When is a Social Program Ready for Rigorous Impact Evaluation?

Background
The federal government increasingly is looking for strategies to identify promising social programs for broad-scale rollout. Recently, there has been a movement toward requiring demonstrated effectiveness in a rigorous impact evaluation (usually random assignment) as a precondition for broader rollout of a program. However, such rigorous impact evaluations have long timelines and high costs. Furthermore, success rates for programs subject to such rigorous impact evaluation are very low.

There appear to be two strategies for increasing the success rate: (i) screen out programs that are likely to fail rigorous impact evaluation; or (ii) improve the programs such that they are more likely to pass rigorous impact evaluation. However, if only rigorous impact evaluation can establish if a program is truly effective, how can we implement these strategies? A recent article in *Evaluation Review* coauthored by Abt Associates Senior Fellow Jacob Alex Klerman and Diana Epstein of the American Institutes for Research provides a constructive answer to this question. The article suggests an approach based on what the authors call a “falsifiable logic model”—an extension of the conventional logic model commonly used in program development.

Key Points
The article conjectures that too rapid movement from program concept to rigorous impact evaluation is a partial cause of the low success rate of these evaluations. Some programs are undergoing a rigorous impact evaluation too early—before the program satisfies its own ex ante logic model; i.e., the intermediate outcomes that the program developer’s own vision for the program posits will be achieved on the way to the final outcomes of interest.

The authors give five examples of how a program may fail to satisfy its own logic model:

- A program cannot secure partnerships or staff cannot be recruited and retained.
- The program does not attract the target number of clients/participants. While many program models assume that there is a need for the program and that people will enroll, this is not always the case.
- Participants who initially enroll in the program do not complete the expected treatment.
- Clients participate, but the program as implemented falls short of what was envisioned in the logic model.
- Clients show little or no progress on the pre/post measures of the program’s intermediate outcomes.

The essence of the proposed approach is to require program developers to specify these intermediate outcomes—in a falsifiable (often quantitative) form, that can be measured without a comparison group, during or shortly after the program, and at low cost—and then to verify that those intermediate outcomes occur.

Crucially, if such intermediate outcomes are specified, a random assignment study is not needed to detect these forms of logic model failure. Such failures can be detected through a process evaluation of an ongoing program, using only observations on those enrolled in the program, and often only administrative data that should be collected as part of program operation and performance measurement. Random assignment is not needed; neither is a follow-up survey or long follow-up periods.

Experience evaluating programs suggests that a more detailed formative evaluation and falsifiable logic model combined with a careful process evaluation would frequently detect programs that failed, for example, to recruit sufficient qualified program participants or to improve program participants’ intermediate outcomes relative to the program’s goals. The article provides examples of each of these forms of logic model failure.
If the program failed to satisfy the intermediate outcomes of its own falsifiable logic model, it would be deemed unlikely to have positive long-term impacts and therefore should not proceed to the more costly, intensive, rigorous impact evaluation. A careful process evaluation would allow funders to identify programs unlikely to show impacts and therefore not ready for random assignment evaluation—without the high cost and long timelines of a rigorous impact evaluation.

Some programs will proceed to Step C, be improved by the formative evaluation, proceed to Step D, and pass their own falsifiable logic models. With that step achieved, the chances of passing rigorous impact evaluation rise. Alternatively, many programs would fail Step D. Those programs would be deemed unlikely to pass rigorous impact evaluation. Screening out programs that are unlikely to pass rigorous impact evaluation should raise the pass rate of programs that do proceed to rigorous impact evaluation.

**Implications**

Implementing formative evaluations and process evaluations should not be used as an excuse to delay rigorous impact evaluation. Program evaluators should either proceed to the process evaluation and then rigorous impact evaluation, or end the program. Forcing programs that fail their own logic models to reapply for funding for formative evaluation is one possible strategy. This proposed approach is consistent with good science and the reality that most programs subjected to rigorous impact evaluations are not shown to be effective.

However, making evaluation cycles longer might not be feasible for political or practical reasons. The paper suggests a few steps to shrink any increase in timelines:

- Conduct the process analyses relatively quickly and expect prompt reporting of results;
- Issue a joint contract for both the process analysis and the impact analysis with an explicit mid-project review, such that often the impact analysis would not be funded; and
- Opt against repeated cycles of formative evaluation and process evaluation.

Those crafting research strategies will need to weigh a longer timeline for evaluation against its possible advantages: finding more programs that work by helping put scarce evaluation resources towards programs that are truly ready to be evaluated.

**Authors**

**JACOB A. KLERMAN** is a Principal Associate and Senior Fellow at Abt Associates  
**DIANA EPSTEIN** is a Senior Researcher at the American Institutes for Research


[abtassociates.com](http://abtassociates.com)

© 2013 Abt Associates, Inc.