

## Resource Guide 1: General Research Methods

This bibliography, designed for users of the Self-Sufficiency Research Clearinghouse, provides several resources that may be consulted to gain a better understanding of evaluation research, and the different methodological approaches that can be taken in conducting such studies. Several general resources on conducting and interpreting the results of outcome and implementation evaluations are provided. Additionally, there are a number of resources devoted to an explanation of mixed-methods research, as well as performance management and its importance to evaluation research and program development. Researchers may find the reports, briefs, and articles provided in this bibliography useful in helping to plan and conduct studies and analysis of their own. Practitioners and policymakers can utilize these resources to better understand the approaches and data that come out of evaluations. This annotated bibliography is a tool to support practitioners and policymakers in making decisions about funding and collaborating with such researchers to ensure that high quality studies are being conducted and account for effective internal performance management practices.

### General Evaluation Resources

**Allen, T. & Bronte Tinkew, J. (2008). Outcome evaluation: A guide for out-of-school time practitioners – Part 4 in a series on practical evaluation methods. Washington, DC: Child Trends.**

[http://www.childtrends.org/Files//Child\\_Trends\\_2008\\_01\\_07\\_OutcomeEvaluation.pdf](http://www.childtrends.org/Files//Child_Trends_2008_01_07_OutcomeEvaluation.pdf)

Many out-of-school time programs want to learn more about how the children and youth they serve are faring. Outcome evaluations allow programs to do just that. This brief provides a basic review of outcome evaluations, discusses why they are important and when they are useful, and presents guidelines, strategies, and techniques for their use in out-of-school time programs. The brief also highlights the experience of an out-of-school time program that has conducted an outcome evaluation and provides a list of helpful evaluation resources.

**Bloom, H. (November, 2010). Nine lessons about doing evaluation research: Remarks on accepting the Peter H. Rossi Award. Boston MA: Association for Public Policy Analysis and Management Conference**

<http://www.mdr.org/publications/575/presentation.pdf>

Howard Bloom discusses nine lessons about how to successfully conduct evaluation research. The lessons are based on his experiences performing large-scale, rigorous evaluations. Suggestions include considering the importance of design, keeping a strong contrast between treatment and control groups, seeking frequent and continuous feedback, and clear articulation of research questions.

**Bloom, H.S. & Michalopoulos, C. (2010). When is the story in the subgroups? Strategies for interpreting and reporting intervention effects for subgroups. New York, NY: MDRC**

<http://www.mdr.org/publications/551/full.pdf>

This revised working paper examines strategies for interpreting and reporting estimates of intervention effects for subgroups of a study sample. The paper considers why and how subgroup findings are important for applied research, alternative ways to define subgroups, different research questions that motivate subgroup analyses, and the importance of pre-specifying subgroups before analyses are conducted. It also considers the importance of using existing theory and prior research to distinguish between subgroups for whom study findings are confirmatory (hypothesis testing), as opposed to exploratory (hypothesis generating), and the conditions under which study findings should be considered confirmatory. Each issue is illustrated by selected empirical examples.

**Chinman, M., Imm, P., & Wandersman, A. (2004). *Getting To Outcomes 2004: Promoting accountability through methods and tools for planning, implementation, and evaluation*. Santa Monica, CA: RAND**  
[http://www.rand.org/content/dam/rand/pubs/technical\\_reports/2004/RAND\\_TR101.pdf](http://www.rand.org/content/dam/rand/pubs/technical_reports/2004/RAND_TR101.pdf)

The primary purpose of this manual is to help communities improve the quality of their programs aimed at preventing or reducing drug use among youth. Funders are increasingly mandating “accountability” for the public or private funds they provide by demanding high-quality outcome data to determine the success of programs. This manual describes a community planning, implementation, and evaluation model—organized as ten accountability questions—to help your agency, school, or community coalition conduct needs assessments, select best practice programs that fit your community, and to effectively plan, implement, and evaluate those programs. With high-quality process and outcome data, your group will be more likely to get long-term funding for these approaches.

Although this manual was originally developed to help communities plan and carry out programs and policies aimed at preventing youth drug use, it may also be useful for prevention efforts targeted at other youth behavior problems such as crime, teen pregnancy, or delinquency.

