



The purpose of the IOCE (International Organization for Cooperation in Evaluation) is to promote an international movement that enhances the development and strengthens the capacity of professional evaluators and evaluation organizations around the world. As an evaluator, you may be an active member of one or more professional evaluation association/society/networks. If those organizations are not yet members of the IOCE you should encourage them to join.

# Evaluating program impacts and effects

## A list of references and resources

### IOCE<sup>1</sup>

### October 2011

Asher, M. (1983) *Causal Modeling* (Sage). (an introduction to causal/statistical modeling)

Asian Development Bank (2006), *Impact Evaluation: Methodological and Operational Issues*. (a brief introduction, interesting discussion of common objections to impact studies)

August 2011

Bamberger, et al (2006) *RealWorld Evaluation*, Sage. (an excellent overview of how to undertake evaluations of development programs while facing various types of constraints, also includes a discussion of the most commonly used designs for evaluating the impact of development programs)

Bamberger, M., Rugh, J. Mabry, L. (2006) *RealWorld Evaluation*. Thousand Oaks, CA: Sage Publications.

Barca, F. (2009) *An agenda for a reformed cohesion policy*

---

<sup>1</sup> Thanks to many sources for this bibliography but in particular the following. The EES statement (2007) drafted on behalf of the EES Board of Directors by a workgroup led by Burt Perrin and the piece by Ben Ramalingham in 2011, learning how to learn: eight lessons for impact evaluations that make a difference (ODI, London). Also, thanks should go to Scott Bayley (also for some annotations) and Jim Rugh for their generous contributions.

(EU: [http://ec.europa.eu/regional\\_policy/policy/future/barca\\_en.html](http://ec.europa.eu/regional_policy/policy/future/barca_en.html))

Barlow & Hersen (1989) *Single Case Experimental Designs*, (Pergamon)

Becker, D (2000), Discussion Notes: Causality, <http://web.uccs.edu/lbecker/Psy590/cause.htm> (a brief summary of different philosophical perspectives on causality)

Bhola, H.S. (2000). A discourse on impact evaluation: a model and its application to a literacy intervention in Ghana. *Evaluation*, 6(2), 161-178.

Borton, J. (2004) 'The Joint Evaluation of Emergency Assistance to Rwanda', *Humanitarian Exchange*, Issue 26, March.

Boruch, 2005, *Randomized Experiments for planning and evaluation: A practical guide*, Sage. (a good introduction to the topic)

Brady, H (2002) *Models of Causal Inference: Going Beyond the Neyman-Rubin-Holland Theory*, Paper Presented at the Annual Meetings of the Political Methodology Group, University of Washington, Seattle, Washington. (paper reviews four of the more common theories of causality)

Briceño, B. and Gaarder, M. (2009) 'Institutionalizing Evaluation: Review of International Experience'. 3ie Research Paper. London: 3ie ([www.3ieimpact.org/reports/India\\_Report\\_DFID.pdf](http://www.3ieimpact.org/reports/India_Report_DFID.pdf)).

Brinkerhoff, D. (1991) *Improving Development Program Performance: Guidelines for Managers*, Lynne Rienner. (includes a discussion of the most common causes of performance problems in development programs)

Buchanan-Smith, M. (2003) *How the Sphere Project Came Into Being: A Case Study of Policy Making in the Humanitarian Aid Sector and the Relative Influence of Research*. Working Paper 215. London: ODI.

Campbell, D. and Stanley, J. (1963) *Experimental and Quasi-experimental Designs for Research*, Rand McNally. (the all-time classic text, discusses the strengths and weaknesses of various research designs for assessing program impacts)

Campbell, D., T. (1979). "Degrees of freedom" and the case study. In Donald T. Cook and Charles S. Reichardt, eds., *Qualitative and quantitative methods in evaluation research*, Thousand Oaks: Sage Publications. Reprinted in: Donald T. Campbell. (1988). *Methodology and epistemology for the social sciences; Selected papers* (E. S. Overman, ed.). Chicago: University of Chicago Press.

Carter N, Klein, R and Day P, (1992) *How organisations measure success: the use of performance indicators in government*, Routledge. (offers a useful typology of different types of PIs and how they can be used/misused)

Catley, A, et al. (2008) *Participatory Impact evaluation: A Guide for Practitioners*. Medford, MA: Feinstein International Center, Tufts University  
(<http://wikis.uit.tufts.edu/confluence/display/FIC/Participatory+Impact+Assessment>).

Center for Global Development (CGD) (2006) 'When Will We Ever Learn? Improving Lives through Impact Evaluation'. Report of the Evaluation Gap Working Group. Washington, DC: CGD ([www.cgdev.org/content/publications/detail/7973](http://www.cgdev.org/content/publications/detail/7973)).

