identify opportunities to scale.

We saw examples, through funding partnerships, of IDRC supporting new programs that build on the lessons and results of previous investments. Climate and Resilience (CLARE) is an example of a program that involves a funding partnership to support longer-term investment for scaling. The CLARE program, which is being designed with the UK Foreign, Commonwealth and Development Office (FCDO) to build on lessons from CARIAA and the Climate Change Adaption in Africa program, is supporting a series of projects stemming from previous programs to follow through on potential for scaling results and promoting uptake.

Leveraging partner relationships: IDRC's strong partnerships with other funders – built partly through the co- and parallel funding models – have been leveraged to support continued funding for scaling beyond initial IDRC investments. For example, one grantee respondent expressed appreciation that IDRC actively makes connections between grantees and other funders for future funding to scale up activities after the project ends. However, this was not always possible, and another grantee expressed disappointment that activities ceased when funding discontinued: 'They have funding restrictions – they cannot fund implementation and scale up is about implementation which needs more resources' (SSI, grantee). Another grantee told us:

> *Although in our case IDRC programme officers have sought co-funding in order to try and scale successful research, data and knowledge generation, but this has been ad-hoc and often collapsed […] IDRC could create either an exchange space whereby grantees in need of specific support for scale up can resort to get experiences from other grantees, or access potential funders (who fund aspects that IDRC think are important but that IDRC as a funder does not fund). (SSI, IDRC staff)*

7. Food, Environment and Health (FEH) was the only 'core' program out of the five programs included in the case study sample, and engaged in a mix of co- and parallel funded partnerships. The rest of the programs were co-funded partnerships. Co-funding is where a program is part funded by IDRC and part funded by other donors, but where IDRC administers all funding to grantees. Parallel funding is where a program is part funded by IDRC and part funded by other donors, but where each donor administers the funding it provided.

Hence, while funding partnerships seem to have happened more systematically at a corporate or program level, leveraging this kind of funding for grantees at the project level is typically up to the propensity of the responsible officer.

Partnerships with other actors (besides funding partners): Partnerships are also used to support scaling beyond IDRC's mandate – emphasising the need for scaling to be a coordinated effort among different actors with different remits, capacities and expertise. One example of this is the Teacher Professional Development at Scale (TPD@scale) Coalition, for which the Centre is providing technical assistance to a government partner to scale education innovations (see Chapter 4 for more on coordination).

> *We work directly with the government, what we are doing is more like technical assistance to the government. Government does the implementation – the Ministry of Education doesn't think about scaling – they have schools, and they need to reach more learners. And what I like about the approach in the teacher professional development work, is that our own costs end up being very low. We provide research technical assistance, we provide models and tools, and the government pays for the scaling, and we may pay for some pilots in certain places. The only way we are going to get tech innovation in scale is for government to implement it, and the entry point has been in government. (SSI, IDRC Staff)*

Another example comes from the private sector engagement and scaling case study, in which IDRC collaborated with the Climate Technology Initiative Private Financing Advisory Network (CTI PFAN) to explore whether and under what conditions private sector financing could be mobilised to scale climate adaptation initiatives. In so doing, the project developed and shared a 'pipeline of bankable climate adaptation projects' with CTI PFAN's network of impact investors (SSI, IDRC Staff).

3.1.2 Investing strategically across the portfolio to support scaling

KEY TAKEAWAY: Programs have found that within their portfolios they need to invest in both the supply and demand side of the scaling pathway simultaneously – building the scaling eco-system at both ends. Portfolios have also been leveraged to support scaling, for example, by testing solutions in different contexts, learning across projects, syndicating efforts and identifying opportunities to scale. This requires different sizes of investments, depending on a number of factors such as the nature of the research project, the stage of scaling in the pathway, the maturity of the field and the level of risk.

IDRC's book *Scaling Impact: Innovation for the Public Good* (2019) recognises a portfolio approach as 'a strategic means of coordinating multiple innovations to optimise impact and opportunity' (McLean & Gargani, 2019, p. 66). It may lead to greater overall impact by 'syndicating efforts' around a common purpose and 'incremental change can be leveraged when multiple innovations are coordinated to work together' (ibid).

All five programs examined in the programming for scale case study took a portfolio approach – ranging from 28 to 107 projects. For example, the FEH program worked to scale up research results of specific interventions for disease prevention across geographical locations, and CARIAA supported collaborative research on climate change adaptation across 'hot spot' regions, developing and testing new analytical approaches, evidence and innovative opportunities for potential scaling up and out. In addition to the potential of greater overall impact by syndicating efforts and leveraging incremental change, a portfolio approach has allowed these projects to test approaches and then look across projects to identify where there is the most opportunity to scale.

> ***In a young field, if you focus all the investment in building up the knowledge base, then there may not be demand built up to use the knowledge when it has been developed. Programs have started to think about the scaling pathway less as a linear process and more as something that has to occur simultaneously.***

Another way we have seen programs using a portfolio approach to support scaling is to invest in projects at both ends of the scaling pathway – the supply and demand sides (see Figure 1). Two IDRC respondents (one survey respondent and one interviewee) mentioned LVIF as a good example of this, with one highlighting that LVIF is '… developing an interesting innovation pipeline/ecosystem (with funds to support project teams in delivering innovations)' (Survey, IDRC staff). The need for programs to think

about the scaling pathway not as a linear process but as something that has to occur simultaneously emerged most prominently in programs working on field building (such as LVIF and FEH), as they found that in a young field, if you focus all the investment in building up the knowledge base then no demand is built up to use the knowledge when it has been developed. Hence, there are signs that some programs are realising the need to strengthen the eco-system for scaling at both the supply and demand ends of the scaling pathway, and they are leveraging investments across a portfolio to do this.

> *Field building is often spoken about as the left-hand side of the pipeline [scaling pathway], but we don't build the other side - which is going to take our research and take our ideas. We are doing a lot of product development, but we are really looking at the two big valleys of death - developing the product and delivering the product – and there is field building to be done on both sides. (SSI, IDRC staff)*

An example of how IDRC has learnt about the importance of investing in both the supply and demand side for achieving outcomes at scale comes from the Open Data for Development (OD4D) program. OD4D was initially a two-phased program, (2015–2017 co-funded with the World Bank, Global Affairs Canada and the United-Kingdom's Department for International Development, and 2018–2020 with funding from the William and Flora Hewlett Foundation and Global Affairs Canada). A summative evaluation of the first phase of the OD4D program found that its intended outcomes on the demand side lagged those on the supply side because the program did not invest enough or for long enough and did not focus sufficiently on the transition to scale (Acevedo Ruiz & Pena-Lopez, 2017). Given that the field was in early stages of maturity, the evaluators concluded it was reasonable for OD4D to invest in the supply side initially, but more focus needed to be paid to the demand side and possibly earlier in the process. In response to this evaluation, Global Affairs Canada, the Hewlett Foundation and IDRC have invested for a further 22 months in a third phase of the program (2020–2022), with a focus on addressing sector specific demands, through further research and the provision of technical assistance to governments for the scaling of activities in least developed regions.[8]

Portfolios can be built strategically to spread investments across the scaling pathway – building demand for the new knowledge or innovation as it is being developed; and

8. Open Data for Development (od4d.net)

different sized grants within a portfolio provide flexibility to support scaling as necessary.

IDRC is still experimenting and learning about the need to invest along the scaling pathway and how to do this strategically. In the OD4D example, the program learned about the need to expand its investment from the supply side to the demand side over a number of phases based on evidence (including evaluation results). The LVIF example shows a more deliberate attempt to invest strategically along the scaling pathway to speed up the time to scale results.

Another factor raised by staff in relation to scaling and portfolio programming was that it is helpful to have flexibility to fund different sized grants within a portfolio, based on the risk, the stage of innovation, the maturity of the field and other factors. One respondent specifically spoke about designing large programs for funders who want to make large investments, but big programs can be 'like tankers' in that they are not flexible or adaptive. The respondent suggested a portfolio approach could be used to break programs up into more, smaller yet inter-connected investments. CLARE is now being designed to have a range of different sized investments, drawing on lessons from CARIAA, which had four large consortia that proved to be somewhat inflexible.

3.1.3. Flexible funding

KEY TAKEAWAY: We heard from staff that flexible funding was one of the main ways they felt they were able to support scaling by enabling grantees to take advantage of emerging opportunities. The main flexible funding strategies reported by respondents were synergy and opportunity funds, rapid response funds and IDRC's willingness to be flexible with program plans. However, staff also identified flexible funding to support scaling as a challenge in the staff survey, suggesting that not all staff are aware yet of the tools available and how they can be used.

Scaling is a dynamic process that takes place in complex systems (McLean, Gargani & Lomofsky, 2020). Several interviewees spoke about the challenges of planning for scale, and the need to adapt and be flexible:

> *In a life span for a project that lasts several years, it's very hard to plan in detail for scale especially in a context where technology moves fast. You have to jump at opportunities, these are things that happen unexpectedly and you need to be ready. At the inception stage, you have no way of knowing if the opportunity will arise. It will probably happen without you having thought about it. (Focus group discussions, grantee)*

We consistently heard from IDRC staff that flexible funding was one of the main ways they felt able to support grantees to scale; IDRC staff and grantees told us that the ability to redirect funds to take advantage of emerging opportunities or to provide no-cost extensions were important mechanisms to support dynamic scaling processes.

Examples of flexible funding identified by the evaluation team include:

- **Opportunity and synergy funds/ grants** are additional grants given to grantees to extend or expand their work, or to take advantage of emerging opportunities (Cundill *et al.*, 2019). Interviewees from CARIAA and IMCHA identified these as supporting their scaling efforts:
  - In CARIAA, **synergy and opportunity funds** were built into the program design as an adaptive management tool to help consortia organise collaborative work outputs (Lafontaine *et al.*, 2018). This included taking advantage of windows of opportunity for policy engagement (SSI, IDRC staff).
  - In IMCHA, **synergy grants** were awarded to carry out supplementary research activities on a theme related to their existing IMCHA work. Although no explicit connection was made between synergy grants and scaling in the final evaluation of IMCHA, the evaluation did find that these grants had allowed selected research teams to expand the scope and depth of their work. They had also allowed research teams to work more specifically on gender and equity issues or leadership and management issues that were missing in the original projects (Decosas, Deville & Medina, 2020).
- **Rapid response funds** were mentioned by a number of respondents in the field building for scale case study as an effective mechanism to respond to policymaker demands within shorter timeframes than traditional academic research allows. Having funds to react fast to opportunities helps build relationships with policymakers, supporting demand-driven research and coordination for scaling research results. As one IDRC staff member said:

> *Government thinks researchers are guys who are floating in a cloud and are not useful, so I have to convince them that these researchers are tough consultants and will deliver quick and dirty solutions in 24 hours – I have money for rapid response mechanisms, so that government can come with problems and we can give it to the researchers to come with some solutions. (SSI, IDRC staff)*

Despite staff interviewees citing flexible funding as one of the biggest enablers of scaling, in the evaluation staff survey it was the third most frequent challenge reported when we asked what are the key strengths and challenges of IDRC's response to the strategic objective to scale. The evaluation team did not find any corporate guidance for responsible officers on flexible funding mechanisms (or on building flexibility into

programming more generally). Together, this suggests that not all staff are aware of these mechanisms.

In interviews, IDRC staff and other funders frequently raised challenges and limitations to using flexible funding for institutions with public funding. One external stakeholder to a large IDRC program recognised this challenge when implementing flexible processes and systems:

> *It is challenging … IDRC faces challenges as a quasi-government institution, they are accountable to taxpayers, and this puts limits on their funding. Their spirit and ethos are flexible, but there are some inflexibilities. (SSI, external partner)*

Clearly there is a balance to be struck here between flexibility and accountability, and something for IDRC to consider as it builds programs to support scaling research results.

3.1.4. Knowledge translation and research synthesis

KEY TAKEAWAY: Two aspects of knowledge translation that programs identified as important for scaling are research uptake and synthesis; the evaluation found that research uptake is well supported but synthesis less so. Both activities can promote the transfer of knowledge vertically beyond the immediate research users and horizontally among research teams. To maximise the value of having a portfolio approach, it is important to do synthesis and find opportunities for scaling. However, synthesis across a program needs to be strategically led by programs to help meet scaling objectives. Programs are building in knowledge translation from the beginning and have leveraged this for scaling but capacity to support this work from within programs remains a challenge.

At IDRC, the terms 'knowledge translation', 'communication' and 'synthesis' are often used interchangeably, which can cause some confusion. IDRC's book *Scaling Impact* defines knowledge translation as a 'dynamic and iterative process that includes synthesis, dissemination, exchange and ethically sound application of knowledge' (Canadian Institutes of Health Research, in McLean & Gargani, 2019).

For this evaluation, we refer to **knowledge translation** as the dissemination and exchange of knowledge and other activities to support research uptake (or research into use as it is sometimes referred to). **Research synthesis** is a distinct (but related) activity that is about integrating research results from diverse sources pertinent to an issue, 'The aim of synthesis is to increase the generality and applicability of those findings and to develop new knowledge through the process of integration' (Hampton & Parker 2011, Magliocca et al., 2014, Baron et al. 2017 in Wyborn *et al.*, 2018, p. 1). This is done in many ways at IDRC, but primarily by bringing information together among research

teams or across a portfolio (horizontally), and then communicating to different audiences depending on its purpose. Synthesis can also consolidate research findings from across a body of evidence (i.e. beyond IDRC funding) – like a systematic review.

All five programs examined in the programming for scale case study had knowledge translation mechanisms to support **research uptake** and use beyond immediate users. They developed knowledge exchange, knowledge translation and/or communication strategies and platforms. Some illustrative examples include:

- Knowledge translation was central to the scaling strategy of IMCHA, which was achieved through strategic partnerships with policymakers, and providing training on knowledge translation to all grantees.
- CIFSRF's third objective was to use research results to inform food security policies and programs; the CIFSRF final evaluation reports that at program level, partners and staff organised 49 knowledge-sharing events targeting policymakers (O'Neill & Manchur, 2018).

A key design feature of the Knowledge and Innovation Exchange (KIX) program is linking knowledge and innovation with national level education structures via regional hubs, putting in place a national delegation comprising five key education stakeholders from the Ministry of Education and local education groups for each member country. This is intended to strengthen the likelihood that research is demand driven and facilitate uptake of research findings.

Regarding **research synthesis,** the evaluation team heard repeatedly from IDRC staff that it was a particularly valuable, yet under-resourced, tool for scaling. A number of IDRC staff members and IDRC documents identified synthesis papers as important especially with larger portfolio programs, as they provide an overview of a particular body of knowledge, helping to identify gaps and opportunities to scale research results. Synthesis work also helps build a critical mass of knowledge from disparate research. This can inform IDRC investment decisions to address evidence gaps or advance a research agenda with significant emerging results. We provide some illustrative examples below.

GrOW in particular used synthesis for scaling research results, and had a specialised program officer position that focused on knowledge translation. Synthesis was used to inform the investment decision for GrOW 2, with a more targeted research agenda and a view to scaling. The final GrOW report highlighted that synthesis enabled them to identify key lessons and challenges to inform policy, program design and monitoring measures (IDRC Grow Team, 2018). As one IDRC staff interviewee said:

As we moved from the first to the second phase, we realised we had a critical mass of knowledge and evidence, we have identified the key barriers and tested some solutions, and now we can work with other funders to pool resources together and focus on those solutions that are promising, scalable and see how we can make a mark on a narrower set of issues…Without the synthesis the knowledge is there but it does not become apparent what the next steps are.
(SSI, IDRC staff)

CARIAA invested in collaborative research synthesis as part of its research into use agenda. CARIAA's focused study on its research synthesis approach found that generating syntheses that highlight new frontiers at the climate–policy nexus are the cornerstone of CARIAA's focus on climate change hotspots. Synthesis activities strengthened the impact of the research, specifically by supporting better-informed policy and practice. It included a broad array of activities such as the creation of collaborative spaces and research outputs including academic papers, policy briefs, blogs, videos, maps, conference panels and media articles. It also took place at multiple levels – project, country, consortium, theme, program and international. Creating collaborative spaces required significant resources and time (Cochrane *et al.*, 2017).

Responsible officers can use synthesis strategically to make decisions about scaling across their program portfolios, but seldom have the time, capacity, motivation or incentives to do so.

Cochrane *et al.'s* (2017) reflection on CARIAA's experience found that synthesis can be used strategically to support the attainment of program objectives. A key lesson is that the program management team, which had the opportunity to look across the four consortia, needed to take a greater thought leadership role in leveraging its network and in providing resources to lead collaborative synthesis. The authors also suggested that program leaders could have used synthesis more strategically by situating collaboration and synthesis within the program objectives, to help participants develop a shared vision and understanding of these objectives. This is an important insight – responsible officers can use synthesis strategically to make decisions about scaling across their program portfolios, but seldom have the time, capacity, motivation or incentives to do so.

One IDRC respondent told us that although it helped to have dedicated knowledge translation officers who could support research synthesis work, their primary focus was on research uptake and communication. Another respondent told us that the main measure for responsible officer success is how many projects they have launched, and

that they are not thinking about opportunities for integration and learning across their project:

> *… there has to be a culture change in how POs understand their job description…. and they will need to ask different kinds of questions and to rather see them [individual projects] as cohort of projects.*
> *(SSI, IDRC staff)*

IDRC has recognised this challenge and established a knowledge translation unit within POEV to support programs with this function. Ideally this will increase the capacity of responsible officers to engage more strategically with their portfolios through synthesis. The evaluation also found that embedding knowledge translation officers in programs, such as in GrOW and LVIF, was a good solution.

3.1.5. Program coordination

KEY TAKEAWAY: IDRC programs are designed with internal coordination structures that support projects to collaborate, learn and synthesise results. Structures vary to meet the specific program ambition – including its scaling ambition. Coordinating multiple stakeholders, often across multiple geographies, is inevitably time and resource intensive and can be cumbersome. Building in flexibility is important.

At the program level, coordination is required across a portfolio of projects. IDRC programs tend to be designed to support collaboration, learning and synthesis across projects (Cathexis, 2019). A 2019 meta-review of program evaluations identified different types of coordination structures: consortia, hotspots, researcher–decision-maker research teams, networks, cohorts and south–north research teams. These coordination structures help facilitate much of the knowledge synthesis work highlighted in section 3.1.4, and to facilitate multi-stakeholder collaboration. In the programming for scale evaluation case study, each program had a different coordination structure designed to support cross-program collaboration and learning, as well as meet the needs of program-specific context and ambition. For example, IMCHA was structured with two regional consortia made up of research and policy actors with the ambition to focus on research uptake, whereas CARIAA was structured with four consortia organised around climate change hotspots with the ambition to find solutions that can be scaled.

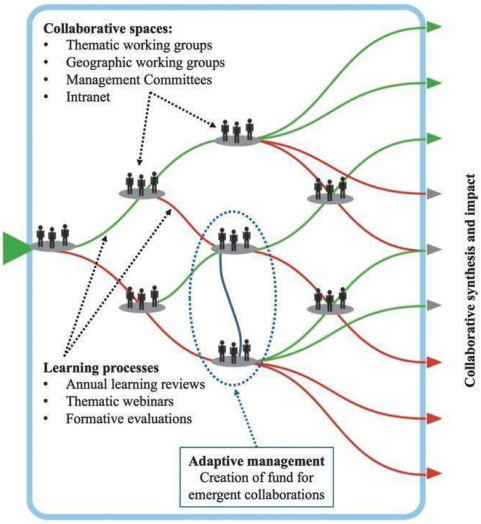

Figure 5: CARIAA collaborative spaces. Source: Investments in process design for transdisciplinary collaboration in CARIAA (Cundill et al., 2019)

CARIAA is a good illustrative example of a program designed with structures to coordinate and operationalise collaboration among projects in geographically and culturally distant teams.

It created formal collaborative spaces, held regular learning processes and created a fund for emergent collaborations that supported the program's adaptive management approach (see Figure 5).

These are useful design considerations for any large scale, trans-disciplinary research endeavours because they support the coordination of the complex web of actors involved in scaling processes, while facilitating collaboration and learning (Cundill *et al.*, 2019).

In CARIAA, collaborative consortia are organised around climate change hotspots to take account of the intersections among ecological, physical and socio-economic systems. This hotspot approach was an intentional design feature to support collaboration and scaling (Cochrane et al., 2017; De Souza *et al.*, 2015 in Cochrane & Cundill, 2018).

Coordination is not without challenges. CARIAA found that a flexible approach is important to support scaling as it allows people and organisations to come together around interests, enthusiasms and timelines in a way that could steer collaboration structures. According to one interviewee from the CARIAA program, some groups 'took off' while others 'died', highlighting the need to be able to adapt the approach when something is not working.

We heard in staff interviews that coordinating too many people and organisations at the program and project level makes decision-making across coordination structures cumbersome. As such, programs and projects become less nimble and responsive to opportunities to scale research results.

3.1.6. Support to grantees

KEY TAKEAWAY: IDRC staff have helped grantees incorporate scaling into their project design and implementation. Discussions about scaling appear to be happening early in the project design process for most grantee survey respondents, and grantees we interviewed appreciated IDRC's hands-on Grants-Plus approach. According to survey respondents, IDRC program staff help them think about different considerations for scaling, although discussions about optimal scale or the negative effects of scaling are less frequent than other considerations for scaling.

IDRC's Grants-Plus model[9]

The grantees we spoke to in interviews were overall appreciative of IDRC's support to them in scaling results throughout the project, describing IDRC staff as 'hands-on' and that they 'walked alongside' research teams, they were 'friendly', 'had the right attitude' and 'understood the context' as well as 'knew how to do research', provided 'useful critique' and were a 'critical friend' (SSIs, focus groups and survey, grantees). On occasion, IDRC involvement extended to stakeholder meetings and even field work. One grantee said IDRC was 'more than a funding agency, they were part of the team', while another said that IDRC was not only a financial partner but a technical partner. This has been referred to in the past as the Grants-Plus model (IDRC, 2011, p. 5).

Grantees also highlighted the usefulness of IDRC-sponsored resources and activities that go beyond research – such as knowledge translation materials, convening of workshops (such as regional workshops), and the role of skilled trainers and facilitators during such events – all contributing to enhanced knowledge and practice in relation to scaling research results.

Considerations for scaling

The survey data provides a broader look at how IDRC staff supported grantees to scale, with 95 grantee respondents from across different programs. The majority of survey respondents were engaged and informed about scaling, with 79% (n=95) rating themselves as very or somewhat knowledgeable about scaling. The majority of respondents (62%) had scaling intentions right from the start of their projects, while 15% said that scaling became part of their discussions with IDRC during the design phase (see Figure 6).

9. The 'Grants-Plus model' is used to describe IDRC's approach as a grant maker, which goes beyond just providing financial support. An IDRC Annual Corporate Evaluation Report for 2010–2011 describes the Grants-Plus model as 'opportunity, engagement and access' (IDRC, 2011, p. 5).

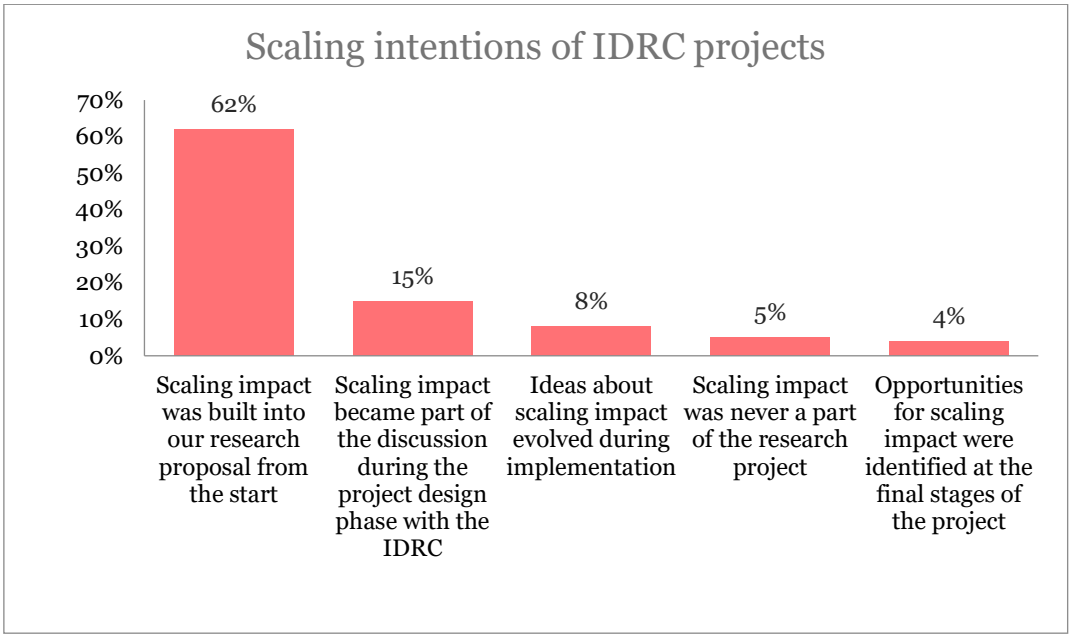

Figure 6: Grantee intentions for scaling in their projects (n=90)

IDRC's book *Scaling Impact* (McLean & Gargani, 2019) recognises inclusion of scaling discussions early on in the project as critical and as one of the key components of good practice of scaling identified in the report by the Scaling-up Community of Practice, especially in the design, testing and proof of concept phases (Kohl, 2021). This coincides with positive results from the grantee survey: 69% said that they had discussed scaling with IDRC in the design phase; 74% said that IDRC had helped them consider a number of different factors related to the implementation of scaling; and 74% said that *more* discussions with IDRC about scaling would have been useful.

