

REPORT FROM THE PRESIDENTIAL STRAND SESSION - 2016 AMERICAN EVALUATION ASSOCIATION CONFERENCE: “Enhancing Collaboration – Design and Evaluation for Innovation at Scale”

Atlanta, Georgia - October 28, 2016

Context: At the 2016 conference of the American Evaluation Association (AEA), Canada’s International Development Research Centre (IDRC) was pleased to host a Presidential Strand session on the Topic of “Impact at Scale.”

Purpose: To further an open and frank discussion of the opportunity and challenge created by “scaling” for both designers and evaluators, and to enhance collaboration between organizations in better understanding this topic.

AEA presented an exciting opportunity for collegial debate and discussion of this issue. The session followed a [panel discussion](#) the previous day with leaders from IDRC, the UK Department for International Development (DFID), USAID’s Global Development Lab, The Rockefeller Foundation, and both Southern and Northern evaluation communities. Panelists spoke to recent work at their respective institutions on the subject, followed by discussion with attendees and each other.

The speakers on this panel were:

Panel Moderator:

- Nancy MacPherson [Managing Director at The Rockefeller Foundation]

Presenters:

- Robert McLean [Senior Program Specialist, Evaluation - International Development Research Centre]
- Penny Hawkins [Head of Evaluation - DFID]
- Joshua Kaufman [Director, Office of Evaluation and Impact Assessment - U.S. Global Development Lab, USAID]

Discussants:

- Mallika R. Samaranayake [President - Community of Evaluators (CoE) South Asia]
- F. Harry Cummings [President - Canadian Evaluation Society]

The special meeting this report summarizes was intended to further discussion stimulated by the panel.

Session Brief: “Enhancing Collaboration – Design + Evaluation for Innovation at Scale”

AEA President John Gargani opened the meeting. John highlighted the importance of the issue of scale to program design and evaluation. He also cautioned evaluators and designers to take a more critical view of the development of the field in regard to scale. He suggested that evaluators and designers hold great power in setting, and re-setting, the future agenda on this topic. He encouraged those in the room to be active in this role.

Discussion Point I – How does scale affect program design? What needs to be considered in designing programs that aim to scale innovations or policy? What do we need to know about context and environmental factors?

Discussion leads:

- Shawna Hoffman, Monitoring and Evaluation Specialist, The Rockefeller Foundation
- Marta Arranz, Senior Advisor Planning, Monitoring, Evaluation and Learning (PMEL) – Influencing / Programme Quality Team, Oxfam GB

Some highlights from the discussion:

- Program implementers desire greater knowledge about scale and strategies on how to accomplish it. Programs/organizations face significant pressure to achieve “results at scale”, particularly from senior levels of organizations, including political leaders. Some participants identified discomfort and professional dissatisfaction with the view that the measurement of “scale” should be quantitative counts of results. We can do much better.
- A concept of ‘optimality’, as espoused in the previous day’s panel presentation by IDRC, was seen to provide a means of thinking more critically about the design of scaling programs. The position of IDRC was that success and scale are not necessarily equivalent and evaluators could help defining what optimality is for a specific program. What scale makes sense and is more appropriate deserves deep evaluative attention.
- A practical suggestion for improvement was presented during this discussion as the “end of program outcome.” An end of program outcome was described as a result that was reasonably achievable by the program, in its lifecycle. This suggestion was intended to embed reasonable expectations, but it does not mean a grander impact beyond program completion would not be aimed for.
- Evaluation can be used to support program quality at the design stage of a program, and through all the implementing process, including scaling. A firewall between Program and Evaluation to protect independence is not always as necessary since the emphasis has shifted from accountability and reporting to learning and use. Evaluators could be more influential when it comes to decisions about scaling (what is getting scaled and why).

Discussion Point II – The most burning questions about scale in both evaluation and program design. What lines of inquiry should we build into evaluations of scaling programs to help improve future direction and program design? What areas of this subject require further investigation, research, clarification?

Discussion leads:

- Eric Abitbol, Senior Consultant, Universalialia
- Shannon Griswold, Senior Scaling Advisor, Office of Evaluation & Impact Assessment at USAID

Some highlights from the discussion:

- Understanding ‘ownership’ of an innovation among intended beneficiaries is an area of important consideration. It is desirable to consider why/how ownership is achieved during design, and to spend effort evaluating why/how.
- How we describe scale in program design, and measure scale in evaluations should be broadened and deepened. The increased acceptance of qualitative measures would help to move in this direction. Rather than searching for numbers, a scaling strategy should focus on impact. Sustainability and systemic change can be embedded in the broader context that allows impact at scale. The entry point has to be a contextualisation and localisation of a scaling-up intervention. Scaling-up happens in complex ecosystem and the program/intervention is only one aspect of this broader context.
- There is a broad range of ‘pathways’ innovations travel to scale. Critical reflection of the pathway on which independent efforts are travelling is appropriate. The difference between market-intended innovations which might ‘go viral’ vs. social change/norm change which might require more ‘force’ or action was elaborated on as an example. Recognizing diversity in possible pathways, and paying attention as an innovation evolves through a journey to scale is prudent.
- Some efforts are truly designed for scale, others scale naturally. We need more honesty about this at the outset.
- As evaluators, our role can be creating learning mechanisms and process to support the decision-making regarding scaling-up potential.

Discussion Point III – What evaluative methods might help us to get at issues of scale?

Are there specific evaluation methods that are particularly relevant for evaluating programs that aim to achieve impact at scale?

Discussion leads:

- Sonal Zaveri, Secretary, Community of Evaluators, South Asia
- Fabrizio Felloni, Deputy Director, Independent Office of Evaluation, IFAD (*regrets*)

Some highlights from the discussion:

- The model of starting with a pilot and moving toward phased-roll-outs is but one means of planning for program scale. In some cases, this approach can be misleading and is an ineffective means of achieving true impact. We should be wary of over-dependence on this ‘traditional’ pathway to evidence-based program scaling. It was suggested we can too often wind-up in a situation with: “many pilots, but no runways.”
- We must focus as much attention on the negative learning as we do the positive. Be wary of looking only for scaling wins; there is a lot to learn from failures.
- Many efforts to scale will/should fail. As a community, we need to be more accepting of this. Through evaluations and program designs that are accepting of failure we can steer leaders who demand scale toward being more accepting of the risks involved. We should of course be prudent about the implications of failure on all stakeholders.

Concluding remarks were presented to the meeting by Penny Hawkins, Head of Evaluation, DFID.

Penny summarized and interlinked several arguments that held significant value in the discussion. Penny noted that evaluators can play an important role in shaping and re-shaping legitimate scale objectives and claims. Penny noted that: change rarely happens in a linear process. But, defining success within a changing context can be improved through processes such as adaptive management. As evaluators and designers, we can and we should broaden and deepen the ways we aim for and measure scale. We should set honest, responsible goals that support our ability to do well in the world. Doing so will be important to move forward positively.