COSA
Committee On Sustainability Assessment

Issue Brief 01

February 2017

Sustainability research that counts

Lessons from implementation research for agriculture

Implementation research is the scientific study of methods to promote the systematic uptake of research findings into policy and practice.[1] Largely applied in the field of health, it has found broader applicability in other domains that use evidence-based research for policy. With the proliferation of rigorous evaluation methods and increasing demands for interventions that demonstrate evidence of what works, there is now a much-needed shift in focus to translate research evidence into policy and practice. For the research for development (R4D) community, this shifting perspective has meant that research quality is judged not only on technical merit but increasingly by demonstrating how much positive impact it has on intended beneficiaries. How can implementation research methods be used to bridge the gap between evidence and practice in the field of agriculture and sustainability?

In this issue brief, we unpack some of the elements critical to the idea of implementation research and highlight some of the unique challenges that occur in thinking about evidence-based policy design for agriculture and sustainability. We proceed by presenting a brief synthesis of the implications and recommendations for evidence-based research for development in the agricultural sector (AR4D) and for the role of researchers.

Gayatri Ramnath
Regional Research Coordinator for Asia

Keith Child
Senior Research Advisor

1 https://www.saxinstitute.org.au/our-work/aci-network-research/network-effectiveness-project/

Unpacking evidence-based research design in AR4D

Traditionally, the perspective adopted by much of the international development community has been what Wigboldus and Leeuwis (2013) call the "push" approach – finding solutions to problems using experiments or pilots and then taking to scale those that are found to work. We examine some of the critical components in this implementation research process chain.

1 Sound evidence: The search for "what works"

In the last two decades, there has been a push to use rigorous evaluation techniques to demonstrate evidence of what interventions work in international development. Several evidence-for-policy research guidelines apply quality and quantity standards to identify what qualifies as strong evidence. Usually aimed at practitioners looking for proof of what social programs work, these guidelines recommend the following approaches: 1. employ evidence from high-quality studies that use valid comparison groups[2] and 2. using studies that have been replicated and show evidence of success in more than one type of setting.

If programs do not meet these requirements, they are not recommended for adoption or scale. The process of looking for this kind of evidence usually involves a meta-analysis of impact assessments (IA) that have evaluated a particular intervention rigorously in multiple settings.

The meta-analysis conducted by The Independent Evaluation Group (IEG) of the World Bank (2011) highlights some of the challenges that make it difficult to derive definitive conclusions about what interventions work in the field of agriculture. Though this study included only those impact evaluations focusing on productivity increases, the authors observe that IA in the agricultural sector with valid counterfactuals – the primary requirement for strong evidence – are scarce.

Of the 271 impact evaluations identified in the literature, only 83 met the inclusion criteria of having a defined intervention and a valid counterfactual. Moreover, the use of randomized control trials – considered the "gold standard" for evidence in some fields of study – was found to be exceptionally rare with only five of the 83 evaluations using this method.

The meta-analysis also found the nature of interventions to be highly heterogeneous thereby limiting the proper use of a meta-analysis as a tool for identifying sound evidence.[3] In addition to the different types of interventions, the IEG meta-analysis found that not all interventions

2 Some guidelines strictly recommend randomized control trials (RCTs) while others recommend using quasi-experimental studies with valid control groups.

3 The eight categories identified are Land Tenancy and Titling, Extension Services, Irrigation, Natural Resources Management, Input technology, Marketing arrangements, Microfinance, and Miscellaneous.

studied the same policy change or used comparable variables for analysis.

Apart from the heterogeneity, specific methodological issues such as self-selection, small sample sizes, and spillover effects are also found to be more pronounced in agricultural evaluations, offering limited consistency when identifying common impacts of interventions (IDB 2010).

Therefore, in AR4D: Sound evidence defined as impact assessments with valid counterfactuals is scarce.

Heterogeneity of interventions and implementation designs make it difficult to aggregate results and draw broad conclusions about what works.

2 Scale

The ultimate purpose of AR4D is to ensure solutions are implemented effectively and made available to a large number of beneficiaries. Though much of implementation research is concerned with the science of taking what works to scale, the idea of scale itself has been interpreted narrowly to mean replication of programs or components of programs. In the following section, we consider nuances of scaling.

Scaling up versus scaling out

Scaling up and scaling out are distinctly different processes with much of the R4D literature focusing on the latter.

Scaling up involves understanding the context within which a program is implemented and processes to broaden its reach. The purpose is to ensure effective uptake at various levels and by different actors.

Scaling out involves replicating programs in other contexts including an understanding of the contextual factors that enabled the program's success at the "source."