In line with IDRC's principled approach to scaling, survey responses from IDRC grantees reflected that they were thoughtful about applying criteria on whether the results of research should be scaled. Only 6% of grantees said that they would be happy to scale research results based purely on successful research outcomes. Most grantees (69%) said they had detailed (18%) or basic (51%) criteria for thinking about whether an initiative that showed positive research results should scale.

IDRC frequently helps grantees think about consultation with user groups, but there is room for further discussion about optimal scale and possible negative impacts.

We also asked respondents whether IDRC played a role in helping them think about whether an innovation could scale, if it should scale, at what scale results would be

optimal, consulting user groups about the research design and impact, and the possible negative effects of scaling. Respondents most frequently mentioned that IDRC had played a role in helping grantees think about consultation with user groups (76%). They less frequently noted that IDRC had played a role in helping them think about optimal scale (66%) and the possible negative impacts of scaling (55%). This indicates room for further discussion about optimal scale and possible negative impacts.

3.1.7. Program staff roles

KEY TAKEAWAY: Scaling requires responsible officers to think and act more strategically and opportunistically. Their role is transitioning from funder and technical partner to knowledge broker, knowledge translator, coordinator and strategic thinker. However, there is not yet a formal recognition of this change in role and some responsible officers feel that they do not have sufficient time, resources or incentives to carry it out effectively.

As highlighted in the previous section, grantee interviewees acknowledged and appreciated IDRC's Grants-Plus approach in supporting their efforts to scale. Several program staff interviewees told us that the focus on scaling within programming had brought new responsibilities to their role, as this quote illustrates:

> *[IDRC is] doing things in a collaborative way which is a big shift – if the ambition is impact at scale and projects working at multiple scales simultaneously then POs need to work completely differently – they play more of a relationship management role, they are knowledge brokers and putting in early warning systems for conflict. (SSI, IDRC staff)*

Changes to the roles and responsibilities of program staff have been recognised elsewhere within IDRC: a recent strategic review of IDRC's Climate Change Program points to a possible '…evolution in the role of IDRC Program Officers, shifting from technical experts and subject area specialists supporting grantee capacity, to knowledge brokers and intermediaries tasked with creating an enabling environment for innovation to emerge' (Harvey *et al.*, 2019, p. 26).

However, a number of staff interviewees suggested that this changing role was not well recognised in IDRC's corporate systems. To verify this, the evaluation team reviewed three job descriptions for program officer, senior program officer and senior program specialist, and found that they mostly cover technical grant administration duties. Coordination was listed as a program officer responsibility, but it was specifically related to 'coordination of research activities of collaborative projects between Canadian and developing country researchers', specifying that their responsibilities are in relation to supporting knowledge uptake, dissemination

and policy influence. There was no mention of relationship management or brokerage, or of supporting scaling.

The findings in the previous sections also suggest that scaling requires responsible officers to be more strategic and adaptive in their programming – thinking about scaling across the portfolio, identifying opportunities for scaling, responding with flexible funding and providing strategic direction for synthesis.

Many of the IDRC program staff we spoke to are motivated by this new role and have embraced the complexity that comes with managing scaling initiatives, as one respondent said:

> *We are relatively small and innovative and the people hired at IDRC are curious and creative and looking for new solutions to new problems, or to emerging problems. (SSI, IDRC staff)*

Responsible officers frequently raised time to be strategic as a major constraint. They recognise that they have little time as they are caught up in grant administration, exacerbated by short funding cycles – which more than one interviewee referred to as 'getting money out the door'. Numerous respondents mentioned that they would be more effective as grant managers of scaling initiatives if they had more skills and time to think and act strategically.

The following quote from an IDRC staff member illustrates this point:

> *In my opinion, more energy, effort and attention should be put to sustained dissemination of these knowledge products/learnings/experiences. However, rarely is there enough time or opportunity as there is always the feeling that the next priority needs to be tackled… This is at the expense of being able to put adequate effort and priority to knowledge translation and dissemination which are key ingredients for achieving impact at scale. (Survey, IDRC staff)*

Others noted that not all responsible officers understand their role in this way and prefer to do grant administration.

There were limited references in the interviews to how programs tried to address these capacity challenges. One option is to have a specialised PO in the grant management team, such as a Knowledge Translation Officer, (a program officer whose primary duties and responsibilities relate to research uptake and communication). Two such examples are from GrOW and LVIF. It could be that externally funded programs have more resources to bolster their teams with specialist roles. In fact, one person in the IDRC staff

survey alluded to this when saying, 'Outside of EFPs, I don't think we had capacity or granting options that were flexible or appropriate for supporting impact at scale'.

Another way that programs have tried to address the capacity challenge is to strengthen the coordination capacity of grantees, or through grants that support a coordination function. A respondent from CARIAA said that they had ensured that there was a coordinator function in each of the four research consortia. IMCHA created coordinating bodies to bring research teams together with each other and with policymakers, they were also responsible for capacity strengthening and knowledge translation – these were called Health Policy Research Organisations (HPRO); one was in East-Africa and one in West-Africa. HPROs also helped ensure the right stakeholders engaged with research teams and convened stakeholder forums to facilitate engagement with stakeholders throughout the research process.

Another key challenge faced by responsible officers related to their understanding of scaling concepts, which inhibits their ability to engage with grantees on scaling matters. This also came up in staff interviews and in the staff survey, with 56% (n=43) of respondents identifying 'understanding concepts of scale' to be a challenge in responding to the strategic objective, as the following quote illustrates:

> *Understanding the concepts of 'scaling' was difficult – how to do it, who to do it, how to measure it was not clear, so it was not well defined in terms of whether scale meant the number of people reached, the technologies developed or whether it was looking at the impact on the people reached, impact of technology on a certain scale – it was not clear what the useful measure would be. (Survey, IDRC staff)*

While there was little guidance on IDRC's scaling work in the first part of the strategic period, respondents do recognise POEV's work to understand scaling, with the publishing of IDRC's book *Scaling Impact* in 2019 and the playbook in 2020. These will be useful to guide the conversation about scaling in IDRC going forward:

> *I want to celebrate the work of POEV team with the theme of scaling and how, in parallel to programs, they have led the thinking and unpacking around how we have tackled scaling definitions and understanding – it allows for diverse, inclusive and coherent understanding of what scale means across programs. (SSI, IDRC staff)*

3.1.8. Gender equity and inclusion

KEY TAKEAWAY: Over the 2015–2020 strategic period, there was significant effort at IDRC to more systematically mainstream gender considerations in its research projects. While gender is a major consideration in the supply side of IDRC research that is positioned to scale, considerations of how scaling itself may affect equity and gender are less prevalent. Staff are not yet able to clearly articulate the link between gender and scaling. IDRC's guiding principles for scaling impact have provided a strong lens for thinking how to integrate gender into scaling strategies, particularly the principles of justification and optimal scale, and could be integrated more into program design.

As a funder, IDRC has long supported gender responsive and transformative research, and consideration of gender is mainstreamed across the Centre's programming (IDRC, 2019b). This evaluation explored how programs included gender in their scaling work.

First, we looked at the relationship between gender and scaling. An IDRC insights paper titled *Transforming gender relations: Insights from IDRC research* (2019b), drawing on a study prepared by Sisters Ink, highlights four ways to scale positive change through gender transformative research based on a sample of 42 gender-focused IDRC projects over the past 10 years:[10]

1. Gender transformative research helps develop gender relevant methodologies, practices or innovations that can be scaled to achieve better outcomes for women beyond the immediate users. For example, in Cairo, activists and researchers recorded incidents of sexual harassment through crowdsourcing and the data was used to influence the introduction of a new sexual harassment policy at the University of Cairo and safe spaces for men and women in the city. The 'Harassmap' model was replicated in 28 other countries.
2. Thought leaders can be important for 'scaling influence', they can use the evidence generated by gender transformative research to inform international policy or guidelines that has a multiplier effect in terms of scaling gender outcomes. For example, an IDRC research project principal investigator became a UN Special Rapporteur and used evidence from an IDRC-supported project to inform a guidance note for the United Nations Secretary-General, potentially influencing dozens of countries.
3. Scaling gender outcomes is achieved through influencing gender norms and cultural practices that inhibit social justice for women. They argue that changing

10. Sisters Ink applied IDRC's continuum of gender integration (derived from IDRC project approval documents) and found that out of the 42 projects, the majority were either gender transformative (16) or gender responsive (17). The rest were gender sensitive (7), or gender aware (2). For a description of each category, see p. 4 here: Transforming gender relations: Insights from IDRC research.

social norms produces deeper and more sustainable change, and that there were '… examples of scaling activities where the participatory, social learning processes used to engage various stakeholders and build local capacity presented real possibilities for scaling' (IDRC, 2019b, p. 20).

4. Evidence could be used to inform global systems such as industry standards or safeguarding measures. The paper gives the example of extractive industries and associated value chains that have relevant global regulations that can be influenced.

IDRC's book *Scaling Impact*, gender is mostly discussed in relation to the first two guiding principles for scaling impact: justification and optimal scale. The search for optimal scale requires balancing the different types of impact produced by the scaling process – the magnitude of impact, alongside the variety, sustainability and equity of impact. Equity of impact is about how impacts are distributed among different groups and whether they create new inequalities for certain groups or replicate or increase existing inequalities (McLean & Gargani, 2019, p. 53). Justification for scale means that the choice to scale must be shared by those who will be affected. Thus, gender and equity considerations are an important part of scaling strategies and decisions to scale.

Scaling Impact gives an example of how gender was considered in the scaling strategy of a gender justice project 'Scaling a Survivor-Centric Approach'. The project integrated gender-sensitivity into its scaling process. It ' …analysed gender barriers and articulated the role of gender at each stage (e.g., at the design of the pilot phase, during pilot testing, and development of the scaling strategy, in partnering and collaborating with key stakeholders, in monitoring and evaluation exercises, etc)' (McLean & Gargani, 2019, p. 121). Importantly, gender had to be considered each time the model was replicated in another district, and both gender transformative and gender accommodative strategies were used in the process of scaling (McLean & Gargani, 2019).

The five programs selected for the programming for scale case study were not 'gender focused' programs (like those studied in the Sisters Ink paper or the gender justice case). But all five considered gender in their programming. Box 3 describes some of the gender-related outcomes and approaches of the five programs mentioned by interviewee respondents, including capacity strengthening for researchers on gender and research.

When we asked IDRC staff and grantees how gender considerations are included in the design and implementation of scaling strategies, most respondents replied that a gender lens is applied to all projects regardless of scaling. None of the interviewees spoke specifically about gender in relation to scaling or scaling processes. For example, we were told that CIFSRF had a gender specialist working with them throughout the project, and specific guidelines were developed to help the project teams and responsible officers to incorporate gender. CARIAA also worked with a gender expert and had a gender and

inclusion working group to share gender-related lessons and approaches; and KIX is looking at equity, inclusion and gender equality in all its projects. The following quote from FEH describes how they included gender in their programming:

> *We worked with … international consultants with experience to help improve capacities and knowledge in relation to gender. We strengthened our own internal tools and learning documents we share with grantees and it is a requirement and expectation of any funding call that it must address gender. (SSI, IDRC staff)*

A 2018 CARIAA paper on gender and social equity did refer to scaling outcomes, but concluded that if '…[climate change] adaptation considers gender and other socio-cultural variables, it can better promote equality and help to improve people's wellbeing' (CARIAA, 2018).

In fact, interviewees in the programming for scale case study sample told us about the many different gender-related aspects, approaches and outcomes of their scaling interventions (see Box 3). All the examples in the box include capacity strengthening for researchers on how to include gender in research for development, and a focus on knowledge generation and to some extent on research uptake.

Thus, individuals with whom we spoke noted that considering gender within research projects can lead to more gender-equitable outcomes and impact. Therefore, if gender is considered within projects that seek to scale those outcomes and impact, then it has the potential to lead to more gender-equitable outcomes at scale. However, as IDRC's book *Scaling Impact* emphasises, it is also important to consider how the equity of impact might be affected by the scaling process itself, and hence it needs to be built into the scaling strategy.

The IDRC programs studied in the programming for scale case study all considered gender in their work, but interviewees struggled to connect this to scaling strategies, and there was little discussion on gender in relation to scaling processes or outcomes in program reports and evaluations.

Hence, although we saw inclusion of gender into the supply side of IDRC research, we do not have evidence from the programs we looked at in the programming for scale case study that gender is being considered in the demand side around decisions to scale, optimal scale and the possible negative effects of scaling for specific groups, including women. Since the Scaling Science initiative and the resulting guiding principles for scaling impact has been an emerging area of work at IDRC over the 2015–2020 strategy period, the evaluation team did not find it surprising that these concepts were not explicitly discussed in the case study programs. We were left to conclude that the case study programs, on the whole, had a comprehensive approach to addressing gender and,

to a lesser extent, other equity considerations, but that the practice of scaling has not been sufficiently developed to allow us to understand the explicit intersection between gender and scaling in program design.

There is also not enough evidence about whether women and other groups are involved in making decisions about what scale is optimal, and whether scaling itself is justified. We found little evidence that the Centre has explicit criteria or frameworks for integrating gender in relation to scaling specifically. The criteria for assessing gender inclusion and gender transformative research do not yet extend this far.

There has been intentional learning about gender and scaling during the strategic period (most notably the Sisters Ink paper and the scaling science work) and the guiding principles for scaling impact in particular offer a helpful framework and language for considering gender within programming for scale. Further socialisation of the guiding principles specifically in relation to gender equity and scaling would strengthen IDRC's work in this area.

IDRC programs have a comprehensive approach to addressing gender and, to a lesser extent, other equity considerations, but the practice of integrating gender and equity considerations into scaling strategies, particularly on the demand side, needs to be strengthened.

BOX 3: Examples of three programs' approach to gender and its outcomes

IMCHA's HPROs encouraged gender integration into the projects' research processes using three strategies: conducting formative research on the gender situation in each country, training on gender and project specific mentoring on gender. The IMCHA summative evaluation (2020) applied a Gender Responsiveness Assessment Scale and found that 12/28 research proposals (43%) were designed to be gender transformative, four were gender specific (14%), eight were gender sensitive (29%) and four were gender blind (14%) (Hera, 2020).

CIFSRF was guided by a gender strategy to ensure gender integration throughout the program funding cycle. The program also supported capacity strengthening and ongoing learning among research partners to strengthen their ability to address gender within their projects. Outcomes of these gender integration strategies occurred at multiple levels. At project level, 78% of CIFSRF projects increased women's access to knowledge, skills and resources. At program level, they led to increasing the gender analysis capacity of research teams, applying a gender lens throughout the research cycle, accounting for progress in meeting gender objectives

and generating evidence on gender-responsive research practices (O'Neill & Manchur, 2018).

CARIAA assessed its research activities according to the extent to which they adopt an intersectional perspective and incorporate gender in overall research design. Diverse groups participated in the research process and training in researching gender was provided (IDRC, 2018a). The CARIAA summative evaluation used a Gender Assessment Framework to assess outcomes which found that gender is discussed and incorporated at several scales across consortia and across hotspots. Regarding outcomes specifically, the evaluation found that gender was incorporated in the consortia's strategies, planning and logical frameworks; quality knowledge on gender drivers and conditions leading to vulnerability in the context of climate change was sufficiently generated and disseminated; gender/social disaggregated data was generated, considered and used in the research and uptake of evidence to different degrees depending on the context; and the CARIAA research community's capacities and involvement in gender have been clearly reinforced (Lafontaine et al., 2018).

3.1.9. Conclusion: Program design and implementation

A flexible approach to implementing the strategic objective encouraged programs to experiment and try different approaches and to adapt existing systems, processes and tools for scaling. As a result, IDRC has 5 years of experience to learn from.

This section highlighted seven areas that emerged from interview and document review data that have enabled programs in their scaling endeavours. Yet, given that the past 5 years have been about experimenting and learning, many of the strengths and challenges we found were two sides of the same coin with enablers yet to be fully recognised, systematised and supported across scaling programs. Here we present a summary of the enablers of scaling – which we believe are strengths to be built on in future scaling programming – and the challenges.

BOX 4: Summary of enablers and challenges of scaling

Enablers of scaling / strengths to build on

- Longer program periods and multi-phase programming that allow more time for promising projects to work towards achieving impact at optimal scale.
- Portfolio programming that enables testing and learning from across different contexts to identify opportunities to scale.

- Investing strategically in both the supply and demand side of the scaling pathway within a portfolio.
- Leveraging IDRC partnerships and staff relationships with other funders and other actors to support scaling.
- A flexible approach to programming, enabling projects to respond and adapt to opportunities and challenges – flexible funding mechanisms are a particularly useful tool for enabling this.
- Investing in research synthesis to learn across a program and identify opportunities to scale research results.
- Flexible coordination structures that enable collaboration, learning and synthesis to evolve and change according to program needs.
- Early discussions with grantees on planning for scaling and on-going support during implementation.
- Program staff embracing new challenges and activities to support scaling – and acting as knowledge brokers, relationship builders and coordinators. Experimentation with specialised functions such as knowledge translation officers and contracting out the coordination has helped bolster the capacity of program teams to support scaling.

Challenges to scaling / areas to improve on

- Scaling research results takes time – even with extended program lengths, grantees did not feel they had enough time. We also heard that support to help grantees secure further funding at the end of the program is ad hoc.
- There is little evidence that flexible funding mechanisms are being used *systematically* to support scaling. Looking ahead, it could be useful for programs to have guidance on what flexible funding options are available and how they might be used to support scaling in a way that addresses possible concerns about transparency and accountability within public-sector funds.
- Not all programs had sufficient resources or capacity for research synthesis work. Program staff often felt they did not have capacity. In one program, the collaborative synthesis was undertaken amongst researchers, but the evaluators found that the synthesis work would have better supported program objectives if it had more leadership from within programs.
- There is a lack of conceptual clarity about scaling at IDRC and some staff felt that this inhibited them from supporting grantees with scaling.
- Questions about at what scale results will be optimal and the possible negative impacts of scaling were the least considered factors in our survey with staff and grantees.

- Not all program staff see scaling and the new challenges and responsibilities that come with it as part of their job; they prefer to stick to more traditional 'grants-plus' duties of supporting opportunity, engagement and access.

3.2. Coordination for scaling

KEY TAKEAWAY: IDRC programs coordinate with a wider set of actors beyond the research community to support scaling research results but bringing together traditional and non-traditional stakeholders is complicated and resource intensive. Nurturing partnerships and relationships is important.

As part of the scaling objective, the IDRC 2015 Strategic Plan committed to working with and connecting to actors to support scaling: 'IDRC will connect solutions with actors who can help advance those solutions to achieve large-scale impact' and 'within 5 years, across its programming, IDRC will be working with public and private sector actors who can advance ideas and innovation through to large scale implementation' (IDRC, 2015). IDRC's guiding principles for scaling impact, developed through a review of IDRC programs, recognise that scaling takes place in complex systems and involves coordinating with a diverse and evolving set of stakeholders. They also refer to the need to plan and adapt for the many actors involved in bringing impact to scale (McLean & Gargani, 2019). While project-level coordination is largely led by grantees, IDRC uses program design features that coordinate the many actors involved in scaling at the program level. In this section, we look at how IDRC coordinates with diverse stakeholders to support scaling and examine more closely how programs have engaged with the private sector to support scaling.

3.2.1. Coordination with diverse stakeholders

IDRC staff consistently reported through interviews that the objective to scale encouraged them to approach coordination differently than they had in the past. As described below, there were a few keys ways that staff reported changing their coordination efforts to facilitate scaling – by engaging with stakeholders beyond research users, often those who comprise the demand side, and reflecting on what would be the best mix of actors to engage given the specific scaling ambition.

Programs are engaging with stakeholders beyond the research community: As one respondent commented: 'We are partnering with non-traditional research organisations like non-governmental organisations (NGOs) who will use the evidence for advocacy and finding/brokering/encouraging new types of partnerships that allow evidence to get closer to users' (Survey, IDRC staff).

In a review of 50 projects, all intended to engage with at least three categories of partners and most projects intended to engage with four or more.

The 2018 mapping study of a sample of 50 IDRC projects with intentions to scale analysed their plans to engage with different partners across seven categories (Sanchez-Swaren, 2018). All projects intended to engage with at least three categories of partners and most projects intended to engage with four or more. Nearly all projects intended to engage with the research constituency (49 out of 50 projects) and government agencies (46 out of 50 projects). Other partners included UN agencies, multilateral organisations or other strategic partners (37 out of 50 projects), civil society organisations (37 out of 50 projects) and the private sector (31 out of 50 projects).

IMCHA is an example of where IDRC created coordinating structures comprising diverse stakeholders. In this case, IDRC funded coordination structures in the form of Health Policy Research Organizations (HPROs). These were themselves multi-disciplinary bodies which connected researchers and government decision-makers. For example, the East Africa HPRO comprised the African Health Policy Resource Centre (a health policy think tank) and two inter-governmental organisations – the East, Central and Southern Africa Health Community and Partners in Population Development. The East Africa HPRO helped ensure the right stakeholders engaged with research teams and convened stakeholder forums to facilitate engagement with stakeholders throughout the research process (EA-HPRO, 2020).

Programs are changing how they think about the 'right mix' of actors for different scaling ambitions. Evidence from the programming for scale case study and the field building for scale case study suggested that the right mix of actors for scaling research results depends on several factors including the scaling objective, where the project is located in the scaling pathway and the maturity of the field in which they are working. Interviewees told us that the right mix of stakeholders can refer to actors from different research disciplines, sectors (e.g., private sector versus civil society organisation (CSO)), fields or geographies, and the right mix of actors can evolve over time.

The following quote illustrates how FEH engaged an evolving set of actors:

> *We are increasingly engaging stakeholders beyond researchers. At the beginning of FEH we put a lot of emphasis on research to policy and this meant moving beyond the academic arena – though we also needed a strong academic constituency and to strengthen the academic base. Since the beginning of the FEH programming we were already trying to partner with donors in order to enlarge the*

funding scope and have more impact. FEH also started to engage with other stakeholders such as a health coalition that is linked to the Inter-American Heart Foundation, and they joined the Community of Practice (COP) as the advocacy branch of the COP, and in each of their countries there is an emphasis on engaging CSOs and decision makers as much as possible [...] All these efforts were oriented to build the field of healthy food systems in LAC (SSI, IDRC staff)

Another example from FEH speaks to the importance of having the right mix of stakeholders during the design phase of a scaling endeavour. For example, the FEH portfolio of projects influencing sugary drink taxes was less successful in Barbados than in South Africa because it was driven by financial policymakers, whereas in South Africa the stakeholders included CSOs, government and academics while also using fiscal and economic levers to scale evidence into policy (SSI, IDRC staff).

Bringing together stakeholders from different disciplines and sectors into partnership formations can help harness efforts and resources within a system, but it also presents challenges and adds to the complexity of scaling at the program and project level.

The benefits of working with actors beyond the research community was highlighted by the CARIAA summative evaluation which found a few cases where research teams from universities worked closely with NGOs and international NGOs, and acknowledged positive benefits from this collaboration. The evaluation recommended more interactions among these different players (Lafontaine *et al.*, 2018).

Bringing together stakeholders from different disciplines and sectors into partnership formations presented challenges and added to the complexity of scaling at the program level, as one respondent explains:

'When an international NGO came in with tried-and-true country engagement and outreach, for many partners this was very foreign and antithetical to what research was about – and it took quite a few tries to get to point where they were understanding one another' (SSI, IDRC Staff).

A lesson from CIFSRF Phase 2 was that choosing appropriate partners from different stakeholder groups and across disciplines (public sector, private sector and civil society) is key to implementing scaling-up activities. CIFSRF was designed on the premise that scaling requires partnerships that can harness the strengths, capacities and resources of

multiple actors within a system. All CIFSRF Phase 2 projects collaborated with various partners from public and private sectors and civil society to spread their innovations, which in some cases enabled projects to reach a larger population base without necessarily greater financial resources (Shilomboleni *et al.*, 2019).

As the examples above demonstrate, the evaluation team found evidence of staff members and programs thinking in new ways about the right mix of actors for scaling research results. However, we also found that this issue does not always get the attention it deserves. One IDRC staff member mentioned in interviews that responsible officers do not always engage directly with the question of who the right mix of stakeholders may be:

> *We have not always engaged with all stakeholders necessary and one of the reasons is partly because of the focus of who is seen as a stakeholder is left to the technical agencies doing the implementation of the initiative, so we each come at it from our own perspectives and some [responsible officers] are more targeted in their approach to their engagement with the project than others. (SSI, IDRC staff)*

Staff noted that thinking about scaling at IDRC has encouraged them to think more about coordinating with actors on both the supply and the demand sides. For example, a vaccine development project is engaging with upstream/supply side actors, while at the same time thinking of vaccine production, distribution and administration of the vaccine, which has required working with downstream/demand side actors. One IDRC respondent noted that constraints on the demand side had inhibited scaling and provided an example from Rwanda where only two veterinary scientists were able to administer a particular vaccine, and this prevented scaling. These ideas are elaborated further in field building for scale evaluation case study.

Staff we interviewed also noted that paying attention to coordination across **geographical locations** (from global to regional to national to sub-national) of the various actors in the scaling system is critical for program success. For example, a key design feature of KIX is linking regional knowledge and innovation hubs with national level structures. This is done by inviting a group of five representatives from the Ministry of Education and local education groups for each member country to join the hub. The hub's role is to share knowledge and innovation, including from KIX, among themselves and into national education processes. This strengthens the likelihood that research is demand driven and facilitates uptake of research findings towards scaling research results.

