**How to evaluate Science, technology, and innovation in a development context? – Discussion on the Beta version of the guidelines**

1. **Do you think the Guidelines respond to the challenges of evaluating the quality of science and research in process and performance evaluations?**

YES

1. **Are four dimensions clear and useful to break down during evaluative inquiry (Research Design, Inputs, Processes, and Outputs)? (see section 3.1)**

YES

1. **Would a designated quality of science (QoS) evaluation criterion capture the essence of research and development (section 3.1)?**

YES. While the QoS evaluation criterion captures the essence of research to a large degree with the available knowledge at a particular point in time, it has not taken care of development aspects (adoption and impact of technologies) adequately.

1. **Do you have experience of using other evaluation criteria to evaluate interventions at the nexus of science, research, innovation, and development?**

YES.

During my service career in the Indian Council of Agricultural Research, the process evaluation criteria are largely taken care of during the project formulation and review stage by experts and Institute Research Council / Research Advisory Council.

The Performance evaluation is taken care of to some extent by Quinquennial Review Teams. But that was not adequate from the policymakers’ and project funding agencies’ point of view, and queries come from them regarding the benefits to society. Therefore, I used to carry out impact assessments of technologies/interventions, under real farm situations in the project area or selected area. This takes into account the increase in income and employment of smallholders besides the extent of adoption of technologies. Towards the last ten years of my career, I used to take the effectiveness/ impact of the research institute at the state and national levels using secondary data which not only shows the extent of adoption at the state/national level but also returns to investments in research at the research institute level.

1. **What are additional data collection and analysis methods that should be elaborated for evaluating science and research in process and performance evaluations? (see textbox 3, figure 8 and tables 5, 6 and 8)**

The steps/criteria listed for data collection are adequate for process evaluations. However, they are inadequate for performance evaluations due to the shorter project period. The full potential of a successful technology or product is generally realized by the farmers and society after the project period is over.

CGIAR deals with complex portfolios and products are subjected to extreme weather events and biotic stresses under real farm situations. Take the case of the plant breeding platform. The development of a crop variety and its release process takes 7-10 years and even more for some crops. The full impact of a variety (in terms of area coverage) takes another 5-10 years in countries where the geographical area is large. Nowadays, the policy of CGIAR is to create more improved intermediate products (through conventional/ biotechnological/ mutation/ hybrid parental lines, etc. methods or a combination of different methods) rather than finished products as they express differently in different environments due to gene and environment interactions. They are supplied to different countries and tested in different environments around the globe in research stations over a period of time. Success and failure stories are there. The prominent products which perform significantly in the environment are only recommended to farmers for cultivation/adoption. The funding agencies are more interested to know how many poor farmers are benefited due to the investments that they have made, which are not available and possible to assess at the end of the project period. It is not possible to get the whole set of impact data during the project period (especially benefits to smallholders) as expected by funding agencies. Therefore, the collection of data on prominent breeding materials supplied to different countries/ research stations and prominent products from such environments/research stations needs to be collected for performance evaluations.

1. **How can CGIAR support the roll-out of the Guidelines with the evaluation community and like-minded organizations?**

More brainstorming sessions with (more) examples/case studies/success stories from different regions of the globe covering CGIAR institutes and NARS under a platform are to be listed to refine performance evaluation metrics. Considering the complexity of the CGIAR portfolios, and time & resource constraints for obtaining impact data, the guidelines may be rolled out. The learnings during the evaluation process will through more insights to refine the methodologies in future.