It is in the context of scaling out that much of the evidence-based literature in agriculture seems unhelpful in determining what works. Even when an intervention has proven technical merit or is scalable in one context, it is still intertwined with highly context-specific sociocultural, geographic, and other factors that might limit uptake and scalability elsewhere. For AR4D research, it is essential to provide a rigorous assessment of what these factors are and how they contributed to the scaling process.

Public versus private goods

Another distinction to be made is the context within which scaling can occur. One is in the public good sense, meaning the primary intent of generating evidence-for-policy is to produce knowledge that can be used for welfare in a broad range of contexts and for the greater, public good. (Duflo, 2004; Deaton, 2009). In this case, the ownership of the evaluation and design rests principally with agencies like NGOs, donors, and research institutions. The motivation of these actors is the public dissemination of evidence on a large scale in order

The process of looking for this kind of evidence usually involves a meta-analysis of impact assessments (IA) that have evaluated a particular intervention rigorously in multiple settings.

to influence policymakers to take up solutions. The research design, theories of change, and evaluation methods reflect this approach.

This can be contrasted with evidence generated primarily for private consumption - i.e., understanding what works but not necessarily with the explicit purpose of producing "international public goods." Rather, the focus is on the client- or customer-centric objectives that are aligned with strategic business or supply chain interests. Here, uptake of research is aimed at key decision-makers in the private sector or within the agencies that commission these studies.

Dimensions of Scaling

Lastly, when discussing scale, it is important to note that scaling can occur in contexts with varying degrees of complexity. Much of the R4D literature focuses on the "push" approach (Wigboldus, Lewis 2013), taking an intervention that has demonstrated positive impact in the pilot stage to scale up and out, in the sense of a technology transfer. This view is problematic in agricultural implementations for two related reasons: the contexts in which scaling occur are complex, and this complexity undermines the predictability and degree to which the scaling process can be controlled or directed. Wigboldus and Leeuwis (2013) discuss four distinct dimensions in which scaling can occur based on the degree of social and technical complexity (Table 1).

Table 1

Scaling Context	Social Complexity (Stakeholder disagreement)	Technical uncertainty	Type of Scaling
Simple	Low	Low	Push: We have something that we would like to go to scale, and we will work hard to make that happen.
Technically Complicated	Low	High	Pull: We have an aspired future in mind and seek to scale up and out that which we think will help make that future reality.
Socially Complicated	High	Low	Plant: We have something we would like to go to scale, but such scaling can only happen if we connect other factors and work with other (development) actors.
Wicked Problems	High	High	Probe: We have an aspired future in mind but are unsure about what scaling processes would be involved in moving toward that future so we will have to navigate and adapt as we go.

Each of these dimensions relates to different levels of complexity and uncertainty in a system and translating research into policy action will require adopting different strategies in different scenarios. When thinking about evaluation, it is important for researchers to develop a rigorous analysis of the contextual factors affecting scale and implementation at the outset to increase the usability of evidence from research. In the next section, we consider examples that illustrate how research organizations can position research for better use.

3 Positioning research for use: framework and tools

The positioning of research and evidence for use is complementary to the process of scaling and can lead to the effective uptake and implementation of interventions. Indeed this is what development donors mean when they apply collective pressure on development agencies and the research community to prove that they can bring about positive change for intended beneficiaries. This implies that traditional criteria for judging research merit such as - citations, peer reviews, and other bibliometric approaches are not measuring up as useful measures of research effectiveness.

This perspective is not new. For instance, in 2005 (Fixsen et. al), a meta-analysis of implementation literature found that two of the methods most commonly used by researchers--information dissemination through publishing and 2. training, were ineffective implementation methods in several fields of study. The issue, in large part, is that the criteria for judging research effectiveness/quality has been the preserve of scientists who tend to judge research quality according to scientific values like internal and external validity, research design and implementation, and replicability rather than on research use, uptake, and impact. Within the scientific community, these latter criteria are seen largely as “somebody else’s problem.” The message from the donor community, on the other hand, is clear: science and scientific values “can no longer be considered a mostly academic enterprise divorced from societal concerns about social goals” (IDRC, 2016).

So how should research be positioned so that it improves research use, uptake and ultimately, impact? We consider two examples that reconcile conflicting perspectives. First we present the IDRC RQ+ framework that offers clear criteria to assess the quality of research conducted by institutions; and second, we present Lean Research (LR) as a method to put these assessment criteria into practice.

The RQ+ Approach

The Canadian International Development Research Centre (IDRC) in its Research Quality Plus (RQ+) approach describes positioning research for use as “the extent to which the research process has been managed, and

research products/outputs prepared in such a way that the probability of use, influence and impact is enhanced." The RQ+ focuses on:

1. Knowledge accessibility and sharing: the extent to which research products are targeted to specific users, to what degree they are conveyed in a manner that is intelligible to intended beneficiaries, and whether they are appropriate for the socio-economic conditions of their context.