While scaling might require coordination for dealing with complexity, establishing large complicated and complex systems and structures is very challenging. A number of respondents mentioned that this was something that IDRC is learning to improve, and it has found a number of ways to support coordination and ensure that it contributes to scaling efforts. For example, building on lessons from CARIAA, the CLARE program is being designed to have a variety of different size projects, including smaller groupings, which could increase the flexibility of the program to respond to scaling opportunities and hence may be better suited for scaling.

Another lesson from the CARIAA program was that having dedicated coordinators within a consortia was a crucial design feature to allow for a new kind of agility and nimbleness to achieve successful outcomes (Cundill *et al.*, 2019). Additionally, within IDRC, a dedicated program officer was responsible for each consortium. For CIFSRF, each project had a dedicated coordinator to ensure that projects remained on track and were well documented. This role absorbed some of the managerial burden, enabling team leaders to focus on the big picture – making the connections needed to drive innovation and scaling (O'Neill & Manchur, 2018). IMCHA created East and West African HPROs as coordination structures. The EA-HPRO coordinated African and Canadian research teams on 19 research projects in six countries.

Representatives from KIX noted that they have been mindful that coordination is a large effort, and hence they intentionally designed coordination structures to not put too much burden on IDRC and its various partners responsible for coordinating all its 'pieces'. A key design decision was to bring in 'regional learning partners' (RLPs) to coordinate policy engagement and research uptake with country-level stakeholders. Four RLPs were identified through a separate targeted grants process. It is too early to tell whether this has been effective.

Important lessons identified from the CIFSRF program were that productive partnerships need to be nurtured, which requires time and space for partnership building. Partners need to be supported to have a common vision; strong leadership from each partner; and flexibility and equity within partnerships (O'Neill & Manchur, 2018). We were told that IDRC's responsible officers spend time building relationships, looking out for conflict and nurturing partnerships, and we have seen from the previous section that grantees appreciate the additional support provided by IDRC responsible officers.

A number of logistical and administrative challenges related to coordination. The CARIAA summative evaluation highlighted some serious challenges with collaboration including the withdrawal of partners leaving some issues unaddressed or excluding some countries; underperformance of consortium partners; tensions around finances and contracting arrangements; and retention of coordination personnel (Lafontaine *et al.*, 2018, p. 47). Just the issue of contracting can inhibit coordination and the solution is not

necessarily clear. For example, in CARIAA instead of allocating funding to one lead institution, each core partner in a consortium was given an individual grant. This was a risk-management strategy to ensure funds were spent effectively. However, it led to legally binding partnership agreements that can undermine collaborative outcomes. Cundill *et al.* (2019) reflect that the individual grants provided an unanticipated disincentive for core partners to collaborate as they reported to the funder individually rather than collectively. On the other hand, requiring grantees to manage coordination of multiple partners was a challenge raised by a number of grantees, one of whom called coordination 'a necessary evil'. Another respondent said:

> *One thing I would mention which is very administrative – a grant with a lot more partners could sometimes be difficult – for partners to take on the full coordination it becomes difficult. (SSI, grantee)*

As can be seen from these results, coordination of diverse stakeholders, across regions and sectors, is challenging and needs careful thought to maximise the value and minimise complications.

3.2.2. Engagement with the private sector

KEY TAKEAWAY: Engagement with the private sector is an example of IDRC connecting with a more diverse set of actors to support scaling. The private sector itself is diverse ranging from large multinationals to micro-enterprises in low- and middle-income countries. IDRC's programs and projects have worked together with private sector actors in various ways that they say have added value and supported scaling efforts. There are a number of challenges in working with the private sector, many of which stem from the different paradigms of private sector and development actors. The newly finalised IDRC Private Sector Engagement Strategy should help provide necessary guidance for engaging the private sector.

Coordination with the private sector was identified by IDRC in the inception phase of this evaluation as a particular area of interest for scaling research results. The 2015–2020 Strategic Plan committed IDRC to 'connect solutions with actors who can help advance those solutions to achieve large-scale impact', specifically mentioning the private sector. The data in this section comes from the private sector engagement and scaling evaluation case study.

The private sector is defined broadly as encompassing large multinationals to small or even micro-enterprises in developing countries. It also includes businesses or industry associations.

There has been a fair amount of engagement with the private sector, and IDRC has recently launched a Private Sector Engagement Strategy. IDRC staff from three programs

(Foundations for Innovation, Agriculture and Food Security, and Employment and Growth) said that private sector engagement was included in their program design in response to the strategic objective to scale. A 2019 presentation to IDRC's Board of Governors reported that approximately 25% of IDRC's research projects included engagement with the private sector, noting there is a substantial amount to learn from these experiences (IDRC, 2019). The 2018 mapping study of 50 IDRC projects with intention to scale found that two-thirds of the projects in the sample intended to engage the private sector (Sanchez-Swaren, 2018). This could suggest that private sector engagement is an approach being used more intentionally to scale research results.

In Table 2 we present several examples of engagement with private sector actors using a typology presented in IDRC's Private Sector Engagement Strategy. Further examples are provided in the private sector engagement and scaling evaluation case study.

Table 2: IDRC's typology for private sector engagement with examples*

Typology	Program
T1: Facilitating sector-level research with and on private sector actors	Pathways to Resilience in Semi-Arid Economies (PRISE) This research project was designed to support climate-resilient economic development in semi-arid lands. Private sector engagement included private sector actor participation in research as well as private sector participation in research dissemination and discussion. Intermediary organisations, or what respondents referred to as 'relay partners' were also included in PRISE for this purpose; that is, these organisations provided research team members with access to private sector networks for research participation and dissemination.
T2: Driving the development of new solutions with private sector actors in a sector	Livestock Vaccine Innovation Fund A number of strategies to engage the private sector included a) the allocation of IDRC funding to private research institutions; b) IDRC facilitating partnerships between research teams and private sector actors who have the necessary skills and expertise to support the research process; and c) IDRC requesting that grantees seek and on-board private sector partners.
T3: Informing policy agendas to enhance	Governance, Ethics and Conflict of Interest in Public Health (GECI-PH) initiative

private sector actors' positive contributions	The GECI-PH network provides an important mechanism through which prioritisation of private sector interests in health policy formulation and consumer communication can be addressed. It also serves as a monitor of private sector attempts to influence the health research agenda, and dissemination of health research findings, in ways that support potentially harmful policies and practices.
T4: Convening cross-sector partners to drive collective action	Women in Trade Knowledge Platform This project's outputs included a multi-stakeholder platform to foster dialogue and policy discussions towards inclusive trade policies and practices that might support women-led businesses and their access to international markets.

*The typology is taken from IDRC's Private Sector Engagement Strategy 2030 (IDRC, no date). Examples were identified in the private sector engagement and scaling evaluation case study.

Benefits of engaging the private sector

Overall, staff and grantee interviewees were positive about private sector engagement, feeling that it added value to their programs and supported scaling efforts. Interviews highlighted four specific benefits:

- **Improved access to financial and technical resources**. For example, in the LVIF program, partnerships with private sector actors provided the necessary skills and expertise to support the vaccine research process from proof of concept to vaccine trials and production. In the Mobilising the Private Sector for Adaptation Finance project, IDRC collaborated with private sector actor Climate Technology Initiative Private Financing Advisory Network (CTI PFAN), which provided technical support on developing a climate change-appropriate business model.
- **Improved access to private sector networks for research, dissemination and advocacy.** A good illustrative example is PRISE: in seeking to generate new knowledge about how economic development in semi-arid regions can be made more equitable and resilient to climate change, PRISE used private sector networks to gather data for research and as an audience it sought to influence with its research results. Interviewees from three projects: PRISE, Access to Finance for SMEs in Least Developed Countries, and Policy Analysis on Growth and Employment (PAGE II) also noted that private sector networks can provide leverage for project-related advocacy.

- **Improved access to private sector channels for distribution of innovations to end-users**. A good illustrative example is the Scaling up the Production and Distribution of Double-fortified Salt (DFS) in India project, where working with Fair Price Shops supported the distribution of DFS to more than 50 million people in three Indian states.

The private sector engagement and scaling case study presents evidence that **engaging the private sector has assisted with sustainability**, as one grantee commented:

> *There were quite a few examples where policies like Development County Integrated Development Plans in Kenya and some National Development Planning in Burkina Faso were changed based on analysis and recommendations and through involving the private sector voices. So strangely enough, engaging the private sector brought a lot of sustainable changes in the policy arena. (SSI, grantee)*

In another example, the PAGE II program had an effective private sector engagement strategy for a vocational education intervention in Kenya that led to more sustainable funding from the government:

> *We worked on a vocational education project where we restructured the vocational training to meet the private sector's need. So, we engaged the private sector around the reform. We linked the academia with the industry. We were the broker between the demand of the private sector and the supply of what academia can produce and this led to government giving more funding for public and private initiatives on vocational education. (SSI, IDRC staff)*

Challenges of engaging the private sector

For many programs and staff, engagement with the private sector was relatively new and presented challenges. Many of the challenges stem from the perception that the sectors have different motives – it is generally understood that research for development is for the public good, while the private sector is motivated by profit. They are also not used to working together and have different vocabulary, tools and incentives. The staff we interviewed also noted that private sector actors are not always interested in working with research projects, particularly when the return on investment is perceived as low, risks are perceived as high, and project timeframes are lengthy.

Many of the challenges stem from the perception that the sectors have different motives – it is generally understood that research for development is for the public good, while the private sector is motivated by profit.

- **Varied capacity of IDRC staff to engage effectively with the private sector:** A review of private sector engagement for the Employment and Growth program (Tewes-Gradl & Elliot-Gaved, no date) reports varied capacity and understanding among IDRC staff on how to engage effectively with companies. The results of the evaluation survey with IDRC staff indicated that fewer than 5% (n=43) of staff respondents agreed to a great extent that 'IDRC had a good understanding of how to support engagement with the private sector' (19% agreed somewhat, 16% were not sure, and 61% agreed very little or not at all).
- A business case is an example of a tool widely used in the private sector to make the case for investment, which is not always familiar to researchers. The need to provide strong business models to engage the private was often mentioned by both IDRC staff and grantees. Part of building a solid business case is also about being able to demonstrate investment returns of an innovation for end users. This is not always easy to demonstrate in financial terms. For example, some of the CIFSRF 2 projects were unclear on the profitability and net investment returns for small-scale farmers from support provided to increase food production through private sector-led ICT-based extension services.
- **Intellectual property rights**: Another key challenge emerging from the evaluation data is questions of intellectual property (IP) rights that private sector players would usually retain to support profits. IDRC has been working on an IP strategy focusing on product patents, issues related to public access, and publications. Currently, IP concerns are largely managed on a project-by-project basis. LVIF, for example, engaged the Canadian Intellectual Property Office to provide its grantees with training in IP-related issues. LVIF has also worked with a group of consultants, including an IP expert, to obtain guidance.

There is compelling evidence in the private sector engagement and scaling case study that working with business-facing or intermediary organisations can help to bridge the gap between IDRC and its research partners and the private sector. Multiple IDRC staff interviewees referred to intermediary or business facing organisations – including non-governmental and professional organisations, networks, funds and forums – as important private sector actors; examples include the Private Financing Advisory Network, the Multilateral Investment Fund, Canada's Trade Facilitation Office,

WEConnect and the World Economic Forum. These organisations facilitate linkages, dialogue and cooperation among private sector, public and civil society stakeholders. They may also provide financial and technical assistance to the private sector.

> *Now, instead of giving our grants to research institutions, we give them to a business-facing or intermediary organisation as those are better placed to understand private sector research needs. So, the research is not only addressing the right questions but also the idea that new business models can go to scale. (SSI, IDRC staff)*

One example of this is from PRISE, which sought to generate new knowledge about how economic development in semi-arid regions can be made more equitable and resilient to climate change. Collaboration with intermediaries – including NGOs like the Kenya Markets Trust – gave the research team access to private sector networks to leverage interest and participation in the project both as research informants and as audiences for research results.

Coordination with the private sector is clearly understood as being important for scaling research results in many ways, and the newly finalised Private Sector Engagement Strategy 2030 will be important to help guide how staff understand engagement with the private sector.

3.2.3. Conclusion

Building collaboration structures to support scaling needs careful thought. Bringing together traditional and non-traditional partners to address challenges holistically and systemically can help to find more impactful solutions, however, supporting this collaboration is not easy.

BOX 5: Summary of enablers and challenges for coordinating scaling

Enablers / strengths

- IDRC has multiple examples of complicated programs with a variety of coordination arrangements that it has been learning from to support scaling of research results.
- Using an organising principle such as geographic location or hotspot has been effective for encouraging collaboration, learning and synthesis to support scaling.
- Coordination structures need to be nimble, responsive and flexible; they need to be able to adapt the approach when something is not working. While the evaluation did find evidence of this flexibility, this was not always the case.

- Having sufficient coordination capacity – either within IDRC or at grantee level, or both depending on the complexity of the project, is important. IDRC has experimented with a number of different options for this.
- Programs are supporting scaling by working with an evolving set of actors in a system as the research process unfolds. They are also realising that bringing in the right partners is critical; and the earlier the better.
- The creation of collaborative spaces, holding regular learning processes and the creation of a fund for emergent collaborations to support a program's adaptive management approach is another key design feature that enables scaling of research results. They are useful design considerations for any large, transdisciplinary research endeavour because they support the coordination of the complex web of actors involved in scaling processes, while facilitating collaboration and learning.
- Productive partnerships need to be nurtured.

Challenges

- Large coordination structures can become inflexible and inhibit the adaptation and quick response needed to support scaling.
- Complicated coordination structures add complexity to scaling processes and may not be the best way to coordinate different actors in large programs.
- Coordination takes time and resources; if not properly resourced it can take attention away from other important activities such as research itself.
- Thinking strategically about the right set of actors does not consistently get the attention it deserves from responsible officers.

3.3. Learning about scale

Learning about scaling has been a key feature of the implementation of IDRC's strategic objective to 'invest in knowledge and innovation for large-scale positive change' for two reasons. First, the strategic objective included an ambition for IDRC 'to be recognised for sharing its learning in scaling up solutions'. Second, as discussed in Chapter 2, IDRC's experience of scaling was limited to a small number of programs at the start of the strategic period and for many programs, this was a new way of working.

As a departure point for thinking about how IDRC has learned about scaling, in the staff survey the evaluation team asked respondents about the most important sources for learning about scaling.

Figure 7 shows that the top two answers relate to learning within programs: learning from grantees (67%) and learning from colleagues in the same program (49%). While

learning from outside their program (e.g., the Scaling Science work (26%), other programs (21%) or external experts (23%)) was significantly lower. The least important for learning about scaling was monitoring information (16%). This starts to tell a story about learning from efforts to scale that we look at in more detail in this section: first looking at opportunities and challenges with learning *within* programs, then learning *across* programs.

This section draws on the evaluation case studies (particularly the case studies on organisational learning about scaling and programming for scale), the organisational review, the staff survey and the grantee survey. This includes interviews focused on learning with 18 staff randomly selected from across the Centre, interviews with 10 senior staff, where learning was one of several topics covered, and interviews with 8 program staff and 8 grantees from 5 programs, which covered how those programs approached learning. Out of approximately 90 eligible program staff, 43 responded to the staff survey, and 95 grantees responded to the grantee survey, both of which included questions about learning.

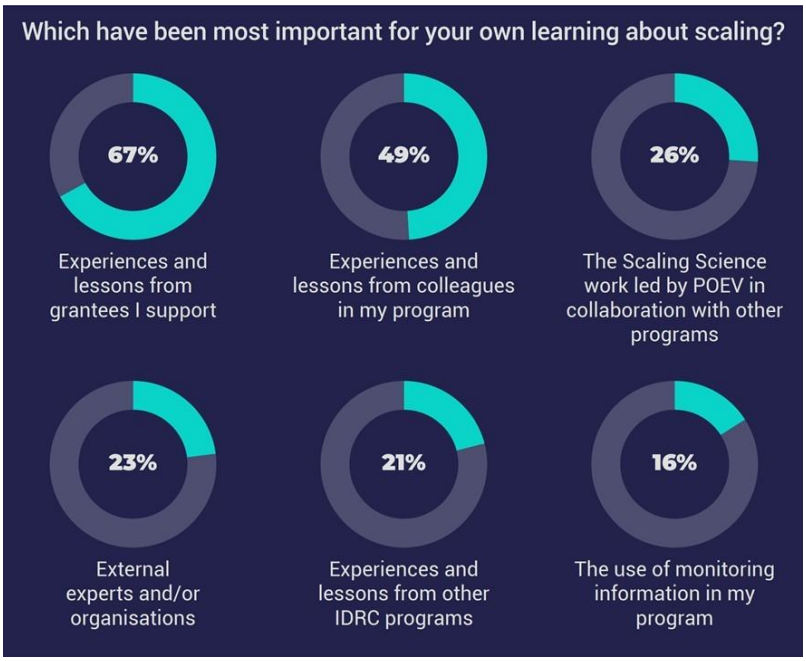

Figure 7: IDRC staff survey response to 'Which have been most important for your own learning about scaling? Select the top three.' (n=43)

3.3.1. Learning within programs

KEY TAKEAWAY: Learning about scaling tends to happen mostly within programs, suggesting that for IDRC staff, scaling is an applied concept that they are learning about through their practice and, most importantly, with grantees. Program meetings and workshops with grantees, as well as one-to-one discussions, are particularly fruitful spaces for learning about scale, alongside systematic learning reviews and synthesis papers – particularly for externally funded programs. The most pressing challenge to learning IDRC staff face is time pressure. Program leaders creating formal space for program officers to learn about scaling helped in some cases, but often learning (about anything) gets squeezed out for other corporate priorities.

IDRC staff play important and complementary roles with respect to learning about scaling. The evaluation findings highlight the roles of responsible officers, program leaders and program management officers. We found that grantees were, overall, very appreciative of IDRC's support and most of the grantee survey respondents reported that they engaged in learning about scaling with IDRC. When asked in the grantee survey whether they had discussions with IDRC about how to conduct ongoing assessment, learning and adaptation of the research process to consider what scale would achieve optimal impact, 78% (n=81) of respondents answered either somewhat or to a great extent. The following quote from a survey respondent illustrates this: 'The IDRC team was an excellent ally that allowed us to learn, grow, apply our knowledge and new knowledge and this has strengthened us as academics'. However there is demand for an even greater focus on learning in projects, 79% (n=86) of grantee survey respondents told us that they would have found it useful to have more discussions with IDRC about scaling research results in their projects.

The role of responsible officers in programs was highlighted by staff interviewees as particularly important for learning about scaling – in engaging with grantees and in looking across the projects in their portfolios. But we found that it is not just responsible officers who play an important role in learning. Program leaders have the authority to provide the formal space (including funded time) for staff to learn about scaling research results and to put learning into practice. Several interviewees emphasised the importance of the program leader role, with some attributing the learning within their program to the program leader's decisions. For example, one interviewee said, 'my program leader's zeal for scaling meant I had to put scaling into my strategy … and they follow up'.

Program management officers (PMOs) often organise learning within the program, synthesise data from various sources and produce reports for program leaders, program area directors and senior management. Therefore, PMOs are knowledgeable about the projects in their program's portfolio and are well connected. One interviewee said that

PMOs act as brokers between program officers, a role which is not formally recognised, and a PMO we interviewed said, 'if people want to know about the project, they come to me'.

The opportunities for learning about scale within programs discussed below emerged from interviews with staff and help to understand how and where learning happens. However, it is clear from these interviews that there are many challenges to effective learning about scaling in programs and these are also discussed below.

Opportunities for learning about scaling within programs

Staff interviewees described the spaces, places and processes where they had opportunities to discuss and reflect on their work on scaling. The evaluation team has categorised these opportunities into four groups, none of which are exclusively for learning about scaling but nonetheless play or have potential to play a role in learning about scaling. There is no hierarchy in the order to which these are presented.

Program meetings and workshops with grantees: Staff interviewees highlighted that program meetings for networking and exchange among grantees were considered fruitful for learning about scaling. For example, IMCHA hosted three 'all-of-IMCHA' meetings across the life of the initiative and the final meeting emphasised learning about scaling among the projects supported. In CIFSRF, a series of scaling-up workshops was hosted for grantees to come together to support learning, collaboration and synthesis and these were well appreciated by grantees for capacity strengthening and experience sharing (Universalia, 2016). For CARIAA, learning took place in collaborative spaces such as annual learning reviews and working groups. For example, interviewees noted that the research into use working group met face-to-face every six months and included discussion about scaling research results and how to achieve impact at multiple scales.

Systematic learning reviews and synthesis: Some externally funded programs, such as CARIAA and CIFSRF, undertook periodic reviews across their work with grantees to identify and exchange key lessons about scaling research results. For example, one of the aims of CIFSRF Phase 2 was to develop models for scaling that could be used by others, and for that purpose it engaged in research about scaling at the same time as scaling research results (Universalia, 2016). This resulted in important lessons about scaling that contributed to IDRC's understanding and to the field in general, particularly through peer-reviewed journal articles written by the program team.

A similar component has been designed into the KIX program, called Research on Scaling the Impact of Innovation in Education (ROSIE), which involves a third-party specialist research organisation to learn alongside grantees and synthesise lessons from the program. ROSIE aims to build grantees' understanding and capacity on how to scale positive impact and answer questions such as: How do you adapt proven approaches in

low(er) income countries? What does it take to scale them and how do you accelerate innovations to reach the largest number of children and sustain impact?

Interviewees agreed that synthesis across projects improves the overall quality of learning about scaling and the evaluation found evidence of this in the programs that included synthesis, such as GrOW, CARIAA and CIFSRF (see section 3.1.4). However, several interviewees in the organisational learning about scaling case study said there was little formal synthesis of knowledge across projects in the programs they worked in, with interviewees suggesting that the reason for this was a lack of resources and tools to support these kinds of activities.

Program evaluations: All programs were expected to commission evaluations and the topics or portfolios to be assessed were decided by programs themselves. Core programs tended to have a single summative evaluation, although not all core programs were evaluated in this strategic period, and externally funded programs tended to commission mid-term and end-term evaluations. CARIAA, however, took a staged approach to evaluation, which was perceived to have supported scaling by allowing them to develop learning questions throughout implementation. Evaluations were conducted by external consultants with the process offering an opportunity for knowledge and information from program staff and grantees to be formally documented and shared. Key findings were often shared through brown-bag lunches or webinars (which we discuss below).

When it comes to evaluating scaling specifically, all evaluations reviewed in the programming for scale case study reported on scaling results but varied in their treatment of scaling. A strong example is the evaluation CIFSRF commissioned to assess its contribution to food security, which assessed each project in terms of the prospects for scaling and the likelihood of sustaining the activities started under CIFSRF Phase 2. The contribution analysis methodology used was particularly effective for this kind of evaluation, which had a narrow set of questions applied across a portfolio of projects that are each targeting different points in a theory of change. Another example is the IMCHA summative evaluation, which included a sub-question about achievements with regards to scale up of research results. The findings presented for this question were identified from survey responses from grantees, and the evaluators reported that respondents conflated knowledge translation (i.e., presentation of findings to decision-makers) and scaling and ultimately, they found that it was too early to evaluate the success of scaling.

Project approval documents (PADs) and project completion reports (PCRs): Both types of project management documents play a role in learning about scaling—they operate as tools for planning, a mechanism for gathering data and as an opportunity for reflection as each has a section on the strategic objective to scale. The PAD asks for information on how the project intends to contribute to the strategic objective and the PCR solicits information on how the project contributed. PCRs also provide space for

responsible officers to document results and learning, which is then reviewed and approved by program leaders and directors. While other program management reports exist, these were not mentioned by interviewees as opportunities for learning. Interviewees mentioned a limitation of PADs and PCRs is that they promote an approach to learning that is focused on individual projects and they do not facilitate learning across a group of projects.

Challenges to learning about scaling within programs

While the interviews with staff highlighted several opportunities for learning in programs, they also highlighted several challenges, which, according to other external sources on organisational learning reviewed for the case study, are common in many public organisations.

Extreme time pressure: The challenge most mentioned by IDRC staff was extreme time pressure, with learning (about anything) often being dropped to focus on other corporate priorities. This was summarised by one interviewee as 'getting money out of the door' (and accounting for it). This finding is generally in line with Jones (2012), who suggests that there is a common problem in development agencies where the pressure to spend funds and the associated accountability mechanisms diminish learning and evidence-informed decision-making.

No mandate to learn about scaling: While scaling research results was a strategic objective, in practice staff received no mandate to learn about scaling at the program level. This meant it was easy for other priorities to take precedence, for example, gender mainstreaming was mentioned by several interviewees as being given more attention than scaling. Interviewees also pointed out that program meetings were more likely to spend time on planning than retrospective learning.

Limited use of monitoring data for learning: The staff survey highlighted a potential gap in the use of monitoring data for learning (see

Figure 7), with just 16% of staff respondents listing it as one of three top sources of learning. Interviewees suggested that during the earlier part of the strategic period, IDRC emphasised collecting and reporting primarily quantitative data in relation to scaling research results and this constrained some staff from thinking more holistically about scaling. As one interviewee stated, the targets focused attention on achieving at the expense of learning. This finding resonates with Guijt (2010) who says, 'official policies that profess the importance of learning are often contradicted by bureaucratic protocols and accounting systems which demand proof of results against pre-set targets. In the process, data are distorted (or obtained with much pain) and learning is aborted (or is too haphazard to make a difference)'.