Chambers, R. (2009) *So that the Poor Count More: Using Participatory Methods for Impact Evaluation in Designing impact evaluations: different perspectives*. 3ie Working Paper 4. London: 3iE ([www.3ieimpact.org/admin/pdfs\\_papers/50.pdf](http://www.3ieimpact.org/admin/pdfs_papers/50.pdf)).

Chatterji, M. (2007). Grades of evidence: Variability in quality of findings in effectiveness studies of complex field interventions. *American Journal of Evaluation*, 2(3), 239-255.

Chelimsky, E. (1997) Thoughts for a new evaluation society, *Evaluation*, 3:p 97-109.

Cook T and Campbell D (1979) *Quasi-experimentation*, Houghton Mifflin. (contains a useful review of different theories of causality and how to test for causal relationships as well as the application of quasi-experiments for impact evaluations)

Cook T, Shadish, W and Wong V, 2008, 'Three Conditions under Which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings from Within-Study Comparisons', *Journal of Policy Analysis and Management*, 27, 4, 724-750. (this reference recommends regression-discontinuity designs, matching geographically local groups on pre-treatment outcome measures, and modeling a known selection process)

Cook, T. D. (1997). Lessons learned in evaluation over the past 25 years. In Eleanor Chelimsky and William R. Shadish. (eds.) *Evaluation for the 21st Century: A handbook*. Thousand Oaks, CA: Sage Publications, 30-52.

Cook, T. D. (2000). Towards a practical theory of external validity. In Leonard Bickman (ed.). *Validity and social experimentation: Donald Campbell's legacy Vol. 1*. Thousand Oaks, CA: Sage Publications, 3-44.

Cracknell, B. E. (2001). Knowing is all: Or is it? Some reflections on why the acquisition of knowledge focusing particularly on evaluation activities, does not always lead to action. *Public Administration and Development*, 31, 371-379.

Cracknell, B. (2000) *Evaluating Development Aid: Issue, Problems and Solutions*. Sage Publications, New Delhi. (interesting discussions of the issue of evaluating for learning vs evaluating for accountability and the politics of evaluation)

Davidson, E. J. (2004) *Evaluation methodology basics: The nuts and bolts of sound evaluation*, Sage. (suggests 8 techniques for causal inference)

Davidson, E.J. (2000). 'Ascertaining causation in theory-based evaluation'. In P.J. Rogers, T.A. Hacsí, A. Petrosino, and T.A. Huebner (eds.), "Program theory in evaluation: challenges and opportunities". *New Directions in Evaluation*, Number 87:17-26, San Francisco, CA. (provides an overview of various methods)

Davis, J (1985), *The Logic of Causal Order*, (Sage)

Deaton, A. (2009) *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*. Princeton, NJ: Research Program in Development Studies, Center for Health and Wellbeing, Princeton University ([www.princeton.edu/~deaton/downloads/Instruments\\_of\\_Development.pdf](http://www.princeton.edu/~deaton/downloads/Instruments_of_Development.pdf)).

Deaton, A. (2010) *What Can We Learn From Randomized Control Trials?* Chicago: Milton Friedman Institute ([http://mfi.uchicago.edu/events/20100225\\_randomizedtrials/index.shtml](http://mfi.uchicago.edu/events/20100225_randomizedtrials/index.shtml)).

Dept of Finance and Administration, 2006, Handbook of Cost-Benefit Analysis, Australian Government Publishing Service. (an easy to read introduction). This book is available free on the net at:

Dept of Finance, 1987, Evaluating Government Programs, Australian Government Publishing Service. (introductory, includes a useful table comparing different types of research designs)

Dhaliwal, I. (2011) Conditional cash transfers and health: unpacking the causal chain (<http://nonie2011.org/?q=node/88>).

Donaldson, S Christie C and Mark, M (2009), What Counts as Credible Evidence in (*eds*) Donaldson, S Christie C and Mark, M *Applied Research and Evaluation Practice* (Sage).

European Evaluation Society, 2007, Statement: The Importance of a Methodologically Diverse Approach to Impact Evaluation.

Foresti, M. (2007) 'A Comparative Study of Evaluation Policies and Practices in Development Agencies'. Report for AFD Evaluation Department. London: ODI.

Gilovich, T (1991) How we know what isn't so: The fallibility of human reason in everyday life, (Free Press, New York)

Glazerman, S. Levy, D. & Myers, D. (2002), Nonexperimental Replications of Social Experiments: A Systematic Review, Mathematica Policy Research Inc. (this research paper concludes that more often than not, statistical models do a poor job of estimating program impacts).

Guba, E and Lincoln, Y (1989) Fourth Generation Evaluation, (Sage). (the authors argue that 'cause and effect' do not exist except by imputation, a constructivist perspective))

Hallam, A. (2011) 'Harnessing the power of evaluation of humanitarian action: An initiative to improve understanding and use of evaluation'. Draft ALNAP Study. London: Overseas Development Institute.

