



Standing  
Panel on  
Impact  
Assessment

2020

## **SPIA Approach to Impact Assessment for CGIAR**

SPIA TECHNICAL NOTE N. 8

STANDING PANEL ON IMPACT ASSESSMENT (SPIA)

## 1. Precedents

The CGIAR portfolio spans a wide spectrum of research, development, and policy engagement activities, involving a multitude of disciplines and approaches, and a large set of causal pathways, and possible outcomes. The commonality across this wide portfolio is that CGIAR aims to provide public goods, which potentially benefit many populations in different parts of the world. It is the global public good nature of CGIAR activities that provides the theoretical rationale for support of the system. The theoretical foundation for CGIAR is unassailable, and moreover, the results of the Green Revolution have provided very strong historical support for investment in a system of international agricultural research centers (e.g., Evenson & Gollin, 2003; Gollin, Hansen & Wingender, 2018; Fishman et al., 2017). Yet evidence for the large-scale impact of more recent investments in CGIAR is scarcer and maintaining confidence in the system requires an updated approach to impact evidence.

SPIA's mandate—providing evidence on the impacts at the system level of a set of agricultural research-for-development (AR4D) centers and programs that are themselves portfolios of research and development programs and projects—is possibly unique in international development. This challenge is complex and distinct from the evaluation of an individual research or development intervention. There are, however, clear parallels to methods for establishing impact of a portfolio of investments in development innovations (Kremer et al., 2019).<sup>1</sup> With CGIAR innovations aiming to improve outcomes across different impact areas, and given the new standards for rigor in impact assessment recognized by the 2019 economics Nobel prize, establishing evidence of impact at the CGIAR system level in the 21st century requires a new approach (Stevenson, Macours, & Gollin 2018). It implies stepping away from traditional aggregate rate-of-return calculations, which typically rely on aggregating the result of a large set of studies with no regard for whether the evidence can credibly be considered as causal or indeed relevant for real-world conditions (Stevenson, Johnson & Macours 2018).

## 2. SPIA's Current Approach

SPIA's approach specifically recognizes that CGIAR's core activity is AR4D, and therefore faces the uncertainties inherent in scientific progress, in addition to those associated with development investments and processes. All good research contributes to knowledge, but only some of that knowledge leads to insights or innovations that ultimately can be scaled and contribute to real-world impacts. Also, while AR4D relies on specific theories of change (ToC) to link research to impacts at scale, the ToC necessarily make many assumptions along the long pathways to impact. Even so, at the system level, evidence is needed on whether, taken together, the benefits of the different activities of the system exceed the investment costs in the system.<sup>2</sup>

---

<sup>1</sup> For a presentation of the approach developed by Kremer (one of the 2019 Nobel laureates) and co-authors for USAID's Development Innovation Ventures (DIV) see <https://www.youtube.com/watch?v=95tINyMc7dc>. There are also parallels with the approach used by the Global Innovation Fund. <https://globalinnovation.fund/>

<sup>2</sup> And possibly, whether they do so to an extent similar or higher than alternative development investments (in education, health, infrastructure, etc).

To provide evidence of the impact of investments in the system, we start from the fact that the returns to research investments are likely very skewed: i.e., even if many research investments may not translate into measurable impacts at scale, a few innovations can have very high returns. Hence, obtaining reliable estimates of the impacts of possible big win innovations, together with reliable estimates of the scale and duration of such impacts, is sufficient (but not necessary) to justify investment in the system as a whole. It is, however, inherently backward-looking. Innovations from research not only take time to be widely taken up, they often need to be in use for a long time so that both the direct and indirect benefits are fully realized—not only for the adopting households or communities but for the economy as a whole. As a result, the most rigorous and reliable estimates of past impacts were obtained in economic, social and ecological conditions that are very different from the world we face today, let alone 20 years into the future when today’s research will be generating impacts. Accountability-oriented studies can hence provide important lessons about how development occurred, but the extent to which those lessons can be applied to the current choices facing CGIAR and partners may be limited. This is further true because AR4D is not a linear process and learning from smaller wins (or even failures) can be instrumental in achieving the big wins.

Therefore, to inform current decision making, a second category of SPIA impact studies seeks to validate key aspects of the ToC of recent research innovations that are in the process of going to scale. Evidence from impact studies can help refine and validate the ToC, by providing evidence on factors influencing uptake and use. Learning studies provide further feedback into the ToC by quantifying short-term benefits and trade-offs that derive, directly or indirectly, from the innovations. Carefully chosen to complement studies conducted by centers and CGIAR Research Programs (CRPs), these studies aim to fill important evidence gaps in the implicit system-level ToC. The results can feed back into research program design, as well as to the design of cost-effective dissemination and scaling programs. As such, learning from these (often smaller scale) studies can be critically important for getting to big wins.

As these different objectives require different empirical approaches, and since they logically have different implications for decision making processes in the CGIAR, we discuss them in turn.