Avoidance of challenges in formal learning spaces: Several interviewees said that there was a tendency to avoid talking about challenges and difficulties in formal spaces such as in PCRs and large group meetings. One interviewee remarked, 'in the PCR, being frank is not well received … we get told to limit or edit it … it's about risk management'. Problems and difficulties are inherent to IDRC's 'research for development' mandate which sets out to support a wide range of organisations in diverse contexts. 'Things will go wrong' as one interviewee said, and others remarked that acknowledging when they do is useful for learning about scaling. Therefore, while some interviewees pointed to the value of external evaluations in acknowledging difficulties and lessons learned and others suggested that they were able to discuss difficulties in their work informally within established teams, the tendency to avoid this in more formal, internal spaces may be limiting potentially valuable learning.

Skills and competencies for learning lacking: Finally, some interviewees mentioned that they lacked the skills and competencies for learning about scaling. This is not just about training on scaling, which was mentioned by some as a gap, but also about training for skills such as facilitation, synthesis and data visualisation, which can support more effective learning within programs.

3.1.2. Learning across programs

KEY TAKEAWAY: Learning across programs was less prevalent than learning within programs and more challenging. Successful cross-program learning tended to be based on good personal connections among individuals. Routine and mandatory spaces where people can interact across program boundaries (such as annual program meetings and brown-bag lunches) provided opportunities for learning about scaling, but these spaces were limited. The Scaling Science initiative was an important systematic effort to learn about scaling across the Centre that has contributed to greater understanding of scaling in some programs, although there is more work to do to share the learning from the Scaling Science initiative with staff.

According to the staff survey and interviews with staff, learning across programs was less prevalent and more challenging than learning within programs. As discussed in Chapter 2, IDRC has embarked on a learning journey on scaling and although overall its understanding of scaling has evolved greatly over this strategy period, it has not been a consistent experience across all programs and staff. Processes were in place for learning about scaling across programs, for example, but these tended to benefit those directly involved. Additionally, existing organisational spaces were used to facilitate learning across programs but there were a number of barriers to effective learning, as discussed below.

In addition to the program roles described above, interviewees identified the roles of POEV, directors of program areas and regional directors as important for cross-program learning. As discussed below, POEV's key contribution was in facilitating knowledge exchange through the Scaling Science initiative that it hosted. Directors of program areas played a role in facilitating connections between program level and corporate level, particularly, as mentioned by interviewees, around setting expectations and developing systems for data collection and monitoring. Regional directors support cross-program learning by connecting programs around important themes in the region and convening regional meetings.

Opportunities for learning about scaling across programs

Interviewees described various spaces offering opportunities for staff to interact across program boundaries for learning about scaling, including systematic learning processes such as the Scaling Science initiative, events such as annual learning meetings and brown-bag lunches, as well as other organisational factors such as program officer collaboration arrangements and regional offices.

Scaling Science: IDRC's Scaling Science initiative, led by POEV but with several programs substantively involved, featured research on scaling (through case studies of IDRC's programs) and communications (through a book, published articles and internal webinars and discussions). In particular, POEV initiated an IDRC working group known as the 'Scaling Science Critical Friends' to select and review case studies that provided a valuable space for cross-organisation learning for those involved. The scalingXchange aims to facilitate learning about scaling with IDRC grantees and improve IDRC and other funders' support for scaling.

According to several interviewees, the Scaling Science project has been one of the most important initiatives for developing IDRC's understanding of scaling. In particular, the concept of optimal scale was helpful for some staff in overcoming the prevailing idea that scaling is only concerned with achieving bigger numbers.

Interviewees suggested that the Scaling Science case studies helped to share experience from IDRC's work, and POEV's engagement in global communities of practice helped propagate lessons from other organisations. Several interviewees suggested that the Scaling Science initiative had helped clarify the terminology and concepts around scaling and open the definition of scaling beyond technological innovation to include policy influence, which has brought more programs into the scaling discussion.

However, the learning from the Scaling Science work has not fully permeated across the Centre. There are pockets of enthusiasm and appreciation, which our interviews picked up on, but these tended to be staff directly involved with the Scaling Science case studies and working group. Many of the interviews with randomly selected staff for the case

study on organisational learning did not mention the Scaling Science work, and some only mentioned a Scaling Science webinar they had attended but did not elaborate further. In addition, many of the staff we interviewed and program documents we reviewed presented scaling in ways that are inconsistent with the lessons from Scaling Science. For example, many sources talk about scaling-up innovations whereas Scaling Science emphasises scaling impact and optimal scale.

The staff survey findings also suggest that conceptual challenges remain when it comes to staff understanding of scaling. So, although Scaling Science is helping improve this, the ideas have not been around for long enough to have a marked effect on program practice and work remains to socialise the findings.

Annual learning meetings: IDRC's all-staff Annual Learning Forum was last held in 2011 and the Annual Program Meeting has continued in its place, the last of which was in 2019. These meetings took place at headquarters in Ottawa and brought together program staff covering various topics over 2–3 days. These were an opportunity 'to take a deep dive into a topic of corporate interest' and to learn about what other programs were doing. One interviewee recalled a session on gender transformative research with a presentation, panel and break out groups that did much to raise awareness. Another recalled a session on scaling impact that provided an opportunity to discuss what it meant and how it might be operationalised. One interviewee suggested the annual learning forum was IDRC's 'most structured learning process'.

Brown-bag lunches and webinars: Over half the interviewees for the case study on organisational learning about scaling mentioned brown-bag lunches and/or lunchtime learning seminars as spaces offering opportunities to learn, including around scaling. Traditionally, participants connect in groups from regional offices or staff travelling overseas – teleconference style rather than webinar style. However, Covid-19 has transformed these into formal webinars. They are organised by programs and have featured the results of IDRC project evaluations, presentations by grantees on an issue/project and the work of external experts on an issue of current interest. Some interviewees referred to a webinar where results from the Scaling Science initiative were shared with staff, which offered a formal opportunity to provide feedback and reflections.

Collaborating PO: Incorporating a collaborating PO (co-PO) into a project provides opportunities to work with colleagues from other programs in a formal way as a potential avenue for sharing learning on scaling. However, three interviewees suggested that working together in this arrangement tended to be a struggle as there were few incentives to be a co-PO and the investment in time did not always lead to quality learning or other benefits, and the co-PO is sometimes left out of the loop.

Regional offices: Some interviewees noted that regional offices were a site for cross-program learning. In one case, all responsible officers who managed projects across different programs in the region were brought together periodically. One interviewee stressed that regional offices also enabled IDRC staff to connect directly with grassroots organisations in the developing world, providing a means to connect senior staff to country realities.

Informal / ad-hoc discussions: Several interviewees suggested that conversations among staff from different programs located in the same physical location (for example, in Ottawa or in regional offices) pre Covid-19, often happened informally along hallways and in stairwells and during mealtimes and breaks. One interviewee said, 'different people gather at lunchtime or during breaks, this was a good way to meet people and raise questions', while another said, 'we were working on the same floor, a few offices apart … organic discussions are more likely to happen if located in the same space …'. Some interviewees said they often met other staff socially, which in turn led to more structured/intentional interactions.

Challenges with learning about scaling across programs

As with learning within programs, there are challenges with learning across programs, predominantly related to organisational factors.

Different approaches to scaling: One major factor discussed earlier in the evaluation that came up several times in our interviews with staff about learning is that programs had their own approaches to and understanding of scaling and these were not always consistent across programs.

Program silos: As one interviewee put it, 'programs often do their own thing, they are not looking horizontally, they are focused, single minded'. There is sometimes competition of ideas among programs which inhibits learning, for example one interviewee said, 'how do we accommodate different epistemologies, different types of data and types of processes? We've struggled, and this is far from being resolved. Is there a hierarchy? How does it all come together?'

Overall, the differences between programs seems to make it more difficult for staff to collaborate, as pointed out by one interviewee, 'you have to work hard and have good chemistry with counterparts and shared understanding about the problem, the context, and the grantee …'

Different starting positions: In some interviews, staff said that the way scaling was introduced at the start of the strategic period made it appear more relevant for programs that align more closely with natural sciences (such as agriculture and vaccines) than

social sciences (such as governance and social services). For some IDRC programs that have a strong emphasis on informing policy change, it was not immediately apparent how they would approach scaling, despite experiences from programs like IMCHA that were readily exploring scaling through policy. This is illustrated by one interviewee who remarked: 'Sometimes it's easier to scale in agriculture or financial technology, but not on social or political issues … if you're talking about policy adoption, can you see that as scaling impact?'

Too much knowledge is held privately: Finally, staff interviewees highlighted the considerable knowledge held in reports and by individuals across IDRC, which could aid their learning about scaling research results. However, on the whole, accessing this knowledge often proved challenging. As one interviewee said, 'learning depends on how much you know about what programs and teams are doing on similar issues … but it can be difficult to get a sense of what's going on … hard to get details about who to approach about what …'. This is consistent with Moynihan and Landuyt (2009) who suggest that organisational learning in the public sector can be aided by systems that collect, store and disseminate information. While IDRC has a document management system based on SharePoint, we were told by some staff that this is not used for finding learning from other programs. Other interviewees pointed out that internal communication was often the limiting factor – that the knowledge exists but is not always made apparent to staff who could use it.

3.3.2. Conclusion

Learning has been particularly important in implementing the strategic objective to scale because, as we noted previously, while the strategic objective committed IDRC to delivering solutions at scale, it did not constrain or prescribe programs on how to deliver it. Thus, it was important for programs to develop and strengthen their own approaches and to share that experience with others.

This section highlighted various formal and informal spaces and processes that offer opportunities for learning about scaling within and across programs, as well as challenges to learning in both cases. We found that learning has occurred more frequently within programs than across programs and tends to be more conducive in less formal processes (such as the Scaling Science Critical Friends group) than more formal processes (like the PCRs). Externally funded programs were more likely to provide substantive learning opportunities than core funded projects, as the former were able to allocate more resources for periodic reviews to generate syntheses and convene partners. In Box 6, we present a summary of these strengths and challenges.

BOX 6: Strengths and challenges relating to learning about scaling

Strengths/enablers for learning about scaling

- Program leaders have authority to provide space (including funded time) to support responsible officers to learn about scaling and to put learning into practice.
- Program staff learn from grantees- program meetings and workshops with grantees, as well as one-to-one discussions, are particularly fruitful for learning.
- Systematic learning reviews and syntheses undertaken by some externally funded programs help to identify and exchange learning.
- CARIAA's staged approach to program evaluation supported learning about scaling by allowing them to develop learning questions throughout implementation.
- The Scaling Science initiative facilitated a systematic effort to learn about scaling across programs, helping some staff to learn and develop their understanding of scaling through engagement in the study (in particular, the working group developed to select and review case studies) as well as through the published results.
- Some regional offices brought together programs in the region periodically, and supported staff to connect with grassroots organisations, which supported learning.
- Informal conversations between staff in the same office (pre-Covid-19) facilitated learning between staff from within and across programs.

Challenges /areas to improve on for learning about scaling

- Time pressures facing program staff, with learning often dropped to focus on what is perceived to be more pressing corporate priorities.
- While PADs and PCRs provide space for reflection on scaling within a project, there is limited formal support for staff to undertake systematic synthesis of knowledge across projects.
- Monitoring data does not appear to be used for learning within programs, with just 16% of staff survey respondents listing it in the top three sources of learning.
- Limited routine and mandatory spaces for staff to interact across program boundaries.
- Tendency to not want to discuss challenges in formal reporting mechanisms.
- Different understandings of and approaches to scaling across programs, coupled by a tendency for programs to focus on their own topics and approaches without looking horizontally at what other programs are doing.

4. POSITIONING IDRC'S APPROACH TO SCALING

This chapter positions IDRC's approach to scaling within the wider research for development field. It looks at what IDRC can learn from how other research funders approach scaling and how other research funders perceive IDRC's niche within the development research field.

This chapter draws on interviews and survey data from individuals working in the research for development field, including multilateral organisations, development banks, bilateral organisations, foundations and research institutes. Sixteen people completed the survey (referred to as respondents) out a shortlist of 38 people invited. Fifteen people were interviewed (referred to as interviewees) out of a shortlist of 20 people invited to interview. All but three of the interviewees were also invited to complete the survey so there may be some overlap between respondents and interviewees. The survey results were anonymous so the extent of the overlap is unknown.

4.1. Other organisations' approach to scaling

KEY TAKEAWAY: Many of the funders interviewed and surveyed have a strong alignment with IDRC's Scaling Science approach, which emphasises similar considerations regarding coordination and context sensitivity. A key difference, though, is that for most of the funders we asked, scaling predominantly means reaching more people and achieving lasting impact. Like IDRC, organisations are integrating scaling into their organisational systems and processes but they experience challenges when reporting on scaling.

4.1.1. What scaling means to other funders

For the majority of survey respondents and interviewees, scaling predominately means reaching more people and achieving lasting impact:

- Of the 15 interviewees, 10 indicated that they associated the term scaling with the ability to effect change on a bigger scale and among a larger number of beneficiaries. Almost half the interviewees indicated that scaling should facilitate longer lasting or sustainable change.
- Many interviewees cited the Sustainable Development Goals (SDGs) as the type of change being sought: *'Our goal is to achieve the SDGs. And we have a strong opinion that that will not happen unless people work with a much stronger scaling-up mind-set'* (SSI, external stakeholder).

- In the survey, we asked respondents what scaling typically means to their organisation; 15 out of the 16 respondents said reaching more people and 14 said having longer-lasting impact. Fewer respondents selected quality of impact (11 respondents), variety of impact (10 respondents) and equity of impact (8 respondents).

This perspective is shared by some IDRC staff we spoke to but is counter to the perspective developed in IDRC's Scaling Science work, which emphasises optimal scale – that is aiming for impact on a bigger scale is not necessarily better.

4.1.2. Paying attention to context and complexity

We asked funders which factors they expected grantees to consider when applying for a research grant, on the premise that scaling requires consideration of factors beyond the actions of immediate users. The responses suggest that funders value attention to context and complexity of scaling. The survey results show that most respondents expect grantees to consider alignment of the research to the policy context (11/16 to a great extent), regulatory frameworks that could enable or hinder uptake of an innovation (10/16 to a great extent), cultural and gender norms that might affect scaling (10/16 to a great extent) and existing approaches or innovations (10/16 to a great extent).

Interviews were consistent with this, with interviewees mentioning policy context and enabling environment for scaling as important criteria when reviewing grant applications with a view to scaling research results. In addition, interviewees mentioned other criteria such as potential partners to scale impact, opposition to the initiative, government buy-in, cost-effectiveness of the solution and local ownership. One funder incorporates these and other factors in a dedicated grantee guideline document on how to address issues of sustainability and scalability when applying for grants.

Two thirds of the interviewees (10/15) also highlighted the importance of coordination with other actors to support scaling. For example, four interviewees said they wanted the grantees to think about who the champions for scaling are, who they will work or partner with and who might oppose the scaling initiative. An additional two interviewees said they wanted the grantees to outline the stakeholders who can create access to funding and markets, and how they are going to work with the private sector. One funder requests grantees to set up a technical advisory group with key stakeholders, including government officials, and have several rounds of stakeholder workshops to ensure contextual factors are considered.

This indicates that many of the funders interviewed and surveyed have a strong alignment with IDRC's Scaling Science approach, which emphasises similar considerations regarding coordination and context sensitivity.

4.1.3. Integrating scaling into the organisation

For many of the interviewees, scaling research results is seen as intrinsic part of their work in the development sector and hence it is well-aligned to their respective mandates and missions. Most of the survey respondents reported that they had incorporated scaling into their organisational strategies (10/16). This filters down to some extent into other aspects of organisations' work. For example, more than half the survey respondents include scaling in project monitoring and evaluation (M&E) activities (9/16), while half reported including scaling in their calls for proposals or funding applications. Interviews corroborated this, with interviewees reporting similar adoption of scaling in organisational processes. Three interviewees, for example, reported the roll-out of capacity development initiatives to equip staff and grantees with the necessary knowledge and skills to integrate scaling into their work, as evidenced by the quote below from a bilateral research funder.

> *There has been a lot of work on scaling impact. It has been integrated more systematically in our management. For example, we have an orientation for program managers and advisors on how to include scaling up and how to build scaling strategies into their projects. (SSI, external stakeholder).*

The majority (12/16) of the survey respondents said they use indicators, metrics and progress markers to monitor the scaling of research results while four said they do not. This was also echoed by the qualitative interviews which revealed that typical indicators would be dissemination of research results and uptake of recommendations by policymakers or others. Most interviewees, however, said that monitoring of scaling of research results was difficult and is more likely to be based on anecdotal evidence or other informal methods.

Other interviewees pointed out that monitoring is more straight forward during the research and development stages which they fund as they receive regular reporting from the grantees. The longer-term impact, however, can be difficult to monitor as this is beyond the time frame of the funded work. One foundation has found a way to do longer-term monitoring of impact. It unpacks the impact model with its grantee before signing the funding contract to extract the innovation trajectory and the assumptions in order to go to scale. After the funding cycle of an innovation has ended, staff telephone the previous innovators every couple of years to learn about how the impact model is unfolding.

Other monitoring methods offered by interviewees include using M&E plans with scaling specific indicators and a readiness for scaling tool to monitor scaling of research results. Finally, two interviewees said they plan to use outcome harvesting to track outcomes of the scaling up of research results.

The evaluation team found it interesting that only seven of the survey respondents indicated that scaling was included in their organisational reporting and accountability, despite 10 respondents reporting that scaling was included in their organisational strategy. This indicates to us that some organizations may be experiencing challenges with measuring and reporting on scaling initiatives, which is consistent with the findings of this evaluation regarding IDRC's experience.

As for evaluation approaches, most (12/15) of the interviewees emphasised that they do not have standard approaches but that each evaluation is designed to meet particular needs, drawing on a range of approaches. Others admitted that they have not done many evaluations on scaling of research results. About half of the interviewees (8/15) said they used randomised control trials (RCTs) to measure impact and one combines RCTs with economic modelling. In all but one case where RCTs were mentioned, RCTs were used as one approach among several for measuring impact. Most of survey respondents (11/16) indicated that they use developmental evaluation as an evaluation approach while quantitative impact studies were less frequently used (6/16). Developmental evaluation is most closely aligned to IDRC's dynamic evaluation guiding principle for scaling impact.

4.1.4. How other organisations describe their niche

Interviewees were asked to describe their own organisation's niche in the research funding community regarding scaling impact; 11 out of 15 responded to this question. The common responses related to the processes and methodologies used ('taking proof of concept to a firm idea of what a path to scale looks like'; 'applying a long term and flexible approach as scaling impact takes a long time'), their sectoral focus (such as nutrition or reproductive health) and their ability to broker relationships among other actors including the private sector. They also spoke about their focus on systems change and developing capacity and tools for scaling.

4.2. Perception of IDRC's contribution

KEY TAKEAWAY: IDRC is recognised by other funders of research for development in the scaling community for its principled approach to scaling and that it is a key contributor on scaling that is equitable and responsible. IDRC could learn from others' experiences, but in general, IDRC already has thoughtful and innovative systems and practices to support scaling relative to others.

Most interviewees said they knew about IDRC through its research, publications, blogs and particularly through the Scaling Science initiative. Others said that they knew about IDRC's contribution through the Scaling Up Community of Practice. When asked about IDRC's niche in relation to scaling, interviewees said that IDRC's niche is the combination of being a leading supporter of researchers in the global south, with an explicit focus on scaling impact. Many interviewees acknowledged that IDRC has been at the forefront of embracing scaling impact; for example, one interviewee said:

> *They have a tremendous niche in scaling impact. They are on the ground, have partnerships, engagement and communication, and support research in the global south. They are key in research and in terms of scaling. They provide thought leadership particularly through Robert McLean. (SSI, external stakeholder)*

Survey respondents considered IDRC's contribution to the field of scaling (Figure 8) to be providing opportunities to implement scaling into research (63%), followed by knowledge sharing about scaling (56%) and a focus on scaling impact rather than scaling output (50%). Slightly fewer respondents selected a principle-based approach to scaling (44%).

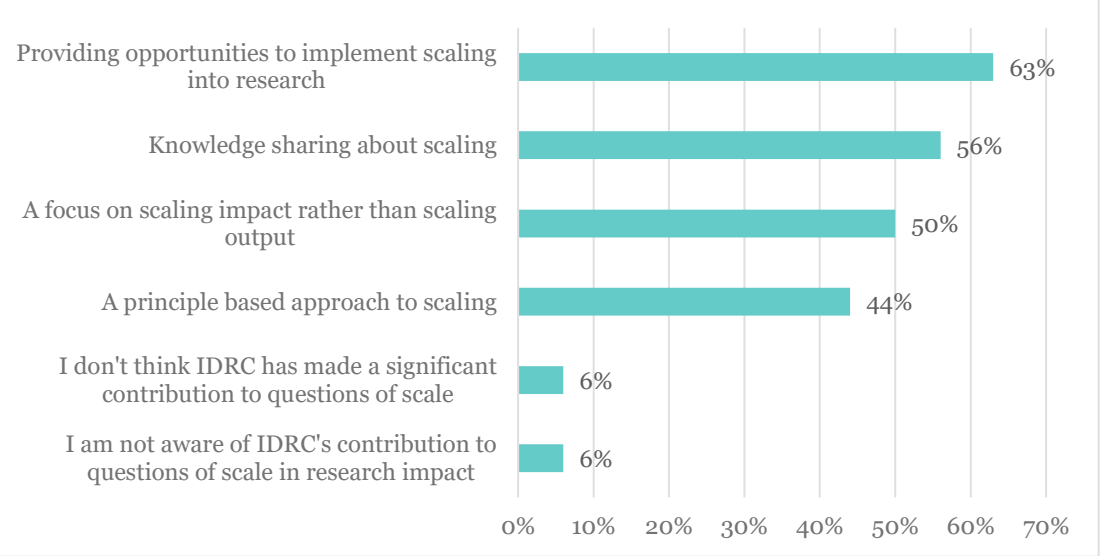

Figure 8: IDRC's main contribution to scaling according to respondents to the funders survey (n=16)

The funders we surveyed and interviewed were appreciative of their collaboration with IDRC and mentioned several strengths that refer to qualities or characteristics that have been identified as important for scaling – flexibility, adaptiveness, supporting innovation and a supportive partnership approach. IDRC was also commended for its high quality, solution-focused research.

IDRC's flexible and adaptive approach: Interviewees appreciated IDRC's flexibility and willingness to support exploratory research and research on issues that are often not considered mainstream. As one interviewee said: 'I certainly appreciate IDRC's flexibility around allowing ideas and practices to emerge and then responding in a fairly agile and adaptable way to interesting opportunities. IDRC has historically been more open to support research that is exploratory, often seeding new ideas, and trying things out' (SSI, external stakeholder).

Access to new tools and knowledge: IDRC's flexible and exploratory approach has allowed it to develop a range of innovative tools and research methodologies that are shared with its partners. Interviewees highlighted this access to new tools and knowledge as a key advantage of engaging with IDRC.

IDRC's approach to partnerships and collaboration: Interviewees noted IDRC's inclusive, supportive and respectful approach to collaboration as a key strength. Two interviewees highlighted this specifically in relation to the Centre's long-term commitment to supporting organisations and research in the global south. This was seen as an important contribution to building research capacity as well as a research 'ecosystem' among southern-based organisations (SSI, external stakeholder). Two interviewees also noted that IDRC's collaborative approach had enabled it to establish a substantial network of expertise, spanning many sectors. Partnering with IDRC enables access to this network, which in turn facilitates network building by and among its partners.

IDRC's reputation and standing in the research sector: Another frequently reported strength in collaborating with IDRC is its level of credibility and reputation as a trusted and trustworthy partner, capable of delivering rigorous research and quality evidence through its grantees. One interviewee also noted that IDRC is well-capacitated and well-resourced, thus allowing for a level of project scope and depth that is often beyond the means and capacity of its partner organisations.

IDRC's support for applied research: Interviewees noted IDRC's ability to support pure as well as applied research as a distinct advantage. For example, one interviewee

commented on the Centre's unique position of being able to '... fund research for research's sake ...' (SSI, external stakeholder), while another noted IDRC's work at the '... intersection between applied and academic research' (SSI, external stakeholder). The Centre's support of applied research was viewed as a key means of addressing 'real problems' and engaging in a constructive, solution-focused manner with local contexts and issues on the ground (SSI, external stakeholder).

4.3. Conclusion

IDRC's emphasis on scaling impact (rather than scaling an innovation) at optimal scale (rather than bigger or more) and its principled approach to scaling sets it apart from other funders that tend to associate scaling with reaching more people. Generally, the evaluation found that the Centre's thinking and practices around scaling are advanced relative to other funders of research for development and it is considered a key contributor to the field of scaling. We believe IDRC can offer the scaling community an opportunity to be more considered and responsible in their approaches to scaling. This is an important contribution in a field that is rapidly expanding and in which many influential funders, donors and multi-lateral organisations are involved.

As for how IDRC could learn from other funders, we identified three potentially useful practices that IDRC could consider:

- The establishment of technical advisory groups at the beginning of projects comprising all relevant stakeholders to better understand the context for scaling.
- Developing scalability tools to assess feasibility of scaling and make them available to grantees.
- Developing clear indicators to monitor scaling impact from the beginning together with grantees and to set aside resources to follow up with grantees after the funding cycle has ended.