Hatry, H (1986) Practical Program Evaluation for State and Local Governments, Urban Institute Press. (a good introduction, includes a review of the circumstances in which it is feasible to use experimental designs)

Holland, P. (1986) 'Statistics and Causal Inference'. *Journal of the American Statistical Association*. Vol. 81 (945-960).

House, E., R. (2006). Democracy and evaluation. *Evaluation*. 12(1), 119-127. Hulme, D. (2000) 'Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice', *World Development* 28 (1): 79-98.

Jones, H. (2009) 'The "gold standard" is not a silver bullet for evaluation', ODI Opinion 127. London: ODI ([www.odi.org.uk/resources/download/2811.pdf](http://www.odi.org.uk/resources/download/2811.pdf)).

Jones, N. et al. (2009) 'Improving Impact Evaluation Coordination and Use'. A Scoping Study commissioned by the DFID Evaluation Department on behalf of NONIE ([www.odi.org.uk/resources/download/3177.pdf](http://www.odi.org.uk/resources/download/3177.pdf)).

Judd, C. and Kenny D, (1981), *Estimating the Effects of Social Interventions*, Cambridge. (heavy emphasis on statistical applications, for the enthusiast)

Kenny, D. (2004), *Correlation and causality*, (a very technical book about analysing causal impacts using statistical models). This book is available free on the net at: <http://davidakenny.net/cm/cc.htm>

Langbein, J (1980) *Discovering Whether Programs Work*, (Goodyear). (good but technical)

Larson, 1980, *Why government programs fail*, Praeger. (The reasons: a faulty theory of change; poor implementation; a changing external environment; or the evaluation itself is faulty)

Mark, M. and Reichardt, C. (2004) 'Quasi-experimental and correlational designs: Methods for the real world when random assignment isn't feasible'. In Sansone, Morf and Panter, (eds), *Handbook of methods in social psychology*, (pp. 265-286), Sage. (useful introductory overview)

Mark, M. R. (2001). What works and how can we tell? Evaluation Seminar 2. Victoria Department of Natural Resources and Environment.

Mayne, J. (2001). Addressing attribution through contribution Analysis : Using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16(1), 1-24.

Mayne, J. (2008) Contribution analysis: An approach to exploring cause and effect, ILAC Brief 16. (an approach that is increasingly popular based on using program theory and shares the same strengths/weaknesses)

McMillan, J. (2007) Randomized Field Trials and Internal Validity: Not So Fast My Friend, (good overview of the limitations). Available free on the net at:

Miles, M. and Huberman, A. 1994, Qualitative data analysis, (Sage) (contains examples of undertaking causal analysis with qualitative data)

Mohr, L. B. (1995) Impact Analysis for Program Evaluation, (Sage) (an advanced discussion of research designs and impact analysis)

Mohr, L. B. 1999, 'The Qualitative Method of Impact Analysis', American Journal of Evaluation, 20, 1, pp. 69-84. (useful introduction to the topic)

Mohr, L. B. (1999). The qualitative method of impact analysis. American Journal of Evaluation. 20(1), 69-84.

Network of Networks on Impact Evaluation, 2009, Impact Evaluations and Development – NONIE Guidance on Impact Evaluations. (a review of the methods commonly used by development agencies). Available free on the net at:

Nisbett, R. and Ross, L. (1985) Human Inference: Strategies and Shortcomings of Social Judgments, (Prentice-Hall) (explains why people struggle to accurately perceive causal relationships)

Norad, 2008, The Challenge of Assessing Aid Impact: A Review of Norwegian Evaluation Practice, (provides a number of examples of problematic impact evaluations along with various lessons for better practice). Available free on the net at:

Nutt, P. (2002) Why Decisions Fail, (Berrett-Koehler Publishers). (interesting review of why strategic decisions often fail, e.g. lack of consultation, poor analysis, faulty implementation,)

OIOS (2008) 'Review of Results-based Management at the United Nations'. Report of the Office for Internal Oversight Services (OIOS). Paris: UN OIOS.

Patton, M. Q. (1987) *How to use qualitative methods in evaluation*, Sage. (excellent discussion of combining qualitative and quantitative methods)

Patton, M. Q. (1990) *Qualitative Evaluation and Research Methods*, Sage. (good all round reference, excellent description of different types of purposeful sampling)

Patton, M. Q. (1997) *Utilization-focused evaluation: The New Century Text*. 3rd Ed. Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2008a) 'State of the Art in Measuring Development Assistance'. Address to the World Bank Independent Evaluation Group, 10 April, Washington, DC ([www.worldbank.org/ieg/conference/results/patton.pdf](http://www.worldbank.org/ieg/conference/results/patton.pdf)).