2. Timeliness and actionability: focus on the intended user setting at a particular time and the extent to which researchers have internalized this in their planning. For example, if the research product is an econometric model for conducting ex-ante impact assessments, does the Ministry of Agriculture (or staff within the intended Ministry) have the technical capacity to run the model and make sense of it? If not, part of the technical innovation should involve targeted capacity building for key ministry staff. This approach is significant because it attempts to measure the extent to which research has been positioned to increase the probability of its use. First, it acknowledges that scientific merit alone is no longer sufficient for judging research quality. Second, and most importantly, it concedes that research effectiveness through scaling up and out is almost always beyond the sphere of control of research staff. Instead, RQ+ focuses on what researchers can be held accountable for and sets clear and transparent criteria for evaluating research quality in light of the new global concern with value for money in research investments.

Lean Research Guide for Smallholder Farmers

A wide range of social research methods are available for researchers to put the principles of implementation into practice. While the RQ+ framework focuses on the users of research products, the Lean Research Initiative (LRI) and the Sustainable Food Lab have developed guidelines for work with smallholder farmers that place research subjects and stakeholders as core beneficiaries of the research process. The LR principles of relevance to stakeholders, respect for human subjects, research rigor, and right-size apply to all the stages of the research process and seek to "produce results that are meaningful not only to the researchers but to research participants as well".

From a scale and implementation perspective, the guidelines emphasize managing stakeholder relationships up front so research subjects and community members can participate in choosing research topics that are relevant to them. The guidelines also focus on how the results of the research should be presented to the community and stakeholders so that they can engage effectively in the decision-making process.

Using approaches like the RQ+ and guidelines from the Lean Research Initiative can help put some of the principles of implementation research into practice.

Scaling out involves replicating programs in other contexts including an understanding of the contextual factors that enabled the program's success at the 'source.'

Conclusion

Evidence-based research in agriculture is complex, heterogeneity is the norm, and implementation is affected by many context-specific variables. For the AR4D community, applying the principles of implementation research to enhance the usability of research necessitates a rethinking of several aspects of research – from what constitutes sound evidence, to the different states of complexity in which scaling can occur and different strategies that need to be used to increase the uptake of solutions.

Central to this discussion is how evidence from research should be positioned to influence key stakeholders. Effective positioning of research findings leads to effective implementation. Apart from ensuring that the research design meets all technical merit criteria, positioning the research effectively for use is the responsibility of the R4D community. Approaches like the RQ+ framework and the Lean Research Initiative for smallholders provide practical guidelines for researchers to enhance research usability by ensuring the research design incorporates a rigorous stakeholder engagement and participatory process.

Bibliography

Duflo, Esther, 2004, **"Scaling Up and Evaluation,"** Annual World Bank Conference on Development Economics 2004, Washington, DC. The World Bank.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., and Wallace, F. 2005. **Implementation Research: A Synthesis of the Literature.** Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Impact Evaluations in Agriculture: An Assessment of the Evidence. 2011. Independent Evaluation Group, World Bank https://ieg.worldbankgroup.org/Data/reports/ie_for_ag.pdf.

Wigboldus, Seerp and Leeuwis, Cees. 2013. **Towards Responsible Scaling In and Out Agricultural Development: An exploration of concepts and principles.** Center for Development Innovation and Knowledge, Technology and Innovation Group, Wageningen.

Robert; Ofir, Z., T. Schwandt, C. Duggan, and R. McLean; McLean, **"RQ+ A Holistic Approach to Evaluating Research,"** 2016, 28, https://www.idrc.ca/sites/default/files/sp/Documents%20EN/research-quality-plus-a-holistic approach-to-evaluating-research.pdf.

Resources

https://www.idrc.ca/en/article/new-evaluation-tool-now-available-assess-research-quality

Winters, Paul., Salazar, L., Maffioli, A. 2010. **Designing Impact Evaluations for Agricultural Projects. Impact Evalaution Guidelines.** Technical Notes No. IDB-TN-198. https://publications.iadb.org/bitstream/handle/11319/1956/Designing%20Impact%20Evaluations%20for%20Agricultural%20Projects.pdf?sequence=1

COSA
Committee On Sustainability Assessment
www.thecosa.org

COSA is a non-profit and independent, global consortium of partners dedicated to accelerating agricultural sustainability through robust information systems that is financed in part by leading agencies including the Swiss Government (SECO), Ford Foundation, and InterAmerican Development Bank.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Federal Department of Economic Affairs FDEA
State Secretariat for Economic Affairs SECO

FORDFOUNDATION

IDB MIN