5. WHAT OUTCOMES WERE ACHIEVED?

The evaluation has considered the question of what outcomes were achieved by integrating scaling into programming from two perspectives. First, we considered external outcomes: the results programs are reporting about how their work and the work of their grantees informs or influences others. The second perspective we considered was internal outcomes: the changes within IDRC to which the introduction of the scaling objective has contributed, and which are considered positive in IDRC's ambitions to scale the research results.

5.1. External outcomes

In this section we present a systematic analysis of external outcomes IDRC achieved by integrating scaling into its programming. External outcomes refer to changes in actors external to IDRC, its programs and its grantees to which IDRC can claim to have contributed through its approaches to scale research results.

Box 7 summarises the data sources reviewed, and the criteria used to identify outcomes, it also notes the number of outcomes identified for each source. See Appendix 2 for a detailed list of all sources and a description of the methodology for the outcome analysis. It is important to note that the sources do not cover all core IDRC programs evenly, as not all programs used Trackify and not all programs had evaluations during the strategic period. Because the evaluation focused on program and cross-program perspectives, only corporate and program level documentation was reviewed. Project level reports such as project completion reports or final technical reports were not within the scope for this evaluation. The findings, therefore, relate to how programs present results relating to scaling and do not take into account grantee perspectives, which may differ.

The analysis focused on two kinds of external outcomes: policy and innovation – as described in the scaling pathway (see Section 1.3.). These categories were selected as they relate directly to the corporate-level scaling indicators IDRC adopted and reflect the way many programs reported against the scaling objective in annual progress reports.

Policy outcomes refer to changes in public policy at sub-national, national and international levels that have been informed by evidence from IDRC research. We distinguish three types of policy outcomes, as defined by a framework used by POEV adapted from Carden (2009):

- Expanded policy capacities of external actors,[11] including for scaling.
- Informed policy dialogues and decision-making processes.
- Contributions to policy implementation or change.

Innovation outcomes refer to the adoption beyond primary users[12] of products, programs, practices or methods developed through IDRC research or adapted/scaled through IDRC research. Drawing on a framework describing stages of innovation, developed by an IDRC working group in 2018, we distinguish the following outcomes:[13]

- Initial adoption of the innovation by end users, beneficiaries or clients. Initial benefits/impacts are being delivered by the innovation.
- Innovation is being used beyond primary users. Impacts at scale are apparent.

Policy and innovation outcomes are not mutually exclusive. For example, innovations are commonly implemented beyond primary users by government agencies, and policy capacity or policy change is often a pre-condition for this. Likewise, the use of an innovation such as a methodology by a government agency may lead to improvements in policy capacity, dialogue or policy change.

BOX 7: Sources of data (with number of outcomes identified in parenthesis) and criteria for the outcome analysis

The following sources were reviewed to identify outcomes:

- 13 program evaluations, selected for their potential to learn about scaling and to cover all IDRC program areas (# of outcomes identified = 185)
- 239 data entries from the Trackify database for the two high-level indicators for scaling (# of outcomes identified = 211)
- 9 program area progress reports to IDRC's Board of Governors (# of outcomes identified = 39)
- Surveys of and interviews with staff and grantees (# of outcomes identified = 5)

11. Although many IDRC programs report grantee capacity (researcher capacity) as an outcome, we do not include this as a scaling outcome; we only consider policy capacity of actors beyond grantees, at the demand side of scaling.
12. We refer to 'beyond primary users' to indicate that outcomes have been achieved by scaling results reached with an initial set of users. In the scaling pathway (see Figure 1) when we refer to adoption/change beyond primary users, we specify 'at optimal scale', which is a term used in IDRC's *Scaling Science* study that challenges the 'bigger is better' scaling model. Optimal scale recognises that scaling produces a collection of impacts and to determine *optimal scale*, we have to consider the trade-offs among different types of impact; sometimes that may mean scaling down. Assessing whether outcomes were achieved at optimal scale is outside of the scope of this evaluation, thus in this section we refer to adoption/use/change beyond primary users.
13. The stages of innovation framework include three stages prior to these (proof of concept, working prototype, user-ready innovation) but we consider these as outputs and have not included such results in the outcome set.

The following criteria were used to identify plausible outcomes:

- Describes the actions of an actor external to the project (i.e., not just about what the project did).
- Describes the contribution of the project (i.e., there must be a plausible link to the project).
- Provides verifiable details (dates, places, events, names or organisations)
- Can be categorised using the outcome types defined above.

5.1.1. Overview of data

KEY TAKEAWAY: Across the sources analysed, 440 outcomes were identified. Using the criteria described above: 355 of these outcomes were policy outcomes and 85 were innovation outcomes. The majority of these occurred at national level. We consider this a considerable result for IDRC, given that this was not an exhaustive assessment of results. The outcomes identified tell a clear picture that IDRC has had many successes in scaling the impact of research.

Figure 9 shows the breakdown of outcomes across the three types of policy outcome and two innovation outcomes. The team identified five times as many policy outcomes than innovation outcomes. This is not unexpected given IDRC's long standing emphasis on policy engagement across its programs. Additionally, programs that seek to scale research results through supporting innovations tend to focus on a small number of innovations (and therefore we would expect a small number of outcomes given that the outcome counts the number of innovations being used beyond primary users) compared with a larger number of policies for programs scaling research results through policy. For example, CIFSRF Phase 2 funded 18 projects, each of which aimed to support a single innovation to scale, compared with CARIAA which funded four consortia, each of which engaged in multiple policy processes at multiple levels in multiple countries.

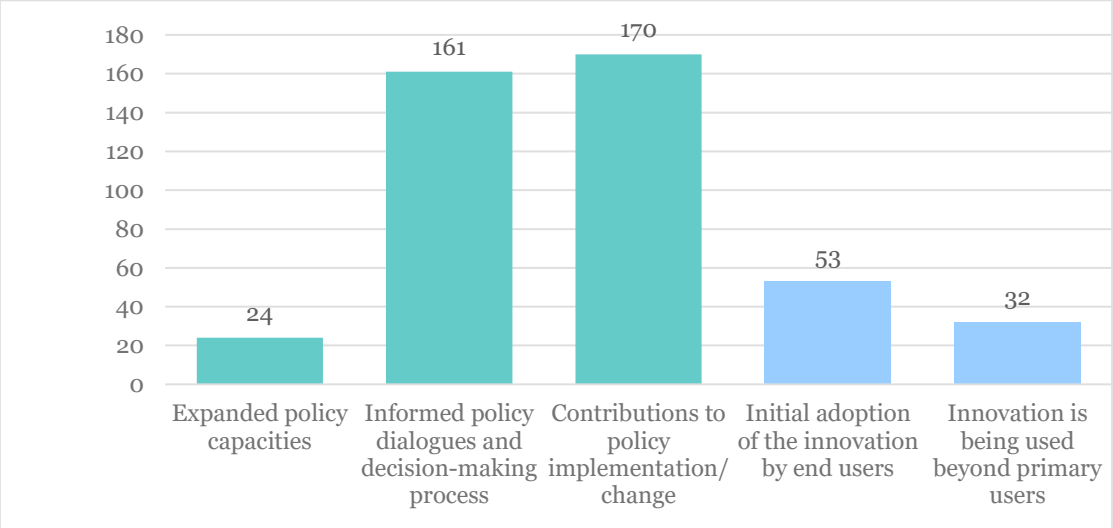

Figure 9: Total outcomes by type (n=440). Blue = policy outcomes, orange = innovation outcomes

We take a closer look at the policy and innovation outcomes later in the report but first we look at the geographic levels at which outcomes are occurring, the distribution of outcome types in IDRC programs and the regions in which the outcomes are occurring.

Geographic level

When we consider the geographic level at which IDRC is scaling (Figure 10), we can see that by far the most outcomes have been achieved at a national level – this is the case whether we take just the Trackify data or just the data from other documentary sources. When we consider that senior staff told us that one of the reasons for introducing the scaling objective was to shift the focus of programs from local levels to larger populations, from that perspective, the scaling objective appears to be achieving this intent. While we did not have data on the geographic scale of outcomes prior to 2015 to compare this with and so could not say with authority that a change had occurred, we could say that the data demonstrates a high proportion of outcomes at national level, which we were told was a desired aim for the scaling objective.

The relatively lower number of outcomes at a regional and global level was not unexpected by the evaluation team purely based on the lower number of opportunities for influence at this level. The outcomes identified at a regional and global level are noteworthy in that they demonstrate IDRC programs' increasing influence and building on work at national level – as highlighted in Box 8 and Box 9.

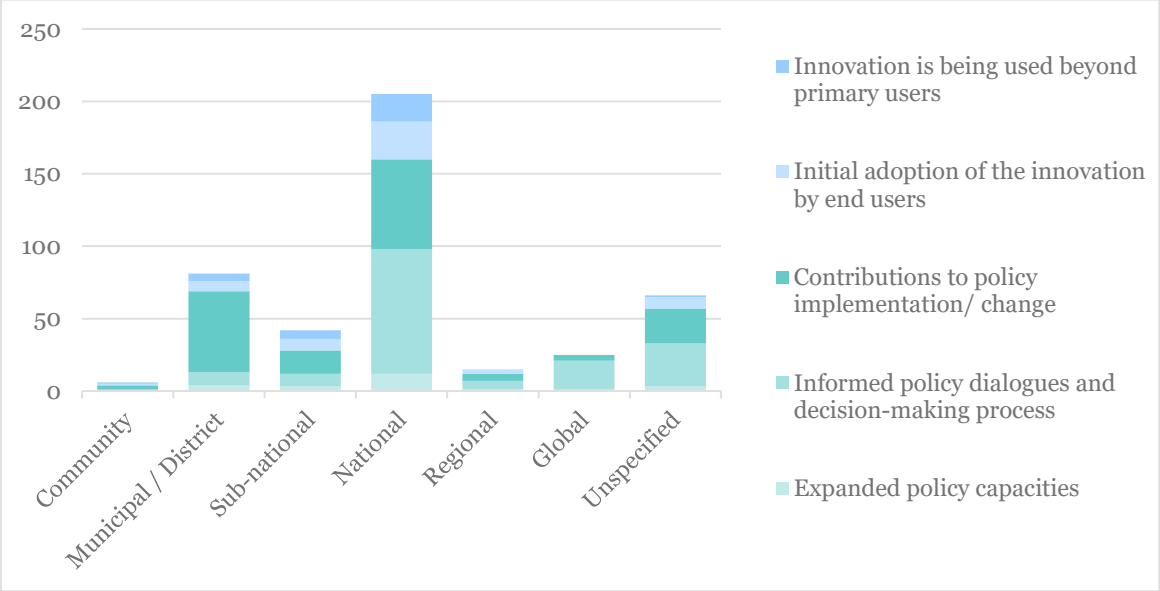

Figure 10: Outcomes by geographic level (n=440)

IDRC programs

Comparing the data across programs was not helpful given that not all programs recorded results into Trackify, and we have not reviewed documentation from every program, hence not all programs have been examined to the same extent to identify outcomes. However, to account for this, we can look at the proportion of different types of outcomes for each program that were identified across all the sources to get a indicative sense of how the programs varied in their balance between policy and innovation outcomes (Figure 11) – note that this is not definitive and only based on this specific evidence base.

Agriculture and Food Security was the only program with more innovation outcomes than policy, which the evaluation team does not find surprising given that most of the outcomes came from CIFSRF which had a big emphasis on innovation. All the other programs had more policy outcomes than innovation outcomes but varied in the balance between these two. This shows that the different approaches taken by programs result in different kinds of outcomes.

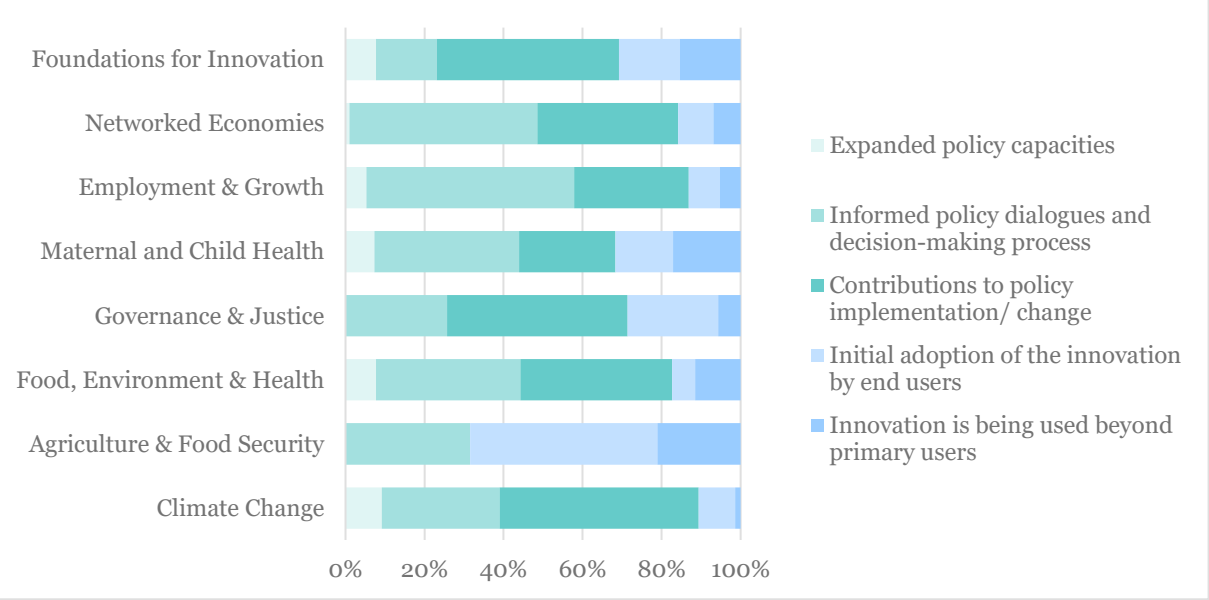

Figure 11: Proportional distribution of outcomes for each IDRC program, using outcomes from all data sources

Regions

We found that the majority of outcomes occurred in Latin America and the Caribbean (LAC) (see Figure 12), with most outcomes in that region identified as contributions to policy change (60/110).36 of those were from one Trackify entry, reporting the development and implementation of Strategic Action Plans for climate change adaptation in 36 municipalities.

The distribution of types of outcome across each region largely conforms to the distribution of the overall set (barring the policy change outcomes in LAC discussed above), with the exception of Asia, which has double the proportion of innovations used beyond primary users (14%) than the total set (7%) and includes innovations from all three program areas.

Looking only at regional outcomes, we found a roughly equal number of regional outcomes in Eastern and Southern Africa, Central and West Africa, the Middle East and North Africa and Latin America and the Caribbean. However, no regional outcomes were identified in Asia.

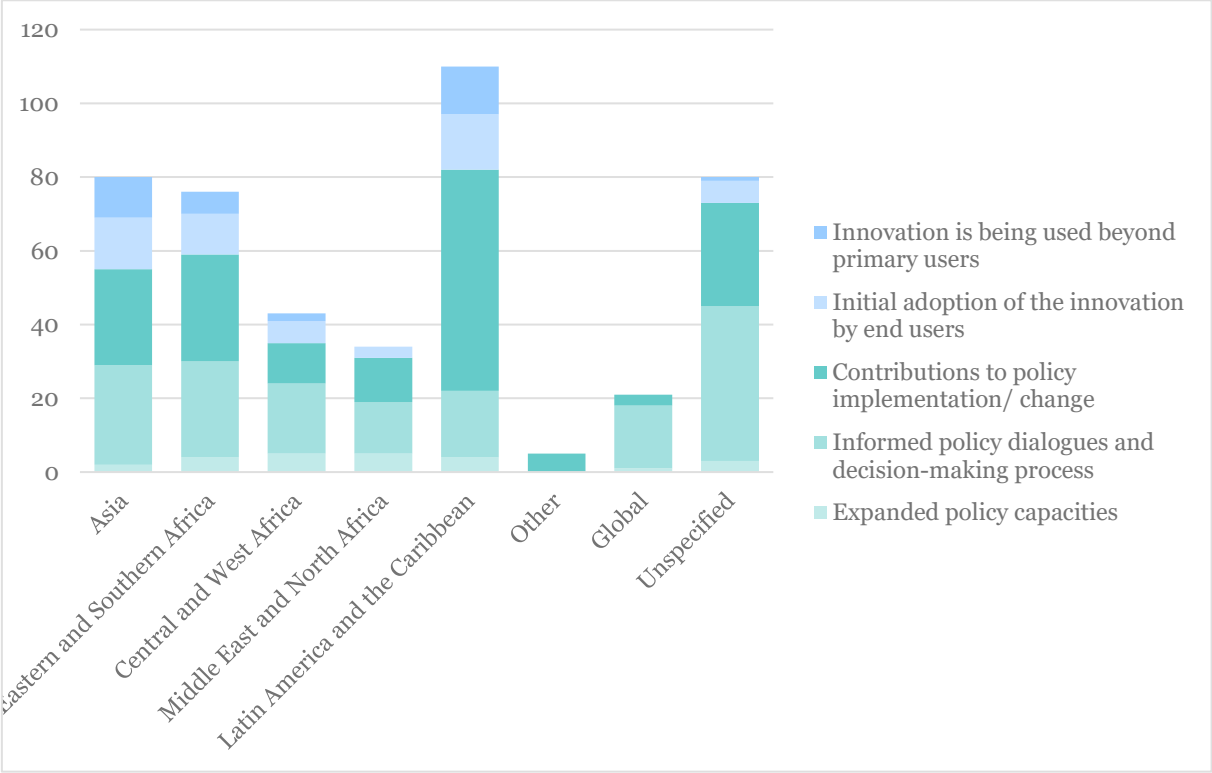

Figure 12: Outcome by region (n=440)

5.1.2. Innovation outcomes

KEY TAKEAWAY: The evaluation has identified 32 innovations being used beyond primary users from the sources reviewed. This is highly notable when we consider the original target for the scaling objective was to support at least 20 initiatives that deliver solutions at scale.[14] In addition, 53 innovations adopted by end users were also identified and are noteworthy given advanced potential of scaling at this stage.

The number of innovations being used beyond primary users identified in this analysis is significantly lower than IDRC reported in its Synthesis Learning Report (IDRC Policy and Evaluation Division, 2019) and the Performance and Learning Report to the Board of Governors (IDRC, 2019a), both of which reported 167 innovations being widely used

14. The original target to support 'at least 20 initiatives that deliver solutions at scale' was tied to IDRC's strategic objective to invest in knowledge and innovation for 'large-scale positive change'. Over the strategic period, the Centre's understanding of scaling evolved and the corporate indicators used to monitor scaling and that were used as the basis for reporting outcomes in this evaluation were broader to track scaling outcomes more generally and not just large-scale impact, thus it is unsurprising that IDRC surpassed this original target.

and adopted. The reason for this is that the 167 innovations were based on raw Trackify data which, as we discuss below, is not a reliable source unless it is analysed and cleaned.

Box 8 highlights five examples of innovation outcomes identified. These have been selected as highlights as they are mentioned by multiple sources and cover work across IDRC's program areas. The first example of potatoes in Colombia demonstrates the importance of supporting scaling through a range of strategies addressing both the demand side of scaling[15] (e.g., business models for distribution and household nutrition) and the supply side (e.g., family farming schools, seed entrepreneurs). The second example of double fortified salt in India demonstrates what can be achieved through coordination with public, private and philanthropic partners. The third example of electronic health records in Peru demonstrates the importance of political leverage for scaling in that one of the research team was appointed the minister of health for a year during the life of the project. It also shows the overlap between policy and innovation outcomes (as described in the scaling pathway) given that in this case a new policy was necessary for the government to implement the system nationally. The fourth example of graduation programs in Latin America again highlights the importance of collaboration with implementation partners – both public and private. The final example of a misinformation reporting system in Kenya and Myanmar demonstrates how research in one country can be built on and replicated in another country – in a completely different context – to achieve similar outcomes.

BOX 8: Five illustrative innovation outcome highlights

The **Expanding Adoption of Nutritious, Disease-resistant Potatoes in Colombia** project, one of the 18 CIFSRF Phase 2 projects, resulted in the development of three new potato varieties with higher yields, more protein, iron and zinc, as well as more blight resistance. The potatoes are estimated to have reached 6.5 million people in Colombia. The success of this project is attributed to its multipronged strategy which included the development of a sustainable business model centred on rural entrepreneurs as specialist seed producers. It also included family farming schools, community garden groups and leadership schools for women in 13 municipalities. In addition to scaling access to the potato products themselves, the project's business model is being scaled across the country by a national organisation (Milena Buitrago Rodriguez, 2018; Wiggins *et al.*, 2018).

Scaling up the Production and Distribution of Double-Fortified Salt in India, another CIFSRF Phase 2 project, took double-fortified salt technology from Canada, which had been proven to prevent anaemia, primarily a women's health issue,

15. Refer to the scaling pathway in Section 1.3 for descriptions of the supply and demand sides of scaling.

and scaled it in three Indian states (Uttar Pradesh, Madhya Pradesh and Jharkhand) reaching 50 million consumers. By working with governmental, private sector and philanthropic partners, the project was able to leverage 25 times the project budget to support scaling. This resulted in the building of a full-scale commercial production plant, contracts with 10 processing companies, standard operating procedures, training of 10,000 women village health workers, training of 8,500 Fair Price Shops and development of a distribution network through those shops. Through the efficient production and distribution process, the project was able to keep costs low; the incremental cost of adding iron to salt is less than CAD $0.25 per person per year (Diosady, Mannar & Menon, 2018; Wiggins *et al.*, 2018).

The **Strengthening Equity through Applied Research Capacity Strengthening in E-health** project piloted an electronic health record system in Peru (called WawaRed) for pregnant women in 15 health centres. The system increased access to health information for women, their healthcare providers and the ministry of health. This is significant because the Peruvian health system is fragmented and provides care through numerous disconnected public and private health establishments. Pregnant women receive care at multiple different centres throughout their pregnancy, leading to many points in the prenatal-delivery-postnatal process that are prone to information errors, which can greatly affect quality of care. These are amplified for women from isolated rural areas where the centres can be in different parts of the country. The pilot's success led to national adoption of the electronic health record system by the government. On 16 January 2017, Peru's Minister of Health, Patricia J Gargía Funegra (who was also one of the researchers on the team), signed a ministerial order to implement the system in 350 health centres countrywide. By 2018, the system had been introduced to 2,240 midwives from 646 health facilities in 18 of Peru's 25 regions and has already registered more than 50,000 women (Trackify, Pérez-Lu et al., 2018).

The **Enhancing Economic Opportunities in Latin America: From Poverty Reduction Projects to Sustainable Livelihoods** project supported the scaling of 'graduation' programs in five Latin American countries. Graduation programs are designed for people in extreme poverty and aim to enhance the assets that allow households to become self-sufficient and build knowledge to cope with shocks without falling back into extreme poverty. Through collaboration between government and NGOs (and the private sector in Paraguay), the pilots implemented by the project have benefited an estimated 139,000 people, 70% of whom are women. By conducting process evaluation, evaluation of effects and life stories qualitative research, the program has generated evidence to support the graduation approach and inform policy making. In Paraguay, the newly elected government (in 2017) decided to scale the program to reach 24,000 families and has committed around US$35 million to the program. As of 2018, Mexico's government had committed to incorporate graduation

into the national strategy, with a potential reach of 1.3 million of the poorest households enrolled in Mexico's social protection program (Trackify, Grantee survey, Fundación Capital, 2018).

The **Scaling Digital Solutions for Conflict Management in Kenya and Burma** project piloted a platform called Una Hakika (Swahili for 'Are you sure?') in Kenya's Tana River county. The platform integrates technology with traditional human networks to counter misinformation by verifying rumours submitted by subscribers and countering false rumours before they begin to spread and in so doing aims to reduce conflict. The reach of the platform expanded substantially over the period of the pilot and by 2017 had 15,000 SMS subscribers and 10,000 engaged on social media. Based on the pilot results, the platform was scaled to neighbouring Lamu County as well to the Democratic Republic of the Congo (under the name Kijiji Cha Amani – Swahili for 'Peace Village') and Myanmar (under the name Peaceful Truth). The introduction of the platform in Myanmar was at the time of the Rohingya refugee crisis and there was a lot of disinformation in social media around this. By 2018, the platform had generated significant demand and had 50,000 users subscribed (IDRC, 2017; IDRC, 2018 and project documents).

Fifty-three innovations adopted by end users were also identified and are noteworthy given the advanced potential for scaling at this stage. One example is the solar power pumping stations developed by the *Himalayan adaptation, water and resilience* project in Pakistan (part of CARIAA), which by the end of the pilot had secured agreement from the government to subsidise 30,000 units (Lafontaine *et al.*, 2018). Another example is the Scaling-up Production of Construction Materials from Recycled Mine Waste project (part of the International Research Chairs Initiative), which piloted ways to recycle mine waste in the manufacturing of construction materials. In response to the pilot, the Moroccan Ministry of Mines, Water and Environment was considering possible support to an eco-industrial complex project that would scale-up the research team's applications (IDRC, 2018c). Third, the Making Growth Work for Women in Low-Income Countries project (part of GrOW) created an app that allows women in Lahore to report on the points in their journey on public transport where they felt the most afraid or vulnerable, generating evidence for use in policy making in areas such as street lighting and police presence. There was interest from the World Bank to use the app in Tanzania (Survey, IDRC Staff).

5.1.3. Policy outcomes

KEY TAKEAWAY: The evaluation identified 170 contributions to policy change. However, it is not always possible to determine the scaling aspects of these outcomes. The number of outcomes relating to policy capacity is low (24) compared with the number of policy dialogue and policy change outcomes and this is considered by the evaluation team as a gap in current reporting of scaling outcomes.

