

---

GOLDBLOCKS RESOURCE

# Resources and Tools for Impact Evaluation



Copyright 2016 Innovations for Poverty Action. Resources and Tools for Impact Evaluation is made available under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/). Users are free to copy and redistribute the material in any medium or format.

FEBRUARY 2016



**ipa**  
INNOVATIONS FOR  
POVERTY ACTION

## Resources and Tools for Impact Evaluation

Goldilocks assembled this set of resources for use in designing and running an impact evaluation. Beginning with the need for a theory-driven evaluation, and ending with a set of concrete tools to use in running an evaluation, they cover a range of practical materials useful for organizations that are considering a rigorous impact evaluation.

This set of resources cover the following topics:

- theories of change and impact evaluations
- determining the right timing for an evaluation
- selecting the right evaluation methodology
- designing and managing a randomized evaluation
- calculating sample size and doing power calculations
- mobile data collection
- administrative data
- data management and analysis

For organizations with strong internal technical capacity, these resources will provide valuable practical guidance for designing and implementing a credible impact evaluation. And for organizations without the internal capacity to design and run an impact evaluation, these resources will help key staff members understand the intricacies of a credible evaluation so they can recruit qualified external researchers to carry it out.

Title	Type	Topic	Audience	Level	What it is
<p><b>Using Randomization in Development Economics Research: A Toolkit</b></p> <p>Esther Duflo, Rachel Glennerster, and Michael Kremer. (2008). NBER Technical Working Paper 0333.</p>	Guide	Randomized Evaluations	Researchers, students, and practitioners	Intermediate knowledge of evaluation and statistics	<p>This is an excellent practical guide for using randomization to measure impact of a program in the field. It explains the rationale for using randomization and discusses ways in which this design can be practically implemented in the field.</p> <p>The guide goes into depth on evaluation design elements, such as sample size requirements, stratification, level of randomization and data collection methods. It also discusses how to analyze data when there are technical issues such as imperfect compliance and externalities, and it discusses challenges involved in drawing general conclusions from randomized evaluations</p>
<p><b>Impact Evaluation in Practice</b></p> <p>Paul Gertler Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. (2011). Washington, D.C.: The World Bank.</p>	Guide	Impact Evaluations	Researchers, students, and practitioners	Basic knowledge of evaluation and statistics	<p>This handbook is a comprehensive, accessible, non-technical introduction to impact evaluation and its practical application in development.</p> <p>The book is organized in three parts: 1) what to evaluate and why; 2) how to evaluate (including different evaluation methodologies); and 3) how to implement an evaluation. The book draws on many years of experience conducting impact evaluations in developing countries, and considers operational challenges and ways of overcoming them.</p>

<p><b>Running Randomized Evaluations</b></p> <p>Rachel Glennerster, and Kudzai Takavarasha. (2013). Princeton University Press.</p>	<p>Book</p>	<p>Randomized evaluation</p>	<p>Policymakers and practitioners</p>	<p>Basic knowledge of evaluation and statistics</p>	<p>The book is an indispensable practical guide for policymakers and practitioners on how to design and conduct impact evaluations of social programs in developing countries.</p> <p>The books walks the reader through the process of doing a randomized evaluation, covering such topics as randomization technique, planning for data collection, designing the evaluation, addressing threats to the validity of the experiment, and analyzing the data. It also discusses how to interpret and draw policy conclusions from the results of randomized evaluations and addresses common pitfalls. The book is rich with practical examples, supplemented with detailed references and further reading materials.</p>
<p><b>Theory-Based Impact Evaluation: Principles and Practice</b></p> <p>Howard White. (2009). <i>Journal of Development Effectiveness</i>. Vol. 1, No. 3, 271-284.</p>	<p>Article</p>	<p>Impact Evaluation Theory</p>	<p>Practitioners and evaluators</p>	<p>Basic knowledge of evaluation</p>	<p>This paper underpins the transportability of the CART in supporting the idea that all evaluations should be based in some theory about how the world works. It suggests that a theory-based approach to impact evaluation tells us why a program works, not just whether it works or not.</p> <p>It outlines six principles to theory-based impact evaluation, and advocates combining rigorous impact estimates with other approaches to unpack the causal chain and understand which program components might work and why.</p>

