Executive Summary

Evaluating the "Innovations" in Social Innovation

From Probes to Scaling

Using an adapted version of NESTA's innovation diagram, this brief introduces and illustrates the different tasks for evaluators at each of the five steps in the experimental and performance phases of the innovation process.

In the **Experimental Phase**, which includes rapid and field prototypes, participants develop and test new ideas, and ***the focus is on learning***.

Evaluation at this stage is focused on the reaction of a wide range of potential users, incorporating adaptations based on feedback, usually in real-time, and determining whether the idea merits further development.

- What did they like/not like?
- What feedback do they have on how it can be improved?
- How did the reactions match/differ from what the team anticipated? How does this affect/relate to the intended purpose of the idea?
- To what extent is the idea likely to generate an **impact**?
- To what extent is the idea likely to be **feasible** to implement?
- To what extent is the idea likely to be **viable** in our community?
- To what extent is the idea likely to **align** (i.e. be "complementary," "duplicative," or "counterproductive") with existing initiatives, strategies, and policies?
- To what extent is the idea likely to be **supported** by key stakeholders?

The merit of further development tends to be quite subjective at this stage, but going through the questions encourages deeper thinking about whether the idea will be able to survive "in the real world" and can provide justification for stopping an idea. ***Stopping development of ideas that are "unlikely" to meet the factors above, as early as possible, is just as important in social innovation as developing the promising ones (fail fast, fail often)***.
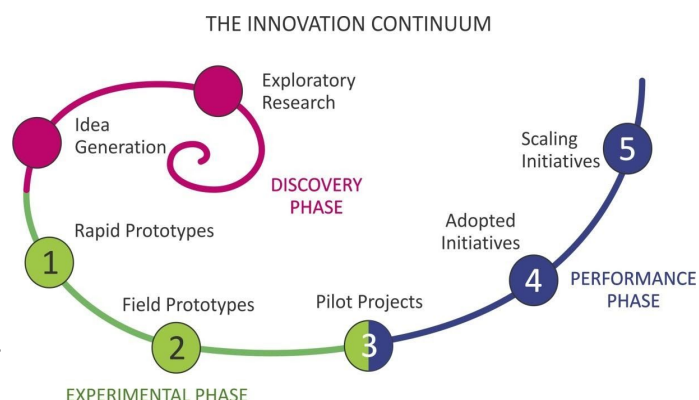
The **Performance Phase** begins by further developing promising ideas into pilots, or a "working model", to ascertain impact and determine whether they should be sustained/implemented and potentially scaled, or discontinued.

The following factors are key to the design and evaluation of pilot projects:

- Sufficient time to track the outcomes the innovation is intended to produce.
- Clear evaluation questions and criteria of "success."
- An evaluation design that generates robust and credible evidence in the eyes of decision-makers.
- Perceived credibility and objectivity of the evaluator.
- Identification of the "real moments" when adaptations are made – which could occur throughout the formal pilot.
- The facilitation of good sense-making and decision-making processes.
- Sufficient resources to make the above possible.

If a pilot demonstrates sufficient success, it is then **"adopted" or implemented**. The term "adopted" is apt because *innovators are often separate from the practitioners/service providers and the innovators have to hand over the idea to a willing "adopter"*.

This is the stage when innovations are expected to generate results. This is why it is critical to provide a robust evaluatory framework to the adopters, based on the learnings from the experimental stages.

Innovation "adopters" are typically concerned with three tasks:

- mobilizing the financial, social, and political support to **sustain** the initiative.
- **adapting** the intervention to reflect the adopters' organization and operating context (continuous improvement).
- managing, tracking, and reporting on the intervention's ongoing **impact**.

The following factors are important to consider in the design and evaluation of adopted innovations:

- **Sustain** - to what extent are we able to mobilize the financial, political and technical resources required to "mainstream" the innovation in our organization or network?
- **Adapt** - what parts of the intervention must be adapted to reflect the operating context of the "adopting" organization(s)? What new capacities, cultures, and structures or processes does the adopting organization require to make the innovation work? To what extent do we have fidelity to the intent, principles, and minimum practices of the original innovation?
- **Impact** - What are the effects, results and impacts of our work over time? How satisfied are the users or beneficiaries of the new practice, model, or service? How can our impact be increased? Costs reduced? Expanded to touch more people?

The final Performance stage is **scaling innovations** - when an innovation is so effective that there is demand for more it, or the benefits are so undeniable that it should scale for broader impact.

Scaling can occur in five distinct, but overlapping ways:

- Scaling Out – the expansion, replication, or dissemination of an innovation to increase the number of people, contexts, or communities impacted. "Impacting greater numbers."
- Scaling Up – changing policies, regulations, structures, and resource flows to support the innovation. "Impacting laws and policies."
- Scaling Deep – changing relationships, cultural values and beliefs, and capturing the hearts-and-minds of the people who need to support the innovation. "Impacting cultural roots."
- Scaling Scree – encouraging, legitimizing, and supporting other innovations that complement the original innovation. "Impacting norms and expectations."
- Scaling Infrastructure – improving the capacity of a community or system to steward and drive the scaling process by changing access to capital, data, talent, knowledge and networks. "Impacting initial conditions."

The following are key considerations to the design and evaluation of scaling innovations:

- **Scaling out** - which parts of the innovation are context sensitive and not easily replicated? Which parts need to be adapted? How far has the innovation spread? What is the ongoing impact of the innovation? Is it growing?
- **Scaling Up** - what is working well and not? Why? How are people reacting to the intervention? How can outcomes be increased? Costs reduced? Quality enhanced?
- **Scale Deep** - to what extent are key stakeholders (e.g., the public, societal influencers, decision-makers) embracing the beliefs, narratives, and values required for the innovation to thrive?
- **Scaling Scree** - what additional ideas, discussions, and experiments have been triggered by the original innovation? To what extent do these innovations complement – or weaken or detract – from the original innovation?
- **Scale Infrastructure -** What resources, skills, networks, and knowledge are required to spread the innovation? How much progress are we making in the creation of this infrastructure? How can it be improved?

When it comes to evaluating social innovations, we have found that it is important for the evaluation to be as learning focused and adaptable as the innovations and prototypes that are being evaluated.