**Harvard Family Research Project. (2005). *Evaluation methodology. The Evaluation Exchange: A Periodical on Emerging Strategies in Evaluation*, 11(2), 1-20.**  
[http://www.hfrp.org/evaluation/the-evaluation-exchange/issue\\_archive/evaluation-methodology](http://www.hfrp.org/evaluation/the-evaluation-exchange/issue_archive/evaluation-methodology)

This is the third issue of *The Evaluation Exchange* devoted entirely to the theme of methodology, though within every issue the authors try to identify new methodological choices, the instructive ways in which people have applied or combined different methods, and emerging methodological trends. For example, lately the authors have seen “theories of change” gain almost buzzword status in the field, and have featured this concept in several recent issues, including this one.

Other topics explored here include the role that evaluation theory plays in methodological choices, the proliferation of outcome models during the recent movement toward more accountability for social programs, how to give the connection between evaluation and learning more than just lip service, and developments in the never ending search to ensure that work of researchers avoids being forgotten on dusty bookshelves (or hard drives) and instead gets the attention it deserves. The issue also highlights cluster evaluation and retrospective pretests as important parts of a methodological repertoire, and has several articles on the relationship between evaluation methodology, training, and minority populations.

**Gennetian, L.A., Bos, J.M., & Morris, P.A. (2002). *Using instrumental variables analysis to learn more from social policy experiments*. New York, NY: MDRC**  
<http://www.mdrc.org/publications/24/full.pdf>

One strategy for discovering the connections between social policy interventions and behavioral outcomes is to conduct social experiments that use random assignment research designs. Although random assignment experiments provide reliable estimates of the effects of a particular policy, they do not reveal how a policy brings about its effects. If policymakers had answers to the “how” questions, they could design more effective interventions and make more informed policy trade-offs. This paper reviews one promising approach to specifying the causal paths by which impacts are expected to occur: instrumental variables analysis, a method of estimating the effects of intervening variables - also called mediating variables, or mediators - that link interventions and outcomes. It explores the feasibility of applying this approach to data from random assignment designs, reviews the policy questions that can be answered using the approach, and outlines the conditions that have to be met for the effects of mediating variables to be estimated. Illustrations of instrumental variables analysis based on data from random assignment studies are also presented.

**Lipsey, M.W. & Cordray, D.S. (2000). *Evaluation methods for social intervention*. *Annual Review of Psychology*, 51(1), 345-375.**  
[http://meagherlab.tamu.edu/M\\_Meagher/%20Health%20Psc%20630/Readings%20630/Lipsey.pdf](http://meagherlab.tamu.edu/M_Meagher/%20Health%20Psc%20630/Readings%20630/Lipsey.pdf)

Experimental design is the method of choice for establishing whether social interventions have the intended effects on the populations they are presumed to benefit. Experience with field experiments, however, has revealed significant limitations relating chiefly to (a) practical problems implementing random assignment, (b) important uncontrolled sources of variability occurring after assignment, and (c) a low yield of information for explaining why certain effects were or were not found. In response, it is increasingly common for outcome evaluation to draw on some form of program theory and extend data collection to include descriptive information about program implementation, client characteristics, and patterns of change. These supplements often cannot be readily incorporated into standard experimental design, especially statistical analysis. An important advance in outcome evaluation is the recent development of statistical models that are able to represent individual-level change, correlates of change, and program effects in an integrated and informative manner.

**Shipman, S. & Wholey, J. (2011). Performance measurement and evaluation: Definitions and relationships. Washington, DC: U.S. Government Accountability Office.**

<http://www.gao.gov/assets/80/77277.pdf>

Both the executive branch and congressional committees need evaluative information to help them make decisions about the programs they oversee--information that tells them whether, and why, a program is working well or not. In enacting the Government Performance and Results Act of 1993 (GPRA), Congress expressed frustration that executive and congressional decision-making was often hampered by the lack of good information on the results of federal program efforts. To promote improved federal management and greater efficiency and effectiveness, GPRA instituted a government-wide requirement that agencies set goals and report annually on performance.