An initial observation when we look at the policy outcomes is that there are more contributions to policy change (170) than to informed policy dialogues (161). This is due to the way the data was collected and does not indicate any broad pattern in how policy changes come about. Roughly half the policy outcomes identified come from Trackify where programs record results against the high-level indicator (number of contributions to policy change) and so there was a bias in the sub-set of data reviewed by the evaluation team towards recording policy change outcomes over other policy outcomes (which would have been recorded under different indicators in Trackify). There were 60 policy dialogue outcomes identified through Trackify and these were the result of reclassifications made by the evaluation team of entries that had been reported as policy change but were more accurately described as informed policy dialogue.

A typical example of this is a contribution to policy change outcome recorded in Trackify that described a project where a grantee was working with the National Commission for Women in India to contribute to an approach paper for a National Policy on Crèches. Based on this outcome description, it is not clear that the IDRC-supported research has contributed to policy change, but it is informing policy dialogue, so the evaluation team reclassified it as such. Seventy-six of the 173 policy change outcomes from Trackify that met our criteria were reclassified in this way.

Looking at the data from sources other than Trackify (program evaluations, program area progress reports, and interviews and surveys with staff and grantees) (Figure 13), the trend is different – there are about 40% more outcomes relating to informed policy dialogue than contributions to policy change. This is consistent with many IDRC program impact pathways which suggest that informed policy dialogue is a contributor to changes in policy and hence the former is more likely to be observed before the latter.

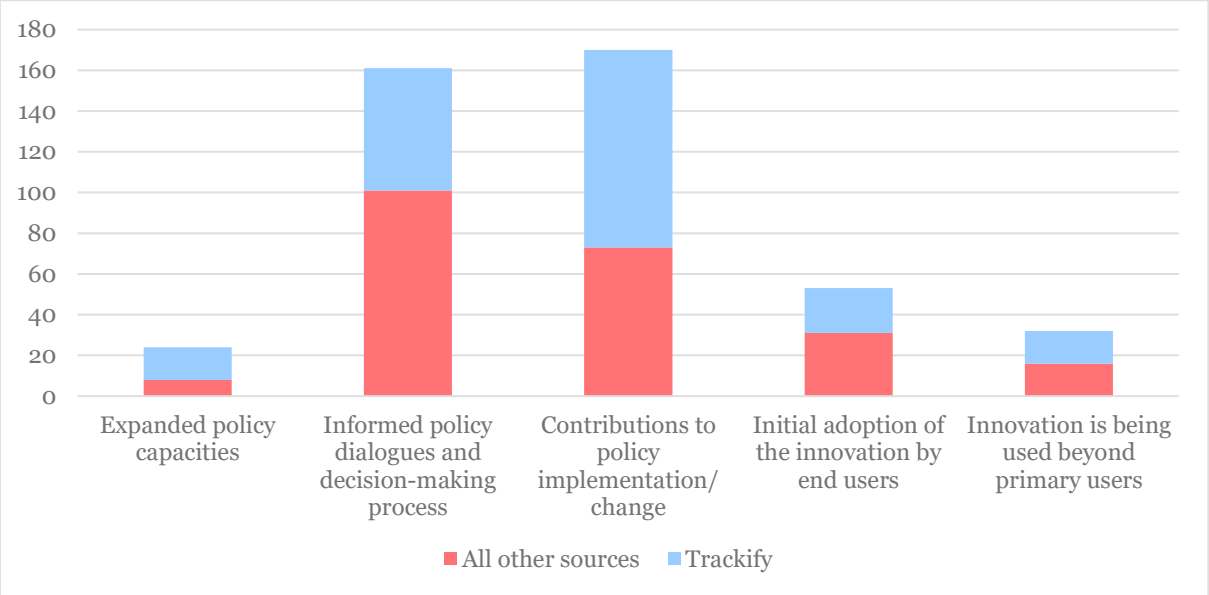

Figure 13: Policy outcomes by source of data

Box 9 highlights four illustrative examples of contributions to policy change at national, regional and global level. These highlights have been selected because they were mentioned by multiple sources (program evaluations, Trackify, staff and grantee surveys and/or interviews) and had sufficient details available to present plausible contribution from IDRC-funded projects.

The first highlight of the sugar tax in South Africa demonstrates how research in one region (initially in Latin America) can provide an entry point for researchers in another region to support policy change. It also demonstrates the possibilities for continued engagement when researchers are funded to evaluate the implementation of policies their research contributed to. This means they are able to look beyond the policy change to the effect of the policy on the lives of citizens – which is the development outcome in the scaling pathway (Figure 1).

The second highlight of the anti-corruption legislation in Argentina demonstrates an approach to policy change that engages with multiple actors in the system – not just the legislators but those directly affected by the legislation (private sector executives) and those responsible for implementing the legislation (prosecutors and defenders). This shows that the project looks not only at policy change but also at what they can do to ensure that the policy has the intended impact at scale – to reduce corruption. It also demonstrates the importance of the systems strengthening part of the scaling pathway (Figure 1).

The third example of regional targets for sodium reduction in Latin America is an example of building on prior research which had success at influencing national policy to achieve regional policy change. The final highlight is an example of influencing global level decision-making, namely two decisions adopted by the 197 parties of the United Nations Climate Change Conference (UNFCCC) made at COP23, which were of particular importance to African countries.

BOX 9: Four illustrative highlights of policy outcomes at national, regional and global levels

National level

Through the IDRC funded project, The Economic and Health impacts of Legislative and Fiscal Policies to Improve Nutrition in South Africans, researchers from the University of the Witwatersrand (Wits) worked directly with the of National Department of Health in South Africa to provide evidence (published 2013–15) to develop a sugar-sweetened beverage (SSB) taxation law. This project followed previous policy research projects in Latin America on the same topic funded by IDRC and, according to the researchers, the success of that work provided legitimacy and interest for them to take up the issue in South Africa.

In March 2016, at the annual budget speech, the South African Ministry of Finance announced the introduction of the Health Promotion Levy, a tax on SSBs. The levy came into effect in April 2018. The project completion report notes that 'although the Department of Health was already considering such a tax, the evidence and dialogue generated by this project was central to the policy process and outcome'.

The taxation rate the South African government adopted was approximately 10-11%, which was less than the 20% recommended by the project. This was attributed to substantial concessions made to the beverage and sugar industries. To maintain momentum, IDRC funded a follow up project with Wits to evaluate the effects of the SSB tax. The evaluation found that announcing and introducing the sugar tax had led to a reduction in the sugar, calories and volume of beverage purchases (Stacey *et al.*, 2021).

As part of the project The Role of the Private Sector in Reducing Corruption in Latin America, a team of researchers from Fundación Universidad de San Andrés in Buenos Aires contributed to implementing a new law on corporate anticorruption in Argentina. The legislation establishes corporate criminal liability for corruption offences and makes it mandatory for certain public contractors to implement anticorruption compliance programs.

As well as drafting the initial bill, the research team also participated in and convened many debates in the lead-up to the law's promulgation. They organised a workshop for private sector executives to debate the technicalities of the proposed law, participated in the legislative debates both at the House of Representatives and the Senate, submitted written evidence to the legislative committee, participated in and spoke at private-sector organised fora, and published numerous policy papers, op-eds and a video. The National Congress approved the law (Law 27.401) on 8 November 2017, and published it in the Official Gazette on 1 December 2017.[16]

In an extension phase of the project in 2018, the team was able to further support implementation of the law by providing training to 72 private sector executives and 25 compliance trainers, as well as to public prosecutors and defenders in separate trainings, each of which was recorded and used as training materials in provincial offices. Because of their advocacy work, the project team was requested to support one of the main state-owned enterprises in Argentina, AySA, the water and sanitation company serving the metropolitan area of Buenos Aires, to build an anticorruption compliance program (project final technical report, Trackify, Abitbol *et al.*, 2019).

Regional level

The Scaling-Up and Evaluating Salt Reduction Policies and Programs in Latin American countries project built on previous research support by IDRC, which had contributed to policy change at national level in Argentina and Costa Rica to influence regional level policy. In this project, researchers from the Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud in Costa Rica, the University of Toronto Department for Nutritional Science and the University of Ontario Institute of Technology worked closely with the Pan American Health Organization (PAHO) to assess and compare the salt content of foods in five countries, analyse the health and economic benefits of population-wide sodium reduction and develop strategies for consumer behaviour change. The close collaboration with PAHO resulted in a commitment from PAHO to update the regional sodium reduction targets. The revised targets were due to be agreed at a meeting in May 2020 hosted by Brazil and PAHO (Staff survey, grantee survey and project final technical reports).

Global level

The Strengthening Scientific Evidence and its Use to Inform Policy Negotiation and Climate Implementation in Africa project supported the African Group of Negotiators Experts Support (AGNES) in its role in advising African institutions and governments at COP23 – the 23rd Conference of the Parties to the United Nations Framework

16. The law is published here: http://servicios.infoleg.gob.ar/infolegInternet/anexos/295000-299999/296846/norma.htm.

Convention on Climate Change (UNFCCC) held in Bonn in 2017. With the support of AGNES, the African Group of Negotiators (AGN) played a critical role at COP23, contributing to two agreements adopted by member states: the Gender Action Plan (GAP) and the Koronivia Joint Work on Agriculture (KJWA). The GAP recognises that climate change is not gender neutral and commits all countries to gender-responsive climate action. The KJWA establishes joint work on agriculture between the science groups and the implementation groups of the UNFCCC. AGNES's role was to convene the AGN's gender and agriculture negotiations prior to the conference to facilitate development of a common approach and preparation of the AGN submission on the two issues. AGNES also provided technical support to the AGN throughout the process of developing the GAP (Lafontaine *et al.*, 2019, Trackify, project interim technical reports).

There is another striking finding related to policy outcomes – the low number of policy capacity outcomes (24 identified in total) compared with the other types of policy outcome. The criteria for this category were strict as we were only looking for strengthened capacity of actors beyond grantees, not the grantees themselves, which are more often reported. The low occurrence of this type may be because programs put a stronger emphasis on influencing intellectual context (such as awareness and knowledge of a topic) than the institutional context (such as capacity to analyse and apply evidence or capacity to engage in policy dialogue), but in the opinion of the evaluation team, it is more likely to be because observing capacity outcomes is more difficult than observing when policy dialogues have been informed and polices have changed. Evaluations, for example, tend to imply strengthened capacity from observed policy change rather than assessing strengthened capacity explicitly.

This points to a potential gap identified by the evaluation; programs are not paying sufficient attention to policy capacity in their monitoring, evaluation and results reporting. This is important not just as a precursor to policy change but also as a significant outcome in its own right, and arguably more important than policy change. We believe if a program can show that it has strengthened policy capacity then it builds a stronger case for more sustainable change than if it can only show contribution to a specific policy change, which may have been an opportunistic policy opening. Some examples of the 24 policy capacity outcomes identified include the contribution of Deltas, Vulnerability & Climate Change: Migration & Adaptation (DECCMA, one of the CARIAA consortia) to the establishment of the Ghana National Expert Advisory Group, the Open and Collaborative Science in Development Network's work to create a policy group on Open Science at the Ministry of Science & Technology in Argentina, and the government of Kenya's launch of a research consortium to support Technical and Vocational Education and Training reforms, which was greatly informed by IDRC funded-research and in fact was led by IDRC at the government's invitation. These are all examples of how

IDRC projects support long-term scaling strategies, which do not yield tangible results at large-scale in the project timeline but strengthen the enabling environment for scaling in the future. This is where field building can potentially contribute to impact at scale as it may help to strengthen the enabling environment for sustainable scaling.

It should be noted at this point that the link to scaling in policy outcomes is less apparent than for innovation outcomes. For example, it was relatively straightforward for the evaluation team to distinguish between an innovation that had advanced to the point of being used by primary intended users and an innovation that reached beyond primary intended users (the last two boxes in the bottom part of the scaling pathway – Figure 1). The example above (Box 8) of the e-health record system in Peru is a good example to illustrate this: the project's primary intended users were the 15 health centres they worked with in the pilot, and this expanded to 350 health centred after the government agreed to implement the system across the country. Primary intended users are typically within the program's sphere of influence as participants in the research, albeit at the final stages, whereas the uptake and use of the innovation is managed by others with little involvement of the researchers (except possibly as advisers or evaluators).

The evaluation team found it more difficult to make this distinction with policy outcomes, at least in the way they are currently recorded in reports and evaluations. The majority of reported policy outcomes describe a change (e.g., a district development plan, adoption of a new national policy, a state-led program) and describe the contribution of the IDRC-funded research project in that change. However, there is rarely an indication as to whether the outcome goes beyond primary intended research users – i.e., if the policy change is at district level, was this the result of work in a different district which has been replicated? Or was the policy change the result of working with one set of actors that then went on to use the evidence to develop or implement the policy with other actors? Both of these would constitute clear scaling strategies for policy change but if approaches such as these are being followed then in most cases, we found that they are not being documented in way that makes it possible to distinguish.

The root of this challenge may be the high-level indicator itself (# new policies implemented or changed). In adopting this as a corporate indicator to monitoring the scaling objective, IDRC made an assumption that all contributions to policy change are relevant for scaling and hence can be aggregated for reporting progress against the scaling objective. However, the *Scaling Science* work, as presented in the *Scaling Playbook*, distinguishes between applied science, which is explained as supporting research use by primary knowledge users (e.g., policymakers), from scaling science, which considers 'the full range of initiators, enablers, competitors, and impacted who will support or hinder downstream results of the innovation' (which in this case is policy change). Therefore, a policy outcome can be considered a scaling outcome if it is clear that it has come about as part of a broader consideration of the full range of actors that

can support or hinder scaling. In the analysis done by the evaluation team, it was generally not possible to determine whether a policy outcome was in fact a scaling outcome from the way in which these outcomes are currently presented as most policy outcomes are mentioned briefly in the documents that we reviewed (that did not include in-depth project level documents like PCRs or FTRs), with little information about how they relate to the program's desired impact.

5.1.4. Equity and sustainability of outcomes

KEY TAKEAWAY: In the program area progress reports, program evaluations, and Trackify data examined, it was not possible for the evaluation team to determine the intended beneficiary groups for 86% of the outcomes identified. The way outcomes are reported in these sources is disconnected from the development outcomes to which the projects seek to contribute. This makes it difficult to determine the significance of the outcome – why it matters, whom it matters for and what might affect its sustainability.

As well as setting out to understand the outcomes that had been achieved through integrating scaling into programming, the evaluation team also set out to assess how valuable and sustainable those outcomes are and for whom. As mentioned in the methodology, this was a difficult question to answer because of the limited data set.

Looking first at the 'for whom' part of the question, we set out to determine the intended beneficiary group for each of the outcomes we identified (see Figure 14). By beneficiary group, we meant the people whose lives are intended to be improved as a result of the outcome – for example, which groups are intended to benefit from a policy change, or through the use of an innovation beyond primary users.

Overall, it was possible to determine the beneficiary group in only 17% (74 of 440) of identified outcomes. The most frequently reported beneficiary group was women (37 outcomes), followed by children and young people (25 outcomes) and then minority groups and other vulnerable populations (12). Looking at the type of outcomes shows a big difference; it was possible to determine the beneficiary group in 47% of innovation outcomes compared to 10% of policy outcomes. Looking at the source of data, the difference between data from Trackify and other sources is not so different; 15% and 18% respectively.

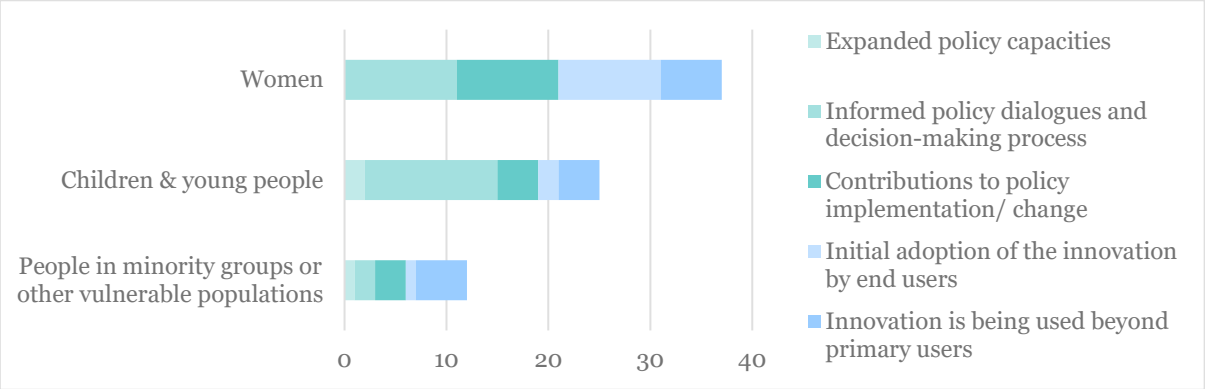

Figure 14: Outcomes by intended beneficiary group (n=74)

The key finding from our analysis is that the way outcomes are reported in evaluations, program reports or Trackify is disconnected from the outcome context, making it impossible in most cases to connect through to development outcomes. This is particularly a challenge with how policy change is reported, which doesn't often include an indication of why the policy matters and whom it matters for.

For example, the summative evaluation of CARIAA documents several policy outcomes such as 'in India, the research has made significant contributions to the implementation of the Maharashtra Groundwater (Development and Management) Act 2009 concerned with more sustainable ground water management and governance'. The policy change is clear but it is not made explicit who is intended to benefit from the new policy or how likely it that those benefits will be realised. Similar outcomes were documented in the evaluations of INASSA and the governance and justice program.

This was also observed by the evaluation team in program reports to the Board of Governors. For example, the Inclusive Economies report from 2018 includes this outcome: 'In Nigeria, the Delta State Ministry of Health set up, in response to research, Committees on essential drug monitoring and Water and Sanitation, and a mechanism to inventory health facility equipment for Primary Health Centres; these actions were complemented by regular facility supervision and increased sensitization and training of facility staff to improve quality of services.' Like the example above, it the change in practice is clear but it is not made explicit why this matters and who this matters for.

Consideration of the effects of policy change is particularly important for scaling when the risk of unintended (negative) impact is higher and trade-offs have to be negotiated. The CARIAA evaluation made a related observation:

> *An important caveat and concern about the [research into use] approach which all consortia face in all areas of research, not only gender, is the assumption that influencing policy will lead to more*

effective, appropriate and gender-sensitive adaptation outcomes on the ground. It is a necessary yet insufficient condition, as other assumptions (and risks) will influence the effectiveness of the policy.

For projects seeking to scale research results through policy, there is a clear challenge in understanding the impact of policy change given that researchers are rarely involved in implementation and may not get the opportunity to evaluate the policy (not to mention the high costs of policy evaluation). Many of the evaluations we reviewed remarked that it was too early to assess development outcomes or impact. However, there is an opportunity, and potentially a responsibility, for projects and programs to engage in foresight analysis to gain more understanding of the potential benefits and risks of the policy changes they are supporting. This kind of analysis is not apparent from the sources we have reviewed.

One explanation of the limitation of the sources is that they predominantly cater to program officer perspectives. Responsible officers commission program evaluations and submit data to Trackify. With each data source, we were therefore looking through a particular window, which resulted in simplification of the outcome information. We would have got a different view if we had looked more closely at grantee reports and/or interviewed more grantees about their outcomes.

The other driver that could explain this finding is that the corporate indicators, which directly shaped the Trackify data and indirectly influenced the other sources (e.g., through program impact pathways and program indicators), led to a greater emphasis on scaling up over scaling deep. For example, the focus of the indicators was on more policy change and more innovations used widely which is scaling up impact. Scaling deep would rather look for changes in quality of impact – variety, sustainability or equity, for example. The indicators are not designed to track these kinds of outcomes and so there is less of an incentive for programs to report them.

5.2. Internal outcomes

5.2.1. Stronger shared purpose and collaboration

KEY TAKEAWAY: The evaluation identified two main positive effects the scaling objective has had on the organisation and the way that programs work: a stronger shared purpose and collaboration, and a change in mindsets for many staff.

The strategic objective to scale required programs to consider how they can contribute to scaling research results. This created a stronger shared purpose around which program staff and the projects they support can identify and mobilise. Several interviewees told us that the strategic objective led to stronger collaboration and fewer

isolated projects pursuing distinct aims. For example, one interviewee said program officer roles had shifted from a focus on individual research projects to collaborative research agendas. Staff interviewees mentioned that the program impact pathways have helped strengthen a shared vision across the program and situate individual projects and partnerships within the portfolio.

Programs have also paid more attention to collaboration and partnerships with other actors, developing strategic relationships with the private sector, advocacy groups, government and other funders. One example provided by a staff interviewee was IDRC's work around open data, which started as a series of bottom-up research projects focussing on national policy. At the time the strategic objective to scale was being implemented, they had the opportunity to leverage that research for influence at global level by working with the World Bank and a network of donors and implementing agencies.

5.2.2. Changed mindsets

Interviews with IDRC staff suggest that the strategic objective has triggered discussion at project, program and corporate levels on what scaling means, how to pursue it and how to measure it. While we find that this has not always resulted in clear answers for programs, it is raising staff awareness around scaling.

Prior to 2015, a small number of programs engaged in scaling. Several interviewees named CIFSRF as a key contributor to learning about scaling; however, it typically approached scaling as something that comes at the end of the research process. Interviewees suggested that learning in this strategic period has shifted the thinking on this and it is now recommended that scaling research results should be considered from the research design stage. It is clear in the design of the KIX program that this learning has been applied.

While pursuing development outcomes is not new for IDRC, doing so through scaling as an explicit strategy is new, and, according to senior staff, it has required a change in culture and mindset: 'The culture has changed to be more open and comfortable with the idea that funding research for development means more than positioning for use and involves activities around the scaling of research' (SSI, IDRC staff).

The scaling objective has prompted programs to think about change differently. Scaling has added a dimension to the existing paradigm of research uptake. By considering how to support scaling research results, programs are shifting their ambitions and re-positioning their spheres of influence. For example, we were told that in one program, pre-2015, it had focused on grass-roots engagement, identifying priorities from communities and building relationships. In considering scaling, it shifted to engage in national policy debates and building a global network, involving higher level research.

The earlier work at community level was necessary but with a scaling lens it was possible to see how the local level efforts could build up to something with more influence.

6.CONCLUSION

IDRC committed to 'invest in knowledge and innovation for large scale positive change' as one of three strategic objectives in its 2015–2020 Strategic Plan. This evaluation set out to assess how well IDRC met this strategic objective and what can be learned from experience to inform the implementation of the next strategic plan. We address the former here and the latter in the next section on considerations.

We have characterised IDRC's experience with scaling as a **scaling journey**, reflecting how the practice of scaling across the Centre developed over the strategic period, and so too its ambition for what it could achieve. The Centre began its exploration of scaling prior to the introduction of the scaling objective and continues this journey into the next strategic period.

Along this journey, IDRC has developed a more nuanced understanding of scaling based on the experience of its staff and grantees across the global south. Yet IDRC's scaling journey in fact comprised many paths, as the Centre took a flexible approach to implementing scaling that allowed programs and projects to interpret and innovate around the strategic objective. Overall, we believe this was sensible given the varied experiences that already existed within the organisation. However, it also introduced challenges for programs. While some staff appreciated the flexibility, others found the lack of conceptual clarity around scaling terminology and approaches at IDRC a challenge in responding to the strategic objective. Lack of conceptual clarity or a common approach also created difficulties for monitoring progress, evaluating the results of the strategic objective and learning across programs.

The evaluation included a systematic assessment of **outcomes achieved through integrating scaling**. Through an assessment of a diverse range of sources including program evaluations, program area progress reports, Trackify and surveys and interviews and staff and grantees, the evaluation identified 440 outcomes that met the criteria. This included 32 innovations being used beyond primary intended knowledge/research users and 170 contributions to policy change. This is a considerable result for IDRC given that the original target for the scaling objective was to support at least 20 initiatives that deliver solutions at scale, but not entirely surprising in the context of IDRC's scaling journey, in which IDRC has refined its understanding of scaling research results (reducing the emphasis on 'large-scale' from the text of the scaling objective, as what defines 'large' was not established). We expect there to be more outcomes than this given that the assessment undertaken for this evaluation was not exhaustive, nor did it cover all programs systematically, additionally some of the sources used were 2–3 years old and it is possible that more has been achieved since that time.

The evaluation found that the introduction of the scaling objective also had **two main positive effects on IDRC**: a stronger shared purpose and collaboration within programs, with fewer isolated projects and a change in mindsets for many staff related to scaling, such as considering scaling earlier to inform research design and process rather than only at the end of the research process.

As programs experimented and innovated with **approaches to scaling**, they used and adapted various tools and processes to support their efforts. The design features that have emerged that contributed to scaling confirm the importance of flexibility and adaptability given that most of the challenges that IDRC is trying to address are complex, and that scaling itself is an intervention with evolving actors and factors. We found that, according to staff and grantees, flexible funding mechanisms were one of the most helpful tools to support scaling efforts, although the evaluation did not find evidence that they have been systematically documented or applied. We found that programs have been experimenting with different approaches to allow more time for research results to scale, although still more time is needed for the implementation phases of programs. Programs have found that investing in the supply and demand side of the scaling pathway simultaneously, with different kinds of investments across their portfolios, supports scaling. Research uptake and knowledge synthesis also emerged as important for supporting programs to scale research results. Research uptake by primary intended -users is a necessary pre-condition for scaling research results at optimal scale, and research synthesis helps programs to make sense of the investments across a portfolio and to identify opportunities for scaling. We found that although the research uptake is well supported in IDRC, research synthesis is less so.

