



## Notes from: Reflection on Methods and Metrics Needed to Support Impact Assessment Plans in One CGIAR

15 September

### Opening presentation

**Intervention:** Worry that impact assessment might not be seen as research.

**Intervention:** Regarding accountability: for some innovations that are ready perhaps it is possible to do accountability studies within 3 years

**Intervention:** Reflecting on the experience in the past two weeks about integrating impact assessment strategies into the initiative proposals. Another issue is the different nature of proposals in the different science groups. This affects ability of the teams to put together rational plans into the fixed template. Trying to come up with rigorous impact assessment plans in a fixed template is challenging for initiatives covering an enormous range of activities.

**SPIA member response:** On collecting good M&E data and good data on rollout, so that a long-term large-scale study can be done. Would like a clarification. Should we think of that as a preparation for the future, but without being a substitute for rigorous learning studies that give results within three years? Is the message that you need both? Another clarification that is needed: Are short-term studies necessarily RCTs? Or are we also open to quasi-experimental approaches?

**Intervention:** Looking at RCTs vs other types of methodology. I suggest also piloting some of the work and then getting into a more detailed RCT in the next phase of the project. How about cost-benefit analysis and estimating value for money? Is it considered impact assessment or where does it fall? It is not as rigorous as impact assessment, but it has value for donors.

**SPIA member response:** On different types of studies. The method you use depends on the research question you are trying to answer. For the initiatives, the research questions should link to the ToC and should test the assumptions in the ToCs. Some of them are more suited to test with RCTs, others require a combination of methods. Some need qualitative methods. It will often make sense to first pilot things at a small scale before designing a large scale RCT. Start thinking about the important questions to answer and work from the start to prepare so that you can answer those questions with the appropriate method. Don't leave these questions for the end of the process when it may be too late to collect some of the data needed. Different initiatives will have different types of questions that are relevant for them. This means that different methods will be appropriate for them. SPIA did not develop the template, but our advice would be to prioritize scientific coherence over fitting perfectly into the template.

Regarding rates of return: Sometimes for a specific innovation we may be interested in calculating, on the farmer level, the cost of using the technology and comparing it to the benefit of the technology. This can be a relevant learning question to ask in order to

understand why certain farmers are adopting and others not. You then need rigorous methods to get to benefits and rigorous methods to get to the cost.

When thinking of the system as a whole, then need to depart from the fact that science is an activity with high degree of uncertainty. Out of all research done only a few projects will lead to big wins. Estimating the benefits of those big wins can justify investment in the system as a whole. This is the idea for the initiatives. Let's not pretend you can estimate rates of return to each activity. Better to document rollout of potential big wins.

On 3 years vs 10 years point, in the PRMF it is clear that at the initiative level it is about studies that test the assumptions underlying the ToC. There will be initiatives that are scaling innovations already, but getting impact estimates on SDG outcomes within a 3-year period is very ambitious. So no need to pretend it is possible. Need a lot of statistical power to get estimates of effects on outcomes such as infant mortality or poverty. An underpowered study is a bad investment. So better to make sure that we have the building blocks.

**Intervention:** Many initiatives have 3 streams of research: 1) Technical research; 2) Piloting->farmer engagement-> marketing activities (typically where we locate all our quantitative impact assessment); 3) Stream with further scaling: engaging partners and getting institutions involved and evaluated business models.

All this is supposed to happen within 3 years. ToC is usually built in a linear way, but we work in parallel. I can imagine links between piloting and scaling (when have initial results from productivity change, they enthuse partners and we can show how things work), but harder to link with the technical research. When we start with the piloting we decide on packages. We have information from technical partners from previous studies, but then when we go through piloting and engage the partners it is much harder to take on new outputs from the technical research and make changes. The whole idea with impact assessment won't work if we change the baseline. How can we be more consistent in this parallel nature of many of these initiatives?

**SPIA member intervention:** A common challenge for impact assessment teams is how to handle country level interventions, lobbying interventions etc. How can SPIA best support these cases where other types of methods are needed? SPIA now has qualitative capacity and can help with matchmaking. Is that something that would be valuable?

**Intervention:** Regarding what role SPIA should play – the time horizon is inhibiting the likelihood that there is integrated impact assessment.

On methodologies, there is a value of pushing RCTs, since it is likely other studies (quasi-experimental or observational) will happen anyway. For some initiatives it would be valuable if SPIA provided guidance on how an RCT can be carried out in a short timeline (what types of interventions and data collection instruments that lend themselves to this type of set-up). Does SPIA have recommendations of how to fund an endline despite the 3-year time horizon?