## **A) Accountability and Scale**

While it is difficult to predict big wins at the moment of investment, they can be sufficient to justify (ex-post) the entire investment in the system, potentially over many years.<sup>3</sup> The simple criterion is that the total cost of investment is lower than the benefits of the few big successes. Only quantifying the benefits of a few innovations will provide a conservative (lower bound) estimate under relatively weak assumptions.<sup>4</sup> As such it explicitly acknowledges that there are likely also many smaller wins and wins for which benefits can be harder to quantify.

---

<sup>3</sup> Studies may also (of course) establish that reach or impacts of certain innovations that were a priori considered big wins are less (or different) than expected, an equally valuable finding.

<sup>4</sup> The approach hence does not rely on extrapolating the gains from big wins to other innovations, but rather it only considers the gains from the big wins, hence assuming the gains from other innovations are zero. Or to be more precise, it assumes any possible negative net externalities of other innovations do not exceed their value to beneficiaries on average.

Importantly, this type of evidence provides a crucial accountability role, but does not aim—and hence should not be used—to provide guidelines regarding prioritization of new research investments.<sup>5</sup>

Documenting big wins requires causal evidence of returns to innovations at scale. As returns are determined, arithmetically, by the number of beneficiaries times the benefit per beneficiary, SPIA hence focuses on:

- 1) Documenting reach: objective evidence that large numbers of people—farmers, consumers, communities—are using the innovations and/or that their use is affecting large areas (e.g., # hectares, # of districts). For most innovations, widespread diffusion is a necessary condition for impact at scale (even if it is not a sufficient condition) and the ability to reach large numbers is an important part of the case for CGIAR. To systematically compile data on adoption and diffusion of CGIAR innovations at policy-relevant scales, SPIA leverages partnerships with national statistical agencies, the [World Bank’s Living Standards Measurement Study](#)<sup>6</sup> team, and the [50x2030 Initiative](#).<sup>7</sup>
- 2) Rigorous estimates of the benefit per beneficiary of innovations that have scaled, which requires credible counterfactuals (i.e., what outcomes would have been in the absence of the diffusion of the innovation). Given the spatial and temporal scales involved in studying big wins, such counterfactuals usually involve quasi-experimental approaches. To be credible and robust to independent scrutiny, the assumptions that necessarily underlie such quasi-experimental approaches need to be supported by:
  - a. Documentation of how diffusion occurred: The data and information needed to show how innovations have scaled, needed to construct a credible counterfactual, are not necessarily readily available, in part because dissemination of successful innovations can happen through a variety of mechanisms (e.g., through national partners, markets, private sector initiatives) and is often not well-documented. Targeted investments in data collection and Monitoring and Evaluation data of large scaling efforts can be used to document the geospatial and time variation of diffusion. At the individual farmer (or community, or consumer) level, adoption is a choice. Therefore, directly comparing adopters and non-adopters does not allow the estimation of impacts (see Janvry et al., 2011, for a detailed discussion). Instead, the variation in the availability of an innovation to potential adopters, when it is well-understood, can provide a starting point for constructing rigorous empirical strategies for documenting impacts at scale.
  - b. Measuring relevant outcome variables at the appropriate geographical and time scales. Advances in remote sensing and increased availability of administrative data can be combined with leveraging information from large representative data collection efforts

---

<sup>5</sup> Indeed, prioritization is necessarily a forward-looking, not a backward-looking exercise.

<sup>6</sup> <http://surveys.worldbank.org/lsm/>

<sup>7</sup> <http://www.data4sdgs.org/50by2030>

and targeted primary data collection, to measure social, environmental, and economic outcomes, and document possible trade-offs and synergies.

## **B) Learning Studies and Decisions to Scale**

Evidence of historic large-scale impacts offers little useful guidance for current or future investments. Consequently, it must be complemented with studies that aim to provide more direct feedback into decision making in research and scaling systems. These include impact assessments aimed at testing specific steps in the causal pathways of particular innovations to inform decisions as part of a stage-gating process. Where possible, such studies should focus on steps that are relevant to the uptake and impact of multiple innovations, so that broader lessons can be drawn. Designed based on the ToC for specific innovations, these learning-oriented studies aim to provide rigorous empirical evidence of causal mechanisms in the chain between research investments and impacts. They can also be specifically designed to measure trade-offs and synergies between different types of outcomes. CGIAR research programs and external researchers often conduct such learning-oriented impact studies<sup>8</sup>. SPIA's contribution is focused on a) methodological improvements and b) coordinated sets of studies focusing on similar evidence gaps in different contexts to enhance the generalizability of findings, making it possible to draw broader (possibly CGIAR-wide) lessons regarding outcomes and causal pathways.