The evaluation highlighted the important roles played by **program staff**. Responsible officers are crucial to scaling efforts and the evaluation found that they have been required to think and act more strategically and opportunistically. We found that many responsible officers are taking on this role, transitioning from funder and technical partner to knowledge broker, knowledge translator, coordinator and strategic thinker. However, there is not yet a formal recognition of this change in role and some staff feel that they have insufficient time, resources and incentives to carry it out effectively. Directors of program area and program leaders play an important role in setting the direction for programs and incentivising program staff to consider scaling. Program leaders in particular were identified as important decision makers when it came to incorporating scaling. Program management officers organise learning within the program and play a brokering role in sharing learning between programs, although this is an aspect of their role which is not well utilised for sharing learning on scaling across programs.

The 2015–2020 strategic plan recognised the importance of **coordinating with actors who can support scaling**, committing the Centre to work with 'public and private

sector actors who can advance ideas and innovation through to large scale implementation'. The evaluation found that IDRC programs are coordinating with a wider set of actors beyond the research community to support scaling research results, including by increasingly bringing non-research actors into program teams – such as private sector partners or policymakers within research project teams. The experience shows that bringing together diverse stakeholders is complicated, resource intensive and takes time. IDRC is still most comfortable working with researchers and the public sector but is increasing its understanding of how to work with the private sector, and many opportunities exist for further learning.

Similarly, programs are experimenting with coordinating structures to support scaling. These mechanisms enable scaling by facilitating collaboration and learning among grantees and facilitating interactions among the evolving set of actors involved in the scaling process. However, coordinating too many people and organisations is also cumbersome, taking up time for other activities and restricting the flexibility that is also important for adaptation.

The 2015–2020 strategic plan also committed IDRC to 'to be recognised for sharing its learning in scaling up solutions'. The evaluation has highlighted numerous opportunities both within and across programs for **learning on scaling**. We found that learning has occurred more within programs than across programs and it tended to be more conducive in less formal processes than in more formal processes. Externally funded programs were more likely to provide substantive learning opportunities than core funded projects, and learning tended to happen on a project-by-project basis rather than across a portfolio of projects. Staff reported that grantees were one of the most important sources for learning about scaling, and monitoring data was the least important way.

The most pressing challenge for learning mentioned by staff was extreme time pressure, with learning (about anything) often dropped to focus on what was perceived as more pressing corporate priorities. Staff told us they had little formal support to help them undertake systematic synthesis of knowledge across projects to aid their learning. They also told us they struggle to connect with and learn from staff beyond their programs, despite the introduction of collaborating officers on projects.

During the strategic period, IDRC undertook an exploration of scaling known as the **Scaling Science** initiative. It involved a review of over 200 projects and resulted in a published article, a book and numerous presentations sharing the learning, as well as a playbook to help researchers and innovators put learning into practice. Through this work, IDRC has developed a principle-based approach to scaling that is different to the dominant approach adopted by many similar organisations. While for many organisations, the scaling-up paradigm is still dominant, with a focus on a technology deployment, IDRC's approach espouses the principle of optimal scale and takes a

systemic perspective. This has helped position the Centre's experience within the growing field of scaling, and grantees and other donors recognise it for this contribution.

While appreciated by others, the Scaling Science approach is not yet fully socialised within the organisation, partly as the work was being developed during the strategic period, although much of it was done in the first half. Some concepts in the Scaling Science study have proved challenging for some staff and grantees to apply, particularly the focus on scaling impact rather than innovations, and to a lesser extent the concept of optimal scale. The reason for this may be that these concepts go against the mainstream discourse on scaling. We found that the principled approach offers a valuable contribution to scaling practice, as there can be no blueprint for scaling research results across diverse programs and contexts.

Finally, the evaluation found that **Trackify** is a useful source of data for results relating to scaling as it contains data that is not reported in other sources, for example, program evaluations and program area progress reports. However, the data is not reliable and should not have been used for corporate level reporting as there are many inconsistencies in how programs report results.

7. LESSONS LEARNED AND CONSIDERATIONS

We suggest the following considerations for IDRC to help build on existing strengths and address the challenges that emerged from the evaluation findings. We believe that many of these considerations will also be relevant and of interest to other organisations seeking to scale the impact of research results.

Corporate and cross-program considerations

7.1. Conceptualisation of scaling

While recognising the benefits of a flexible approach to scaling during the strategic period, having a broad and diverse conceptualisation of scaling makes it difficult to have conversations and learn about scaling across the Centre and with grantees. A unified conceptual understanding or definition of what is meant by scaling, scaling impact and optimal scale, could, for example, make it easier to identify common objectives, make it possible to choose the most appropriate approach and facilitate cooperation among the multiple parties and stakeholders that need to be involved in scaling efforts. Opinion among IDRC staff is mixed on whether it would be helpful to have a standard approach to scaling. We believe the Scaling Science study developed a middle ground, identifying guiding principles for scaling in a responsible way, while leaving space for different scaling strategies depending on the specific program or project context.

Looking ahead to IDRC's next strategic period, could a more unified approach to understanding scaling of research results benefit IDRC? Should the Centre provide enhanced support for staff and grantees to better understand and use the concepts introduced in the Scaling Science work?

7.2. Learning about scaling

The evaluation found that most staff learn from grantees and colleagues within their programs, and the evaluation team believes this should continue to be encouraged to ground learning about scaling in practice, but also that more support for cross-program learning would be beneficial. The Scaling Science initiative has facilitated learning across the Centre but more socialisation is required to build awareness as well as capacity for application. To encourage learning, the Centre and its programs could continue to encourage both formal and informal learning spaces and ensure that time and resources are set aside to promote learning. Learning about scaling will be facilitated by strengthened conceptual clarity of scaling across the Centre. It will be important to discuss challenges and failure to identify opportunities for improvement.

In what ways can the Centre's upcoming learning agenda facilitate learning about key cross-cutting issues such as scaling and enable conversations about challenges and failure in scaling given that scaling involves higher levels of risk in programming and that not all projects could or should scale?

7.3. Monitoring, evaluation and reporting on scaling

Monitoring, evaluation and reporting on scaling has presented some challenges for IDRC both in terms of what and how to monitor and report and assessing impact at scale within program timelines.

To improve the quality of monitoring and the usefulness of reporting scaling outcomes, IDRC could standardise the data collection criteria for the high-level indicators. One suggestion is to document outcomes using the outcome harvesting[17] format, which requires a description of change, a significance statement explaining why the outcome matters and who it matters for, and a contribution statement suggesting how the program contributed to the outcome (considering both the grantee's and IDRC's actions). Significance is particularly important for scaling as it clarifies how an otherwise isolated outcome is part of a more concerted effort to scale research results. It also connects the outcome downstream to those that are intended to benefit from the outcome. Pursuing this may require strengthening of capacity of those involved in program monitoring, including PMOs.

With clearer standards and stronger capacity for recording qualitative outcomes, it may also be possible to define other kinds of indicators for scaling, including policy/institutional capacity, demand for evidence and interrelationships between scaling actors, all of which are important for enabling scaling and can help to clarify outcomes beyond use by primary intended users.

Trackify could be better managed to provide reliable data, for example by creating a process for quality assurance which spans program and corporate levels, setting standards and providing guidance for data entry for each indicator, providing training on these for program staff where necessary, and by making the division of responsibility between program staff, POEV and the Information Management and Technology Divisions explicit.

There has been a range of experience with evaluation of scaling. Program evaluations reported on scaling results but varied in their treatment of scaling. Some programs adapted their evaluation approach to accommodate evaluation questions about scaling.

[17] Outcome harvesting is an approach to evaluation inspired by IDRC's outcome mapping methodology. For more information see: www.outcomeharvesting.net

For example, CARIAA took a staged approach to evaluation, which was perceived to have supported scaling by allowing them to develop learning questions throughout implementation. CIFSRF commissioned an evaluation to assess its contribution to food security, which assessed each project in terms of the prospects for scaling.

Many of the evaluations we reviewed remarked that it was too early to assess development outcomes or impact at scale. We found that particularly for programs seeking to scale impact through policy, there is a clear challenge in understanding the impact of policy change and that policy outcomes tended to be assessed in a way that was disconnected from their context (i.e. why the policy matters and whom it matters for). Similar to the observation above on monitoring, we believe there is an opportunity to expand the scope of program evaluation to examine the significance of policy change and better understand the potential equity and sustainability of outcomes. For example, foresight analysis could strengthen the understanding of the potential benefits and risks of the policy changes programs are supporting.

How can IDRC's scaling outcomes and contribution to impact at scale be tracked and assessed in a more systematic and reliable way? To what extent is it feasible and appropriate to expand the scope of program monitoring and evaluation to better examine the significance of outcomes and incorporate more analysis of the potential benefits and risks of impact?

Considerations in programming

7.4. Time horizons for investments in scaling

Scaling research results takes time. The findings from this evaluation suggest that the whole process of scaling research results for impact at optimal scale could be 10–15 years, if not longer, depending on the maturity of the research field. This means a long-term perspective is important for IDRC investment decisions, as is taking into consideration the program's position on the scaling pathway, what is realistic to achieve within the program timeframe, and what needs to be set up for sustaining scaling efforts beyond the initial investment. Strategies to support this include multi-phase projects and strategic partnerships with other funders, to continue and support scaling efforts beyond the Centre's investment.

Across IDRC program portfolios, what is the right balance for supporting longer-term investments, multi-phase projects and strengthening strategic partnerships specifically with a view to scaling research results? Are there other ways that the Centre can support programs to 'position themselves' to achieve impact at optimal scale, even if scaling impact beyond primary intended users itself takes place after the end of the IDRC funded project or program's lifetime?

7.5. Maximising a portfolio approach for scaling

Scaling research results requires knowledge, capacity and resources a scaling eco-system that span the scaling pathway – on both the supply and the demand side. The portfolio of projects within programs can be used strategically to facilitate this. The evaluation has found that IDRC programs that seek to scale research results have to ensure that sufficient attention is paid to the demand side, through for example, market-based approaches to scaling research results or policy influencing, alongside the supply side generation of new knowledge. In some cases, the field may have to be built to develop demand for an innovation or solution, working with different stakeholders.

Can a portfolio approach be more strategically used to build eco-systems for scaling by investing across a portfolio in projects that support both knowledge and innovation supply and demand solutions?

7.6. Flexible funding practices to support scaling

Flexible funding options support programs to scale research results because scaling is a dynamic process that requires adaptability to build on existing work with high potential or respond to emerging opportunities. Examples of flexible funding mechanisms that the evaluation identified as enabling scaling efforts included opportunity and rapid response funds.

Should IDRC leverage flexible funding mechanisms more systematically across the organisation to scale research results? Could formal criteria and process be beneficial to promote flexible funding and support a more consistent understanding across the Centre of the flexible funding options that can be used to support scaling?

7.7. Coordination for scaling

Coordination is clearly an important principle for scaling impact; ensuring the right actors are involved in the scaling process at the right time. This requires strategic thinking from within programs, and the logistics can be time consuming and resource intensive for both IDRC staff at program level and grantees at project level. It is important to build in flexibility and adaptability into coordination structures in order to facilitate collaboration and response to opportunities that support scaling.

How are the various coordination roles, both internally within a research program and externally with other stakeholders, best filled and by whom to support scaling? How can IDRC further support the competencies, skills and resources required to coordinate multiple stakeholders across a scaling pathway?

7.8. Knowledge translation and research synthesis

Knowledge translation and research synthesis has long been recognised by IDRC as an important function for research uptake, and it is just as critical for scaling research results. However, to support scaling it must be done with the objective of scaling in mind, and this requires strategic leadership within programs to focus the thought leadership of individual research teams. Consistent use of synthesis within programs – looking across the portfolio of projects – has been helpful in identifying opportunities for scaling.

Could IDRC provide additional support, both at the corporate level and within programs, to provide more time and resources for knowledge translation and synthesis work with a focus on scaling?

7.9. Support to grantees

IDRC's Grants-Plus approach is beneficial for scaling and staff have been supporting grantees to implement scaling that is responsible and in line with the guiding principles for scaling impact. The evaluation found evidence that staff are engaging in discussions about scaling with grantees early in the research process, and this practice should continue. Ongoing and evolving conversations about how to integrate the scaling principles into the research and scaling process will be important, in particular, optimal scale and the potential negative effects of scaling have been identified as areas needing improved support at present. These are important areas for designing and implementing inclusive scaling processes that promote equitable outcomes. However, the evaluation found that support for scaling increases program staffs' responsibilities, but these additional responsibilities are not yet formally recognised.

What additional support is required to facilitate more nuanced discussions with and among grantees about scaling (particularly about optimal scale and the potential negative effects of the scaling process) throughout the research process in a way that encourages responsible scaling?

Considerations for external influence

7.10. Positioning of IDRC in the scaling field

IDRC's principled approach is an important contribution and counterbalance to the mainstream thinking of 'bigger is better' and top-down approaches to scaling models. The Centre can use its unique position as a research for development funder to strengthen responsible scaling within the field, with a particular focus on the guiding principles for scaling impact of justification, coordination, optimal scale and dynamic evaluation. IDRC could also position questions of equity more centrally in discussions of scaling within its own programming, considering the existing strength of supporting gender transformative research.

To what extent and in what ways should IDRC continue to, or even strengthen, its influence on the evolving debates and dialogue in the scaling field, with the paradigm of principled scaling?

REFERENCES

Abitbol, Eric *et al.* (2019) 'Governance and Justice Program Evaluation'. Universalia Management Group. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/57597.

Acevedo Ruiz, M. and Pena-Lopez, I. (2017) 'Evaluation of the Open Data for Development Final Report'. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56510/IDL-56510.pdf?sequence=2&isAllowed=y.

Carden, F. (2009) *Knowledge to policy: making the most of development research*. Los Angeles, Ottawa: SAGE ; International Development Research Centre.

CARIAA (2018) 'Novel Insights Brief: Gender and Social Equity'. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57369/IDL-57369.pdf?sequence=2&isAllowed=y.

Cathexis (2019) 'Distilling Key Learning across IDRC's Program-Led Evaluations: Final Report'. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/58310.

Cochrane, L. *et al.* (2017) *Enabling synthesis: lessons from the first three years of CARIAA*. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/56605.

Cochrane, L. and Cundill, G. (2018) 'Enabling collaborative synthesis in multi-partner programmes', *Development in Practice*, 28(7), pp. 922–931. doi: 10.1080/09614524.2018.1480706.

Cundill, G. *et al.* (2019) *Large-Scale Transdisciplinary Collaboration for Adaptation Research: Challenges and Insights*.

Decosas, J., Deville, M. and Medina, M. (2020) *Innovating for Maternal and Child Health Summative Evaluation*. HERA.

Diosady, L., Mannar, V. and Menon, J. (2018) 'CIFSRF final technical report: scaling up the production and distribution of double fortified salt in India (CIFSRF Phase 2)'. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/57332.

EA-HPRO (2020) 'EA-HPRO Fact Sheet', November. Available at: http://www.ea-imcha.com/index.php/en/resources/publicationss/ea-hpro-fact-sheets-1.

Fundación Capital (2018) 'Enhancing economic opportunities in Latin America: from poverty reduction projects to sustainable livelihoods'.

Gonsalves, T. and Baranyi, S. (2003) 'Research for policy influence: a history of IDRC intent'. Available at: www.researchgate.net/publication/265399751_Research_for_Policy_Influence_a_History_of_IDRC_Intent.

Guijt, I. (2010) *Accountability and Learning In Capacity Development in Practice*. London, UK: Earthscan.

Harvey, B. *et al.* (2019) 'Learning, landscape and opportunities for IDRC climate programming'. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/57623.

Hera (2020) 'Issue brief on gender integration'. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/59620/IDL%20-%2059620.pdf?sequence=2&isAllowed=y.

IDRC (2011) *IDRC Annual Corporate Evaluation Report*, p. 5. Available at: IDL-47408.pdf (dspacedirect.org).

IDRC (2015) 'Investing in Solutions: Strategic Plan 2015–2020'. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57216/IDL-57216.pdf?sequence=2&isAllowed=y.

IDRC (2017) 'Technology and Innovation: Progress Report to the Board of Governors'. Available at: http://ic.idrc.ca/programs/si/_layouts/15/WopiFrame2.aspx?sourcedoc=/programs/si/smcboard/TI%20Progress%20Report_FINAL%20Feb%2027.docx&action=default.

IDRC (2018a) 'CARIAA Novel Insights Brief'.

IDRC (2018b) 'Investing in Solutions: Performance and Learning Report 2017-2018 - Presented to the Board of Governors'. Available at: http://ic.idrc.ca/sites/cse/2015-2020/_layouts/15/WopiFrame.aspx?sourcedoc=/sites/cse/2015-2020/Performance/2017%20-%202018%20Performance%20and%20Learning%20Report/Performance%20and%20Learning%20Report%202017-2018.docx&action=default.

IDRC (2018c) 'Technology and Innovation: Progress Report to the Board of Governors'. Available at: http://ic.idrc.ca/sites/projects/_layouts/15/WopiFrame.aspx?sourcedoc=/sites/projects/project/108924%20Evaluating%20Research%20Results%20at%20Scale/TI%202018.docx&action=default.

IDRC (2019a) 'Performance and Learning Report 2018–2019: DRAFT to be Presented to the Board of Governors'. Available at: http://ic.idrc.ca/sites/cse/2015-2020/_layouts/15/WopiFrame.aspx?sourcedoc=/sites/cse/2015-2020/Performance/2018-2019%20Performance%20and%20Learning%20Report/Performance%20and%20Learning%20Report%202018-2019-Draft.docx&action=default.

IDRC (2019b) *Transforming gender relations: Insights from IDRC research*. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57633/IDL-57633.PDF?sequence=2.

IDRC (no date) 'International Development Research Centre, Private Sector Engagement Strategy 2030 (draft)'.

IDRC Grow Team (2018) *Growth and Economic Opportunities for Women Final Report of a 5-year research program 1 June 2013 – 31 May 2018*. Final report. IDRC.

IDRC Policy and Evaluation Division (2019) 'Synthesis Learning Report: Strategic Plan 2015-2020'.

Jones, H. (2012) 'Promoting evidence-based decision-making in development agencies. Background note.' ODI. Available at: www.odi.org.uk.

Kohl, R. (2021) *Crosscutting Issues affecting Scaling: A Review and Appraisal of Scaling in International Development*. Global Community of Practice on Scaling Development Outcomes.

Lafontaine, A. *et al.* (2018) 'Collaborative Adaptation Research Initiative in Africa and Asia (CARIAA) Summative Evaluation'. Le Groupe-conseil baastel Itée. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57296/IDL-57296.pdf?sequence=2&isAllowed=y.

Lafontaine, A. *et al.* (2019) 'Climate Change Program External Evaluation'.

Linn, B. *et al.* (2010) *Scaling up the fight against rural poverty*. Working Paper 43. Brookings. Available at: www.brookings.edu/wp-content/uploads/2016/06/10_ifad_linn_kharas.pdf.

McLean, R. and Gargani, J. (2019) *Scaling impact: innovation for the public good*. London; New York, Ottawa: Routledge, Taylor & Francis Group; International Development Research Centre. Available at: https://www.idrc.ca/en/book/scaling-impact-innovation-public-good.

McLean, R., Gargani, J. and Lomofsky, D. (2020) 'Scaling what works doesn't work: we need to scale impact instead', *LSE Blogs*, September. Available at: https://blogs.lse.ac.uk/impactofsocialsciences/2020/09/07/scaling-what-works-doesnt-work-we-need-to-scale-impact-instead/.

Milena Buitrago Rodriguez, C. (2018) 'CIFSRF final technical report: Scaling up the production of more nutritious yellow potatoes in Colombia (CIFSRF Phase 2)'. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/56982.

Moynihan, D. P. and Landuyt, N. (2009) 'How do Public Organizations Learn? Bridging Structural and Cultural Perspectives'. Available at: https://lafollette.wisc.edu/images/publications/facstaff/moynihan/MoynihanLanduytPAR69(6).pdf.

O'Neill, M. and Manchur, W. (2018) 'Contributions to end global hunger: a synthesis of the Canadian International Food Security Research Fund (CIFSRF)'. Available at: http://hdl.handle.net/10625/57282.

Patton, M. Q. (2002) *Qualitative research and evaluation methods*. 3 ed. Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2008) *Utilization-focused evaluation*. 4th ed. Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2018) *Principles-focused evaluation: the guide*. New York: Guilford Press. Available at: www.guilford.com/books/Principles-Focused-Evaluation/Michael-Quinn-Patton/9781462531820/contents.

Pérez-Lu, J. E., Bayer, A. M. and Iguiñiz-Romero, R. (2018) 'Information = equity? How increased access to information can enhance equity and improve health outcomes for pregnant women in Peru', *Journal of Public Health*, 40(suppl. 2), pp. ii64–ii73. doi: 10.1093/pubmed/fdy177.

Price-Kelly, H., van Haeran, L. and McLean, R. (2020) *The Scaling Playbook: A Practical Guide for Researchers*. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/58780/IDL-58780.pdf?sequence=2&isAllowed=y.

Sanchez-Swaren, E. (2018) *Evaluation on Achieving Results at Scale - Mapping Analysis*. IDRC.

Sauveplane Stirling, V. *et al.* (2019) *External review of IDRC's Food, Environment, and Health (FEH) Program 2015–2020: final report*. Final report. Toronto: University of Toronto Press. Available at: https://idl-bnc-idrc.dspacedirect.org/handle/10625/58260.

Shilomboleni, H. *et al.* (2019) 'Scaling up innovations in smallholder agriculture: Lessons from the Canadian international food security research fund', *Agricultural Systems*, 175, pp. 58–65. doi: 10.1016/j.agsy.2019.05.012.

Stacey, N. *et al.* (2021) 'Changes in beverage purchases following the announcement and implementation of South Africa's Health Promotion Levy: an observational study', *The Lancet Planetary Health*, 5(4), pp. e200–e208. doi: 10.1016/S2542-5196(20)30304-1.

Tewes-Gradl, C. and Elliot-Gaved, F. (no date) *IDRC Employment and Growth Private Sector Engagement Review (unpublished report)*.

Universalia (2016) *Evaluation of the Canadian International Food Security Research Fund (CIFSRF)*. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56897/IDL-56897.pdf?sequence=2&isAllowed=y.

Wiggins, S. *et al.* (2018) 'Understanding the CIFSRF Phase Two portfolio's overall contribution to food security'. Overseas Development Institute. Available at: https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/57230/IDL-57230.pdf?sequence=2&isAllowed=y.

World Health Organization (2010) *Nine Steps for Developing a Scaling-Up Strategy*. Geneva: World Health Organization; ExpandNet.

Wyborn, C. *et al.* (2018) 'Understanding the Impacts of Research Synthesis', *Environmental Science & Policy*, 86, pp. 72–84. doi: 10.1016/j.envsci.2018.04.013.

APPENDIX 1: METHODOLOGY

Overall approach

Mixed-method, utilisation focused strategic review: This was a mixed-method, user-focused strategic review. It comprised discrete studies that together helped answer the evaluation questions. The main components included: (1) an organisational review, (2) a study of IDRC's positioning within the wider sector, (3) perceptions and experiences of grantees, and (4) thematic case studies: field building for scale, programming for scale, private sector engagement and scaling, and organisational learning about scaling.

Summative and learning focus: The evaluation had two main focus areas. First, it was summative – looking back at how IDRC has implemented scaling and what results it has achieved in so doing. Second, it had a learning component – to support IDRC to learn from its experience of scaling to inform future work. Learning was also offered by positioning IDRC's experience with scaling against that of other research funders.

Engagement and sensemaking: Our approach involved engaging with IDRC throughout the evaluation, inviting staff to provide input and comments at critical stages and to join us in sensemaking of the data. An evaluation advisory group comprising representatives from across the organisation participated in an inception workshop and an 'emerging findings' webinar and took part in a 'findings to recommendations' workshop.

Components

(1) Organisational review

Purpose:

- To identify how IDRC adjusted its organisational systems and processes to adapt to the scaling objective.
- To identify strengths and challenges, and to make recommendations for improvement.
- To evaluate the outcomes IDRC achieved.

Methods:

- Document review
- Online self-completion survey of IDRC staff
- Semi-structured interviews with senior management, program managers and other staff

- Analysis of program outcomes.

Sampling and sample:

- 43 IDRC staff responded to the survey that was sent via email to all program staff: 24 senior program specialists, 6 program management officers, 6 senior program officers, 3 program officers, 3 program leaders and 1 director of program areas. The open call survey was promoted via email within IDRC.
- Interviews with 10 senior staff. It was a purposeful sample prepared by the evaluation team in collaboration with the POEV – the criterion for inclusion was to have a diverse set of perspectives from different parts of the organisation involved in implementing the strategic objective to scale.

The approach to the analysis of outcomes is described separately in Appendix 2.

(2) Positioning study

Purpose:

- To identify IDRC's niche and position relative to other research funders.
- To learn from the practice of others and identify recommendations for improvement.

Methods:

- Online self-completion survey of other funders of research for development
- Semi-structured interviews with other funders of research for development
- Document review.