Many analytic approaches have been employed over the years by the agencies and others to assess the operations and results of federal programs, policies, activities, and organizations. Most federal agencies now use performance measures to track progress towards goals, but few seem to regularly conduct in-depth program evaluations to assess their programs' impact or learn how to improve results. Individual evaluation studies are designed to answer specific questions about how well a program is working, and GPRA explicitly encourages a complementary role for these types of program assessment. The GPRA Modernization Act of 2010 aims to improve program performance by requiring agencies to identify priority goals, assign officials responsibility for achieving them, and review progress quarterly. Complete and accurate information on how well programs are working and why will be key to its success.

This glossary describes and explains the relationship between two common types of systematic program assessment: performance measures and program evaluation. Based on GAO publications and program evaluation literature, it was first prepared in 1998. Major contributors were Stephanie Shipman and Joseph Wholey. Please address any questions to Stephanie Shipman at (202) 512-4041 or [shipmans@gao.gov](mailto:shipmans@gao.gov).

**Substance Abuse and Mental Health Services Administration. Supported Employment: Evaluating Your Program. DHHS Pub. No. SMA 08-4364, Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services, 2009.**

<http://store.samhsa.gov/shin/content//SMA08-4365/EvaluatingYourProgram-SE.pdf>

This resource describes how and under what circumstances it is best to collect process measures, collect outcome measures, and implement quality assurance systems within the context of evaluating supported employment programs. Some basic suggestions for using data to improve a program are provided, as well as tips for monitoring outcomes, and conducting a readiness assessment. Included in the appendices are measures that may be used, in addition to protocols for administering such measures.

**U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. 2010. The Program Manager's Guide to Evaluation, 2nd ed. Washington, D.C.**

[http://www.acf.hhs.gov/programs/opre/other\\_resrch/pm\\_guide\\_eval/](http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/)

Good program evaluations assess program performance, measure impacts on families and communities, and document our success. With this information, programs are better able to direct limited resources to where they are most needed and most effective in their communities. To help program managers fulfill these goals, the Administration for Children and Families has developed The Program Manager's Guide to Evaluation. The guide explains program evaluation - what it is, how to understand it, and how to do it. It answers your questions about evaluation and explains how to use evaluation to improve programs and benefit staff and families.

**U.S. Government Accountability Office. (2012). Designing evaluations: 2012 revision. Washington, DC: U.S. Government Accountability Office.**

<http://www.gao.gov/assets/590/588146.pdf>

This methodology transfer paper addresses the logic of program evaluation designs. It introduces key issues in planning evaluation studies of federal programs to best meet decision makers' needs while accounting for the constraints evaluators face. It describes different types of evaluations for answering varied questions about program performance, the process of designing evaluation studies, and key issues to consider toward ensuring overall study quality.

Designing Evaluations is a guide to successfully completing evaluation design tasks. It should help GAO evaluators—and others interested in assessing federal programs and policies—plan useful evaluations and become educated consumers of evaluations.

Designing Evaluations is one of a series of papers whose purpose is to provide guides to various aspects of audit and evaluation methodology and indicate where more detailed information is available. It is based on GAO studies and policy documents and program evaluation literature. To ensure the guide's competence and usefulness, drafts were reviewed by selected GAO, federal and state agency evaluators, and evaluation authors and practitioners from professional consulting firms. This paper updates a 1991 version issued by GAO's prior Program Evaluation and Methodology Division. It supersedes that earlier version and incorporates changes in federal program evaluation and performance measurement since GPRA was implemented.

## Mixed Methods

**Frechtling, J. & Sharp, L. (Eds.) (1997). User-friendly handbook for mixed method evaluations. Arlington, VA: National Science Foundation.**

<http://www.nsf.gov/pubs/1997/nsf97153/>

The first publication, User-Friendly Handbook for Project Evaluation: Science, Mathematics, Engineering and Technology Education, issued in 1993, describes the types of evaluations principal investigators/project directors (PIs/PDs) may be called upon to perform over the lifetime of a project. It also describes in some detail the evaluation process, which includes the development of evaluation questions and the collection and analysis of appropriate data to provide answers to these questions. Although this first handbook discussed both qualitative and quantitative methods, it covered techniques that produce numbers (quantitative data) in greater detail. This approach was chosen because decisionmakers usually demand quantitative (statistically documented) evidence of results. Indicators that are often selected to document outcomes include percentage of targeted populations participating in mathematics and science courses, test scores, and percentage of targeted populations selecting careers in the mathematics and science fields.