Such impact assessments are ideally conducted on promising innovations before they get scaled, and indeed, the insights from these studies can be instrumental to help innovations scale and become big wins. They are often designed prospectively, with the purposeful construction of a counterfactual, when possible through a randomized control trial. They can be specifically designed to fill evidence gaps related to the adoptability of innovations by target users and the size and distribution of impacts of that use on beneficiaries. By rigorously testing whether the key causal mechanisms that are expected to drive adoption or impact are working as expected (and also where and for whom) they provide early evidence of benefits and can help identify key obstacles for scaling of the innovations. As such they allow learning lessons that can feedback to research and forward to dissemination efforts. As for the accountability studies, finding evidence of lack of impacts can be equally valuable feedback into the research and scaling efforts.

## **3. Rigorous Methods, Timelines, and the Importance of Partnerships**

Accountability and learning impact studies are strongly complementary to each other as they test for evidence of impacts at different moments in the research cycle. Both types of studies require the use of empirical methods of the highest standards to provide trustworthy and useful evidence. For this reason, improving the methodological rigor of individual impact studies is an important part of the SPIA mandate. This starts from valid and accurate measurement and extends to credible causal inference and appropriate methods for measuring impacts at scale.

---

<sup>8</sup> <https://storage.googleapis.com/cgiarorg/2019/11/e6acb6fc-presentation-spia-chair.pdf>

The rapid pace of change in impact assessment and measurement methods and tools means that impact assessment in CGIAR needs to stay up to date. To achieve methodological rigor, objectivity, and relevance, SPIA encourages and facilitates partnerships between external impact assessment specialists and CGIAR natural and social scientists.

Generating such rigorous impact evidence is a multi-year process, as it typically requires collecting multiple rounds of data, careful research design planning, and preparatory scoping work. This is even more so when aiming to provide evidence through a balanced portfolio of studies.

#### **4. The Use of Different Types of Impact Evidence for Decision Making in CGIAR**

Not only is it challenging to generate credible impact assessment evidence, but it is also challenging to use that evidence to understand the impacts of past AR4D investments and make informed decisions about future investments. There are many claims made and studies published, and it is not always easy to determine credibility or relevance. SPIA does not have a mandate to manage or oversee the large amount of impact assessment research that takes place within centers and CRPs, or to appraise the evidence quoted in support of a funding proposal, and it does not seek that mandate. However, by raising awareness of the issues involved in assessing research impact, by providing advice about quality standards (for users of evidence), and research designs (for researchers of prospective new studies), SPIA can contribute to improving the generation and use of high-quality impact evidence across CGIAR.

#### **References**

Evenson, R. E., & Gollin, D. (2003a). Assessing the impact of the Green Revolution, 1960 to 2000. *Science*, 300(5620):758– 762. <https://science.sciencemag.org/content/300/5620/758>

Fishman, R., Dar, A., Barnwal, P., Mueller, N. D., von der Goltz, J., and McCord, G. C. (2017). Crop variety improvement and infant mortality in the developing world. Mimeo.

Gollin, D., Hansen, C. W., & Wingender, A. (2018). *Two blades of grass: the impact of the green revolution*. NBER working paper 24744. <https://www.nber.org/papers/w24744>

de Janvry, A., Dunstan, A., and Sadoulet, E. 2011. Recent Advances in Impact Analysis Methods for Ex-post Impact Assessments of Agricultural Technology: Options for the CGIAR. Report prepared for the workshop: Increasing the rigor of ex-post impact assessment of agricultural research: A discussion on estimating treatment effects, organized by the CGIAR Standing Panel on Impact Assessment (SPIA), 2 October 2010, Berkeley, California, USA. Independent Science and Partnership Council Secretariat: Rome, Italy. <https://cas.cgiar.org/sites/default/files/docs/deJanvryetal2011.pdf>

Kremer, M, S Gallant, O Rostapshova, and M Thomas, 2019 “Is Development Innovation a Good Investment? Which Innovations Scale”. Harvard University Working Paper. [https://scholar.harvard.edu/files/kremer/files/sror\\_div\\_19.12.13.pdf](https://scholar.harvard.edu/files/kremer/files/sror_div_19.12.13.pdf)

Stevenson, J., Johnson, N. and Macours, K., 2018, Estimating ex post Impacts and Rates of Return to International Agricultural Research for Development, SPIA Technical Note 6

<https://cas.cgiar.org/spia/publications/estimating-ex-post-impacts-and-rates-return-international-agricultural-research>

Stevenson, J., Macours, K., & Gollin, D. 2018. *The Rigor Revolution in Impact Assessment: Implications for CGIAR*. Rome: Independent Science and Partnership Council (ISPC).

[https://cas.cgiar.org/sites/default/files/pdf/ispc\\_synthesis\\_study\\_rigor\\_revolution.cgiar.pdf](https://cas.cgiar.org/sites/default/files/pdf/ispc_synthesis_study_rigor_revolution.cgiar.pdf)

---

For more information, contact:

Karen Macours, Chair, SPIA

[spia@cgiar.org](mailto:spia@cgiar.org)