**Intervention:** What will initiatives do and what will CGIAR do? Should focus of the initiatives be on the learning studies and then CGIAR takes care of the accountability studies? Does this mean that initiatives shouldn't worry about bigger and longer-term impacts?

**Intervention:** In the new template that was sent out, it is not clear in which category RCTs should be placed. Would be helpful to have guidance there?

Documentation of reach will be important, so what is SPIA's suggestion on how to go about it?

**SPIA member response:** RCTs are typically learning studies (testing short-term impacts), but this doesn't mean they can't be used as building blocks for more long-term studies. There is an example in the SPIA portfolio (livestock insurance), where authors can now look at 10-year impacts building on variation created from the RCT. But we know that for many of the innovations and technologies these types of RCTs don't exist and so it is not always possible to build the long-term impact assessment on these.

However, if we can do the RCT at the moment when we are scaling, we are making things easier for the future. So for the initiatives that do have scaling efforts there is a lot to say for getting an RCT expert involved in talks with scaling partners to see if it is possible to design an RCT.

What is SPIA's role? We can't advise every initiative on exactly what to do, since is outside of both capacity and mandate. However, can try to point you to people with the right expertise inside and outside of the system.

Regarding 3-year time frame. Everybody has an interest in looking at the longer term. If you have a good initial design, it is usually easy to fund an endline.

On what the CGIAR does and what initiatives should do: Guidelines for what design teams should focus on includes learning studies and documenting rollout (what happened where, when, why and with whom). This needs to be documented from the start and happen within the initiative, since that's where this information exists. Requires close collaboration with research teams.

For the part that is about documenting reach, there is economies of scale and more of a logic to have these at the system level. However, documenting reach isn't enough, but needs to be complemented with impact estimates that cannot be designed without information from the initiatives.

**SPIA question:** To those who have been working with IDTs. What type of gaps are there in terms of methods? What type of expertise is needed?

**Intervention:** Wide range of experiences. Different to what degree initiatives are involving plans for causal impact evidence. Some really strong and some just completely ignore causal impact assessment on the ToC.

**SPIA member intervention:** Some initiatives have engaged impact assessment researchers who have an eye for where in the ToC there are opportunities to involve impact assessment. That is critical for successfully building impact assessment into the proposals.

**Intervention:** Some of the work packages are designing the work under quite tight budgets. In that context designing the impact assessment has been seen as secondary.

Regarding question on methodology. Most initiatives are not yet at the stage of designing studies for ToC testing. Even for those who are giving priority to impact assessment plans, they are not yet looking at specific designs. This will have to come at a later stage.

**SPIA member intervention:** Regarding budget: clear timing problem. Most teams don't know what sample size they need or even how many RCTs need to be done, but need to provide budget now. Risk of not having enough money to implement good impact assessment ideas.

**SPIA member intervention:** It is clear that different initiatives will be at different stages. September 30th is not the end of the process. What is it that you think would be helpful from

SPIA when you move into the next stage after Sept 30th? Would matchmaking be appreciated?

Are there any further ideas on how to collectively share experiences and learnings?

**Intervention:** Would be helpful to get messages out for early October that SPIA is there to help for matchmaking or more engagement. Some of the initiatives are lagging behind and would probably appreciate this. There are also a lot of people in the system who are happy to find more work after 2022.

**Intervention:** Is there understanding that there are cost implications for learning studies? Does SPIA have any role in creating awareness among initiative leads?

**Intervention:** If there is no guideline, many initiatives will end up underestimating the budget requirement for impact assessment. Would be helpful to define a minimum.

**Intervention:** We cannot expect that all of the budgets will come from central funds. Bilateral funding can be one solution. In designing and implementing RCT, some initiatives who are not ready to do it may need to identify alternative ways of piloting using qualitative methods or on a smaller scale.

**Intervention:** Have to be careful so that initiatives don't plan to rely fully on bilateral funding.

**Intervention:** Budgeting on MELIA - hoping that monitoring and learning will be enough in the first 3 years. Have to focus on monitoring and learning in the first 3 years and then IA can come later.

**SPIA member intervention:** It is important that monitoring activities are thought through in advance with the lens of an impact assessment design. Documenting the when, where, why and how. So that it is in place later for a study to be done.