<p><b>When is a Program Ready for Rigorous Impact Evaluation? The Role of a Falsifiable Logic Model</b></p> <p>Diana Epstein, and Jacob Alex Klerman. <i>Evaluation Review</i>. 36(5) 375-401.</p>	<p>Article</p>	<p>Impact Evaluation Theory</p>	<p>Practitioners and Evaluators</p>	<p>Basic knowledge of evaluation</p>	<p>The authors argue that too many program evaluations are performed without thorough due diligence about whether the program is likely to work or not. The article articulates the importance of testing first to be sure that the implementation model is sound before moving to test the immediate outcomes and the full impact.</p> <p>To do so, the authors propose a “falsifiable logic model,” which combined with process evaluation methods, can help identify programs that are unlikely to be found effective by rigorous impact evaluation. To demonstrate how this approach can be applied in practice to screen out programs that are unlikely to achieve intended social impact, the authors provide detailed examples of logic model failures that could have been identified by a process analysis and how this approach can be integrated as part of evaluation process.</p>
<p><b>Optimal Design</b></p>	<p>Statistical program</p>	<p>Sample Size calculator</p>	<p>Evaluators</p>	<p>Intermediate knowledge of evaluation and statistics</p>	<p>Optimal Design is a free software that aids in calculating the necessary sample size for an evaluation. It is relatively easy to use, with menu-based dialogue boxes, graphs, etc.</p> <p>However, the software has some limitations. Among these, it is not possible to calculate power for an individually-randomized trial if your outcome variable is binary, and the software will not calculate different-sized treatment and control groups (rather, it assumes 50-50 split).</p>

<p><b>How to Do Power Calculations in Optimal Design Software</b></p> <p>J-PAL</p>	Guide	Power Calculations	Evaluators, Practitioners	Basic knowledge of statistics	This guide, part of a J-PAL course on evaluations, teaches readers how to use Optimal Design software to determine the required sample size for an evaluation. Using the software, readers work alongside the text for interactive learning.
<p><b>SWOG</b></p>	Statistical Program	Sample Size, Power Calculator	Evaluators	Intermediate-advanced knowledge of evaluation and statistics	SWOG is a set of online statistical tools that helps users calculate sample size, statistical power, and a number of other functions for a variety of different evaluation types.
<p><b>STATA</b></p>	Statistical Program	Sample size, Statistical Power, Analysis	Evaluators	Intermediate-advanced knowledge of evaluation and statistics	Stata is a useful statistical program for intermediate and advanced users. Commands such as sampsi, clustersampsi, or power can be used to calculate sample sizes for evaluation designs. It is also very useful for conducting data analysis and for basic data analysis.
<p><b>A Mobile in Every Pocket: What Does It Mean for M&amp;E?</b></p> <p>Asim Fayez, and Guillaume Kroll. (2016). CEGA.</p>	Article	Data collection	Program management, evaluators	Basic knowledge of surveys	This article describes the current state of mobile data collection. It highlights the potential and drawbacks of mobile data collection as part of an M&E strategy, and suggests things to keep in mind when deciding whether to use mobile or paper-based data collection.
<p><b>Using Mobile Phones in Data Collection: Opportunities, Issues and Challenges</b></p> <p>Michael Trucano. (2014). EduTech (blog), World Bank.</p>	Blog	Mobile Data Collection	Practitioners, evaluators	Basic knowledge of data collection	This blog post provides a simple overview of the opportunities for and challenges to using mobile phones in data collection.

<p><b>Reproducible Research: Best Practices for Data and Code Management</b></p> <p>Harrison Diamond Pollock, Erica Chuang, and Stephanie Wykstra. (2015). Innovations for Poverty Action.</p>	<p>Guide</p>	<p>Data Management</p>	<p>Evaluators</p>	<p>Advanced knowledge of data analysis</p>	<p>This guide outlines best practices for proper data and code management, including documentation and security. While this resource aims to help advanced users produce transparent, reproducible resource, its guidelines on documenting and managing data and code may be useful for advanced practitioners and evaluators.</p>
--	--------------	------------------------	-------------------	--	--