The current handbook, User-Friendly Guide to Mixed Method Evaluations, builds on the first but seeks to introduce a broader perspective. It was initiated because of the recognition that by focusing primarily on quantitative techniques, evaluators may miss important parts of a story. Experienced evaluators have found that most often the best results are achieved through the use of mixed method evaluations, which combine quantitative and qualitative techniques. Because the earlier handbook did not include an indepth discussion of the collection and analysis of qualitative data, this handbook was initiated to provide more information on qualitative techniques and discuss how they can be combined effectively with quantitative measures.

Like the earlier publication, this handbook is aimed at users who need practical rather than technically sophisticated advice about evaluation methodology. The main objective is to make PIs and PDs "evaluation smart" and to provide the knowledge needed for planning and managing useful evaluations.

---

## Performance Management, Outcomes and Measurement

**Castillo, I.D. (2012). Let's talk about performance management.**

[http://www.childtrends.org/Files//Child\\_Trends\\_2012\\_02\\_14\\_SP\\_RtoR.pdf](http://www.childtrends.org/Files//Child_Trends_2012_02_14_SP_RtoR.pdf)

This presentation explains what performance management is, its importance to both service providers and funders, and provides several real-life examples. The difference between impacts, outputs, and outcomes and how to use them in performance management, is also explained.

**Castillo, I.D. & Schindler, M. (2012). Measuring success & failure: The right (and most useful) amount of evaluation.**

[http://www.childtrends.org/Files//Child\\_Trends\\_2012\\_02\\_02\\_SP\\_RightEval.pdf](http://www.childtrends.org/Files//Child_Trends_2012_02_02_SP_RightEval.pdf)

This presentation discusses the importance of outcomes and evaluation to service providers and funders, the need to manage expectation for outcomes, and differences between impacts and outcomes. Performance management and evaluation, and the differences between the two, are also examined.

**Introduction to the Evaluation Data Coordination Project. Retrieved from**

[http://www.acf.hhs.gov/programs/opre/other\\_resrch/eval\\_data/reports/common\\_constructs/com\\_intro.html](http://www.acf.hhs.gov/programs/opre/other_resrch/eval_data/reports/common_constructs/com_intro.html)

OPRE funded this study based on the perspective that coordinating outcome measures and data collection across multiple evaluation projects is crucial for making comparisons across evaluations and for facilitating cross-study research, such as meta-analyses, after the studies have been concluded. This coordination will help researchers be more certain that cross-program differences in impacts on the same construct are due to differences in the effectiveness of the programs (as implemented with different populations and in different contexts) instead of differences in how the construct is measured. Coordinating the inclusion of identical, well-established measures across multiple evaluation studies will have a dramatic influence on the usefulness of these data to researchers and policymakers in the years to come. The work undertaken in this process will also enhance data collection efforts in future research.

**Kasunic, M. (2005). *Designing an effective survey*. Pittsburgh, PA: Carnegie Mellon Software Engineering Institute.**  
<http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA441817>

A survey can characterize the knowledge, attitudes, and behaviors of a large group of people through the study of a subset of them. However, to protect the validity of conclusions drawn from a survey, certain procedures must be followed throughout the process of designing, developing, and distributing the survey questionnaire.

Surveys are used extensively by software and systems engineering organizations to provide insight into complex issues, assist with problem solving, and support effective decision making.

This document presents a seven-stage, end-to-end process for conducting a survey.

**Walker, K.E. & Moore, K.A. (2011). *Performance management and evaluation: What's the difference?* Washington, DC: Child Trends.**  
[http://www.childtrends.org/Files//Child\\_Trends\\_2011\\_01\\_19\\_RB\\_PerformMgmt.pdf](http://www.childtrends.org/Files//Child_Trends_2011_01_19_RB_PerformMgmt.pdf)

In previous research briefs, Child Trends has provided overviews of various forms of program evaluations: experimental, quasi-experimental, and process. This brief provides information on performance management—the ongoing process of collecting and analyzing information to monitor program performance—and its relationship to other forms of evaluation.

---

Note: Unless otherwise noted all information is from the publication and/or Web site.