Sampling and sample:

- We drew up a list of 38 relevant external organisations by reviewing and extracting names from the Scaling up Community of Practice's newsletters and soliciting additional relevant funders and research institutes from IDRC staff. The criterion for inclusion was that the organisations had to be funders of development research who were known to be involved in scaling. The survey was sent to the full list of potential respondents and 16 funders responded to the survey: 5 bi-lateral, 4 public-sector funded, 3 multi-lateral, 3 think tank, 1 philanthropy/foundation.
- We conducted 15 semi-structured, qualitative, virtual interviews with interviewees purposively selected by the evaluation team in collaboration with the POEV to cover a range of research funders, including multi-lateral

organisations, bi-lateral organisations, public-sector funded organisations, think tanks and philanthropic organisations. We followed a two-stage sampling process. First, we used criterion sampling to identify a shortlist of these, and the criterion was to select those who were known by IDRC to have been involved in scaling initiatives and who would be good comparators for IDRC. The second stage was convenience sampling – we contacted those organisations / people for whom we could find contact information.

(3) Perceptions and experience of IDRC grantees

Purpose:

- To assess the perception and experience of IDRC grantees in relation to practice of scaling and scaling results.
- To identify strengths and challenges, and to make recommendations for improvement.

Methods:

- Online self-completion survey of grantees
- Focus groups with grantees.

Sample:

- We shared the open call survey via IDRC's newsletter, the OTT and IDRC websites, and social media channels; programs were also encouraged to share it with their grantees.
- 95 grantees responded to the survey: 30 from Latin America and the Caribbean, 20 from Sub-Saharan Africa, 18 from Canada and the US, 7 from the Middle East and North Africa, 7 from West Africa, 6 from Asia, and 4 from Europe and Central Asia.
- 18 grantees participated in five, 90-minute, online focus group meetings. Participants were from four regions: Latin America, Francophone Africa, Anglophone Africa and Asia. Participants were volunteers drawn from two sources. Invitations to participate in focus group discussions were sent to all IDRC scalingXchange participants (which brought IDRCs southern grantees together to discuss scaling); and to survey respondents who indicated that they were also willing to participate in focus groups.

(4) Case studies

Four thematic case studies areas were identified in collaboration with IDRC: (1) field building for scale, (2) programming for scale, (3) private sector engagement and scaling, and (4) organisational learning about scaling.

Purpose:

- To gain a more in-depth understanding of how IDRC has implemented scaling and the outcomes achieved in particular thematic areas.
- To learn more about IDRC's guiding principles for scaling impact, their relevance and application.
- To identify strengths and challenges.

Evaluation case study 1, field building for scaling impact

Methods:

- Review of a small number of prominent articles, IDRC documents and evaluations of IDRC field building programs and projects.
- Semi-structured interviews (SSIs) with IDRC staff, grantees and a co-funder.

Sample:

- 12 interviews: 7 program staff, 4 grantees and 1 co-funder. Purposeful sampling used by the evaluators in collaboration with POEV to select respondents from programs that have been involved in field building. Programs not included in the interviews were included in the document review.

Evaluation case study 2, programming for scale

Methods:

- Review of IDRC program-level documentation including evaluation reports, synthesis pieces and journal articles.
- SSIs with IDRC staff, grantees and external stakeholders.

Sample:

- 20 interviews: 8 program staff, 8 grantees, 4 external stakeholders, such as funders and external service providers. We used a purposeful criterion sampling to select 'best cases' demonstrating good practice and results for scaling impact that could form the basis for learning and recommendations. The sample of programs included those with a mix of projects that are implementing scaling in different ways such as innovations taken to scale (e.g. products, methods, models); scaling impact by strengthening system elements (e.g. policy, finance,

social norms, human resources); or taking tried and tested ideas from one place to another (institution, country, region). They include a mix of externally co-funded programs and one core program.

Evaluation case study 3, private sector engagement and scaling

Methods:

- Review of IDRC, grantee and other funder documents.
- Brief literature review.
- Review of survey results.
- SSIs with IDRC staff.

Sample:

- 15 interviews: 12 IDRC staff (including 1 former staff member) and 3 grantees. A purposeful sampling was used to select 'best cases' demonstrating good practice of either collaborating with the private sector or mitigating the negative effect of the private sector on scaling impact.

Evaluation case study 4, organisational learning about scaling

Methods:

- Review of a small number of prominent articles, syntheses and meta-analyses that explore organisational learning primarily within public institutions.
- SSIs with IDRC staff.

Sample:

- 18 interviews: program officers, program management officers, program specialists, senior program specialists, program leaders and regional director. The sample was gender balanced and included staff with varying lengths of service at IDRC ranging from 1–20 years and covering a number of IDRC program areas and divisions: Agriculture and Environment, Inclusive Economies, Technology and Innovation, the Asia Regional Office and POEV.

Overall sample

Table 3 presents the total sample for all components, including both qualitative and quantitative data collection.

- Internal stakeholders are directly involved in IDRC programs: IDRC staff and grantees.
- External stakeholders are not directly involved in IDRC programs: grantee partners or collaborators, actors in the systems they are trying to influence, or others not involved in IDRC's scaling work.

Table 3: Total sample for all components (excluding the inception phase)

Component	IDRC	Grantees	External
Organisational review	10 semi-structured interviews (SSIs) with senior staff 43 survey responses from program staff	-	-
IDRC's positioning	-	-	15 SSI 16 survey responses
Case study 1: Field building for scale	7 SSIs	4 SSIs	1 SSI
Case study 2: Programming for scale	8 SSIs	8 SSIs	4 SSIs
Case study 3: Private sector engagement and scaling	11 SSIs with current staff, and 1 with former staff member	3 SSIs	-
Case study 4: Organisational learning about scaling	18 SSIs	-	-
Grantee experiences and perceptions	-	5 focus group discussions (FGDs) with 18 grantees from 4 regions	-

		95 survey responses	
TOTAL	55 SSIs	13 SSIs	20 SSIs
	43 survey responses	5 FGDs with 18 grantees	16 survey responses
		95 survey responses	

Three of the case studies had a thematic focus (private sector and scaling; field building for scale; and programming for scale). These three case studies covered 10 programs[18] and 14 projects. Of these 10 programs, 6 were examined at the program level only; and 4 included program and project level. Tables 4 and 5 presents these by program area and case study respectively. The fourth case study on organisational learning did not focus on particular programs or projects but instead randomly sampled the short-list of projects and programs to identify program staff to interview about their experiences with learning about scaling.

Table 4: Projects and programs reviewed in each case study, by program area

Program area	Program level (including co-funded programs)	Projects included in more depth	Case studies
Agriculture & Environment	Food, Environment and Health (FEH) program		Private sector and scaling, Field building, Programming for Scale
	Livestock Vaccine Innovation Fund (LVIF)	Development of Cross-protective Synthetic RNA Vaccine against Foot and Mouth Disease (FMD) project	Private sector and scaling, Field building for scale

17. Programs include both IDRC programs and a sample of the co-funded programs they manage.

		Development of Two Multivalent Rift Valley Fever Vaccines for Improved Uptake in Cattle and in Small Ruminants project InnoVet AMR for Animal Health project	
	Agriculture and Food Security (AFS)		Field building for scale
	Canadian International Food Security Research Fund (CIFSRF)	Scaling up Production and Distribution of Double-fortified Salt in India project Achieving Impact at Scale through ICT-enabled Extension project	Private sector and scaling, Programming for scale,
	Cultivate Africa's Future Fund (CultiAF) (Project level only)	Gender Inclusive Financing for Scaling up Improved Fish Processing Technologies in Malawi project	Private sector and scaling
	Climate Change Program (Project level only)	Mobilising the Private Sector for Adaptation Finance project	Private sector and scaling

	Collaborative Adaptation Research Initiative in Africa and Asia (CARIAA)	Pathways to Resilience in Semi-arid Economies (PRISE) project	Private sector and scaling, Programming for scale
Inclusive Economies	Employment and Growth (EG) program	Promoting Responsible Value chains in India for Effective Contribution of the Private Sector to the Sustainable Development Goals project Policy Analysis on Growth and Employment (PAGE II) project Women in Trade Knowledge Platform to Boost Inclusive and Sustainable Growth project	Private sector and scaling
	Growth and Economic Opportunities for Women (GrOW)		Field building for scale
	Innovating for Maternal and Child Health in Africa initiative (IMCHA)		Programming for scale, Field building for scale
Technology & Innovation	Networked economies (Project level only)	Teacher Professional Development at Scale (TDP@Scale)	Field building for scale

	Open data for Development (OD4D)		Field building for scale
	Global Partnership for Education, Knowledge and Innovation Exchange (KIX)		Programming for scale
	Foundations for Innovation Program (project level only)	Access to Finance for SMEs in Least Developed Countries project Research and Development Impact Vouchers project	Private sector and scaling

Table 5: Programs and projects reviewed in each case study, by case study

Case study	Programs and projects
Case study 1 Field building for scale	• Growth and Economic Opportunities for Women (GrOW) • Innovating for Maternal and Child Health in Africa initiative (IMCHA) (• Food, Environment and Health (FEH) program • Livestock Vaccine Innovation Fund (LVIF) • Agriculture and Food Security (AFS) program • Open Data for Development (Od4D) • Teacher Professional Development at Scale (TPD@scale)
Case study 2 Programming for scale	• Canadian International Food Security Research Fund (CIFSRF); Collaborative Adaptation Research Initiative in Africa and Asia (CARIAA); Food Environment and Health (FEH, including NCDs and rapid response work on Ebola, Zika and Covid-19) • Innovating for Maternal and Child Health in Africa initiative (IMCHA) • Knowledge and Innovation Exchange (KIX)
Case study 3	Programs reviewed: • Food, Environment and Health (FEH) program • Employment and Growth (EG) program

Private sector engagement and scaling	Projects reviewed: • Promoting Responsible Value Chains in India for an Effective Contribution of the Private Sector to the Sustainable Development Goals project (EG program) • Policy Analysis on Growth and Employment – PAGE II project (EG program) • Women in Trade Knowledge Platform to Boost Inclusive and Sustainable Growth project (EG program) • Mobilising the Private Sector for Adaptation Finance project (Climate Change program) • Pathways to Resilience in Semi-arid Economies (PRISE) project (Collaborative Adaptation Research Initiative in Africa and Asia (CARIAA) program) • Gender Inclusive Financing for Scaling up Improved Fish Processing Technologies in Malawi project (Cultivate Africa's Future Fund (CultiAF)) • Scaling up Production and Distribution of Double-fortified Salt in India project (Canadian International Food Security Research Fund (CIFSRF) program) • Achieving Impact at Scale through ICT-enabled Extension project (CIFSRF program) • Development of Cross-protective Synthetic RNA Vaccine against Foot and Mouth Disease (FMD) project (Livestock Vaccine Innovation Fund (LVIF) program) • Development of Two Multivalent Rift Valley Fever Vaccines for Improved Uptake in Cattle and in Small Ruminants project (LVIF program) • InnoVet AMR for Animal Health project (LVIF program) • Access to Finance for SMEs in Least Developed Countries project (Foundations for Innovation program) • Research and Development Impact Vouchers project (Foundations for Innovation program)
Case study 4 Organisational learning about scaling	Not project specific

Documentary sources

The evaluation drew on a variety of documentary sources. **Error! Reference source not found.** describes the main types of documents reviewed across all components of the evaluation.

Table 6: Types of documentary sources

Internal	• IDRC corporate documents, e.g., strategy and results framework • Reports to the Board of Governors, e.g., corporate performance and learning report and program area progress reports • Program area implementation plans • Grant management documents and grantee reports • Evaluations of IDRC programs/grants • Corporate level evaluations • IDRC publications sharing insights from research • Various articles and books on reflections on scale, including the scaling impact book
External	• Published articles or reports on scaling • Grant-making documents from other funders that relate to scale.

Analysis

Qualitative

- We developed a coding framework based on the evaluation framework. We coded the data using Dedoose on the basis of coding reports, coding each case study using the key case study questions that were aligned to the framework.
- We conducted a document review to analyse the relevant documents as part of the analysis phase.
- We wrote up the grantee component and the positioning study as discrete pieces, and then integrated them in the main report.
- We synthesised the coding reports for the integrated report and re-coded the case studies using an evidence matrix in MS Excel.

Quantitative

- We carried out quantitative data analysis using a combination of Kobo Toolbox Data Analyser and ArcGIS.
- We developed a data analysis plan, and cross-tabulated on key questions.

APPENDIX 2: METHODOLOGY FOR OUTCOME ANALYSIS

Data sources

1. Program evaluations

We identified 13 program evaluations from 7 programs through a purposive sampling approach. We first looked for evaluations for all the programs considered in the evaluation case studies. We found 11 evaluations through this approach for all programs except KIX, LVIF and Foundations for Innovation. Over half the evaluations (6) were in the Agriculture and Environment program area so we selected two additional program evaluations from the other two program areas: one core program evaluation and one externally funded program evaluation. This resulted in a sample of 13 program evaluations:

1. Climate Change
    a. CARIAA Staged Evaluation Second Thematic Review – Application of Hotspot Approach
    b. CARIAA Summative Evaluation
    c. Climate Change Program External Evaluation
2. Agriculture and Food Security
    a. Evaluation of the CIFSRF
    b. Understanding the CIFSRF Phase Two portfolio's overall contribution to food security
3. Food Environment & Health
    a. External Review of IDRC's Food, Environment, and Health (FEH) Program 2015–2020 – Final Report
4. Governance and Justice
    a. Governance and Justice Program Evaluation
5. Maternal and Child Health
    a. IMCHA: A Mid-Term Formative Evaluation
    b. IMCHA Summative Evaluation
6. Employment and Growth
    a. GrOW Formative Evaluation for Mid-Term Review
    b. IDRC Employment and Growth Learning Evaluation
7. Networked Economies

a. Evaluation of the Information and Networks in Asia and Sub-Saharan Africa (INASSA) Program
b. Evaluation of the Open Data for Development Program

2. *Program area progress reports to the board*

We reviewed 9 program area progress reports to the board:

1. A&E: 2017, 2018, 2019
2. IE: 2016, 2018, 2019
3. T&I: 2017, 2018, 2019

3. *Staff survey*

The staff survey included two qualitative questions regarding scaling outcomes, neither with a particularly high response rate but sufficient for qualitative review.

1. Q22: Please provide an example that explains how an approach to scaling impact in one of your projects led to better development outcomes? (15/43)
2. Q 24: Please provide an example about what was scaled as a result of a research project. (19/43)

4. *Grantee survey*

The grantee survey included three questions with qualitative responses relating to scaling outcomes, again with a relatively low response rate but sufficient for review.

1. Q17: In your opinion, in what way did the approach to scaling impact in your research project lead to better development impact? (17/93)
2. Q18: Can you provide an example that explains how your approach to scaling impact leads to better development outcomes? (46/93)
3. Q20: Can you provide more detail on what was scaled as a result of your research? (46/93)

5. *Interviews*

Some of the interviews conducted with staff and grantees provided additional details about particular program outcomes.

6. *Trackify*

We extracted two datasets from Trackify and provided them to the evaluation team in June 2020:

1. All evidence values linked to corporate indicator: # Innovations being widely used and adopted (170 outcomes from 99 entries)
2. All evidence values linked to corporate indicator: # New policies implemented or changed (482 outcomes from 140 entries)

Data analysis

1. Qualitative analysis

The team reviewed the program evaluations, program area progress reports, staff and grantee survey responses and interview transcripts to extract distinct outcomes. We used the following criteria to identify plausible outcomes:

1. Describes the actions of an actor external to the project (i.e. not just about what the project did).
2. Describes the contribution of the project (i.e. there has to be a plausible link to the project).
3. Provides verifiable details (dates, places, events, names or organisations).
4. Can be categorised using the outcome types defined in the scaling pathway:
    a. Policy outcomes:
        i Expanded policy capacities
        ii Informed policy dialogues and decision-making process
        iii Contributions to policy implementation/change
    b. Innovation outcomes:
        i Initial adoption of the innovation by end users…
        ii Innovation is being widely used and adopted…

Once identified, each outcome was recorded in a dataset and categorised according to:

I The outcome type (defined above)

II The region where the outcome occurred (Asia, Eastern and Southern Africa, Central and West Africa, Middle East and North Africa, Latin America and the Caribbean, Other, Global, Unspecified)

III Geographic level (community, municipal / district, sub-national, national, regional, global, unspecified)

IV Target beneficiaries (women, children and young people, people in minority groups, unspecified)

V IDRC core program

2. Trackify analysis

Trackify data was first cleaned to remove erroneous entries. We applied the following steps:

1. Entries which could not be identified as an outcome were removed (i.e. there was no description of the actions of an external actor) (238 outcomes were removed).

2. Any entries where the description was insufficient to determine the outcome or the program's contribution were marked as 'insufficient evidence' (125 outcomes marked).
3. Duplicates were removed (including duplicates across the two datasets).
4. The 'Value' field (used to state how many outcomes are reported in a single entry) was corrected where necessary to ensure consistent counting (e.g. some entries used the value field to state how many sites an innovation had been scaled to whereas most reported how many innovations had been scaled – we opted for the latter).

Once identified, we recorded each entry in the outcome dataset and categorised using the same fields as above:

I The outcome type

II The region where the outcome occurred

III Geographic level

IV Target beneficiaries

V IDRC core program.

Trackify data discussion

Our analysis of data from Trackify highlighted several useful findings relating to the quality and usability of Trackify data and the associated corporate indicators. While this does not pertain directly to results achieved through scaling, it does relate to the methods used to monitor and report on progress against the scaling objective.

Figure 15 presents the breakdown of how we classified the outcomes in the two Trackify datasets (innovations and policies) – in the middle row are the two datasets from Trackify with 170 and 482 evidence values[19] respectively, along the top row are the five outcome types we have classified and along the bottom row are the evidence values we did not count as outcomes.

We found that with clear criteria it is possible to identify plausible outcomes from the Trackify dataset and it is a useful source of data that is complementary to and not duplicative of program evaluations. With a review of the indicators, some adjustments to the data entry guidelines to ensure higher consistency and a quality assurance mechanism in place, the system can provide usable data. However, we do not consider

18. An evidence value is an individual result submitted to Trackify. Some Trackify entries report more than one evidence value, for example, one project may report five contributions to policy change in one submission. This would count as five evidence values.

the data as usable in any way without a detailed review such as been done here. We identified the following issues as particularly concerning:

1. 238 of the 652 (36%) of the evidence values were classed as not an outcome. For innovations, this usually means they were classed as either proof of concept, working prototype or user ready innovation, which we identify as an output. For policies, this usually means they are describing research outputs or activities, or policy recommendations.
2. 166 (25%) of the evidence values had insufficient evidence to classify them. This usually means there is enough information in the Trackify entry to see a potential outcome but either it is not clear who or what has changed or how the project reportedly contributed to this change. This could be a factor of over-simplification in the Trackify data entry – which was purposefully limited.
3. 37 (5.7%) of the evidence values were either duplicates (e.g., occasionally the same data was present in the policy and innovation datasets) or the value (the number of results associated with that entry) was incorrectly reported and had to be adjusted (e.g., occasionally the value was given as the number of sites where an innovation was being used rather than the number of innovations).
4. Only 6 of the 170 (3.5%) evidence values tagged as innovations could be plausibly classified as innovations used beyond primary users, suggesting that the program indicators had not been well matched to the corporate indicator. In fact, there were more innovations used beyond primary users identified in the policy dataset (10) than in the innovation dataset suggesting that programs require more guidance on definitions of what counts as an innovation versus a policy outcome, or that the indicators are insufficiently nuanced to capture program outcomes.
5. 97 of the 482 (20%) policy outcomes were classified as contributions to policy change, which means most (80%) of the outcome reported as contributions to policy change could not be identified as such by the evaluators. This could be due to misalignment of program indicators to corporate indicators but in the opinion of the evaluation team is more likely to be due to programs having insufficiently rigorous criteria for data entry.
6. The overlap between results reported in Trackify and those presented in program evaluations was low. Only a few outcomes were common to both with the majority only occurring in one or other source. This suggests there is value in continuing to combine external assessment with internal monitoring and not to rely on either individually.

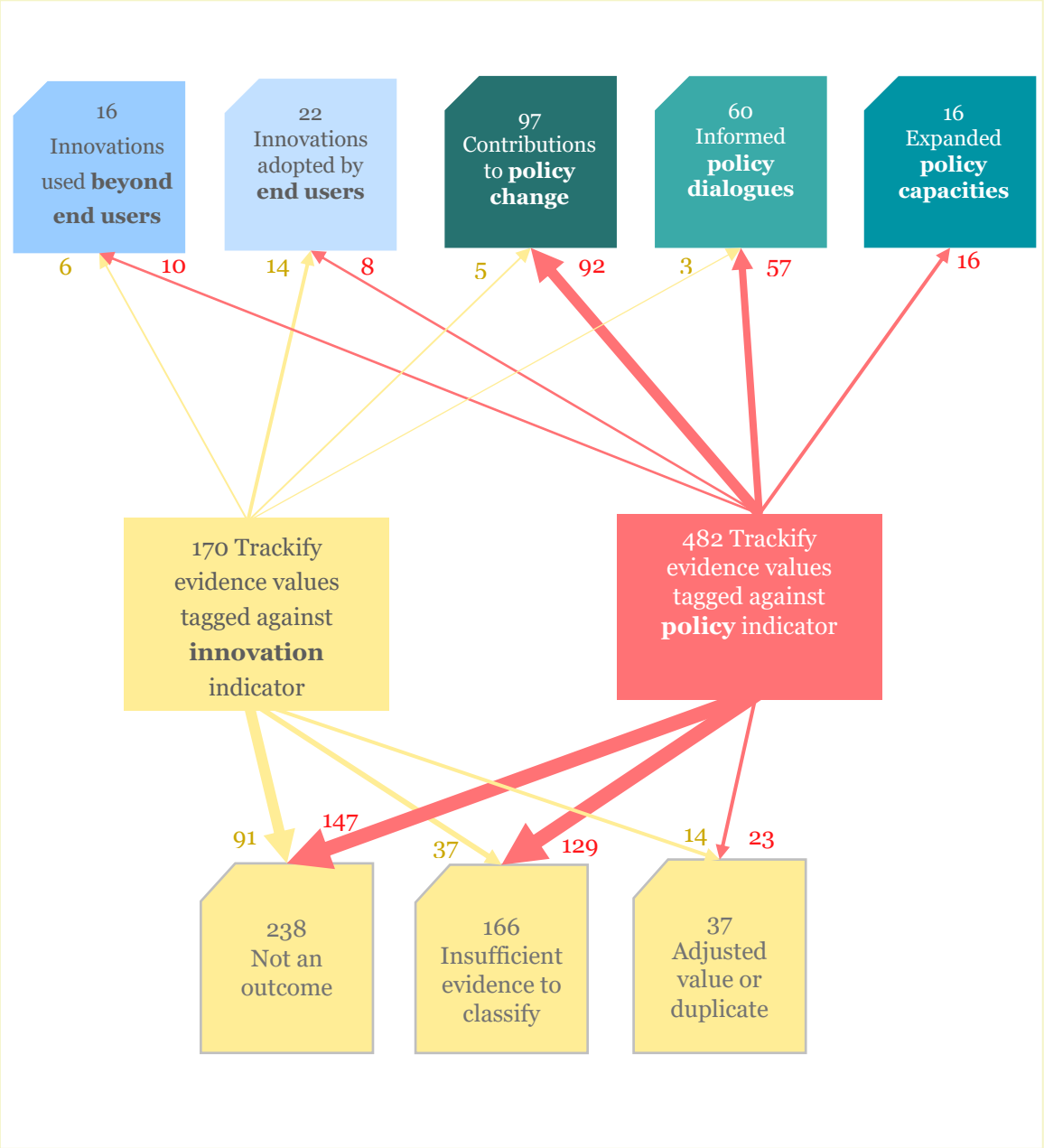

Figure 15: Outcomes classification from two Trackify datasets (arrow thickness approximately corresponds to quantity)

APPENDIX 3: CASE STUDIES OVERVIEW

We wrote four case studies as part of an evaluation of the International Development Research Centre (IDRC) strategy to scale research results, conducted by OTT Consulting in partnership with Southern Hemisphere. Each case study explores a different thematic area related to scaling at IDRC:

- Field building for scale
- Programming for scale
- Private sector engagement and scaling
- Organisational learning about scaling.

They draw on semi-structured interviews with IDRC staff, grantees and other donors of research, document and literature reviews. Case studies should be read in conjunction with the main evaluation report.

Case study 1: Field building for scaling impact

Establishing an important relationship between field building and scaling, this case study proposes that the integrating IDRC's guiding principles for scaling impact into field building efforts can build a strong foundation for scaling impact. After unpacking the relationship between field building and scaling impact, the case study looks at results achieved by integrating scaling into field building efforts and how field building has supported implementation of IDRC's strategic objective to scale.

Case study 2: Programming for scale

This case study looks at what IDRC has done to implement the strategic objective to scale at the program level. The main findings and discussion focus on program design, management and coordination, and adaption of systems and processes for scaling impact. It then looks at results achieved so far by these efforts, identifying 'best cases' that demonstrate good practice and results for scaling impact that could form a basis for learning about programming for scale.

Case study 3: Private sector engagement and scaling

This case study explores IDRC's engagement with the private sector 'to advance ideas and innovation through to large-scale implementation'. It first looks at how IDRC staff understand private sector engagement and strategies programs employed to engage with the private sector related to scaling impact. It identifies benefits and challenges of private sector engagement at IDRC and draws out lessons learned.

Case study 4: Organisational learning about scaling

This case study looks at how and how well IDRC staff and teams learned from their efforts to scale impact within and across programs between 2015–2020. It explores how learning was understood by staff, where learning about scaling impact took place, how scaling impact was understood by staff, how the organisational context affected learning, who the key actors who influenced learning were and the quality of learning that took place.

www.onthinktanks.org