Patton, M. Q. (2008b) *Utilisation-Focused Evaluation*. 4th Edn. London: Sage.

Patton, M. Q. (2010) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.

Pawson, R. (2002a). Evidence-based policy: In search of a method. *Evaluation*, 8(2), 157-181.

Pawson, R. (2002b). Evidence-based policy: The promise of 'realist synthesis'. *Evaluation*, 8(3), 340-358.

Pawson, R., and Tilley, N.(1997) *Realistic evaluation*. London: Sage Publications.

Pearl, 2000, *Causality: Models, Reasoning, and Inference*, Cambridge University Press. (advanced technical examination of causal/statistical modeling)

Perrin, B. (2000). Donald T. Campbell and the art of practical "in the trenches" program evaluation. In Leonard Bickman (ed.). *Validity and social experimentation: Donald Campbell's legacy* Vol. 1. Thousand Oaks, CA: Sage Publications, 267-282.

Perrin, B. (1998). 'Effective Use and Misuse of Performance Measurement'. *American Journal of Evaluation*. Vol. 19 (1):367-379.

Perrin, B (1999) 'Performance Measurement: Does the Reality Match the Rhetoric? A Rejoinder to Bernstein and Winston'. *American Journal of Evaluation*. Vol. 20(1).

Posavac, E. and Carey, R. (2002) *Program Evaluation: Methods and Case Studies*, Prentice Hall.  
(good all round text, includes a summary of the types of evaluation questions that can be answered by particular research designs)

Proudlock, K. and Ramalingam, B. with Sandison, P. (2009) 'Improving humanitarian impact evaluation: Bridging theory and practice' in ALNAP (ed.), *ALNAP 8th Review of Humanitarian Action: Performance, Impact and Innovation*. London: Overseas Development Institute.

Public/Private Ventures (2011) *Priorities for a New Decade: Making (More) Social Programs Work (Better)* ([www.ppv.org/ppv/publications/assets/329\\_publication.pdf](http://www.ppv.org/ppv/publications/assets/329_publication.pdf)).

Ramalingam, B. and Jones, H. with Young, J. and Reba, T. (2008) *Exploring the Science of complexity: ideas and implications for development and humanitarian efforts*. ODI Working Paper 285. London: ODI.

Ramalingam B. (2011) *Learning how to learn: eight lessons for impact evaluations that make a difference* (ODI, London)

Reichardt, C. (2000) 'A typology of strategies for ruling out threats to validity'. In Bickman (ed) *Research Design: Donald Campbell's' legacy*, Sage.

Reynolds, K. and West, S. (1987) 'A multiplist strategy for strengthening nonequivalent control group designs', *Evaluation Review*, 11, 6, 691-714. (an excellent example of how to fix up a weak research design by adding additional features thereby improving your assessment of the program's impact)

Riddell, R. (2007) *Does Foreign Aid Really Work?* (Oxford: OUP)

Riddell, R. (2009) 'Is Aid Working? Is this the right question to be asking?', *Open Democracy*, 20 November ([www.opendemocracy.net/roger-c-riddell/is-aid-working-is-this-right-question-to-be-asking](http://www.opendemocracy.net/roger-c-riddell/is-aid-working-is-this-right-question-to-be-asking)).

Rodrik, D. (2008) 'The new development economics: we shall experiment, but how shall we learn?'. Revised version of paper presented at Brookings Development Conference, 29-30 May ([www.hks.harvard.edu/fs/drodrik/Research%20papers/The%20New%20Development%20Economic%20s.pdf](http://www.hks.harvard.edu/fs/drodrik/Research%20papers/The%20New%20Development%20Economic%20s.pdf)).

Rogers, P. et al (2000) *Program Theory in Evaluation: Challenges and Opportunities*, No. 87. (a series of papers on the strengths and weaknesses of using program theory to assist with causal analysis)

Rogers, P. (2008) Using programme theory to evaluate complicated and complex aspects of interventions., *Evaluation* January 2008 vol. 14 no. 1 29-48

Rogers, P. (2009) Matching Impact Evaluation Design to the Nature of the Intervention and the Purpose of the Evaluation in Designing impact evaluations: different perspectives. 3ie Working Paper 4. London: 3ie ([www.3ieimpact.org/admin/pdfs\\_papers/50.pdf](http://www.3ieimpact.org/admin/pdfs_papers/50.pdf)).

Roodman, D. (2008) Through the Looking Glass, and What OLS Found There: On Growth, Foreign Aid, and Reverse Causality, Working Paper 137, Center for Global Development.

Rossi, P. Lipsey, M and Freeman, H ( 2003) *Evaluation – A Systematic Approach*, Sage.  
(recommended, includes an excellent discussion of different types of research designs and when to use each of them)

Rothman, K and Greenland, S. (2005) 'Causation and Causal Inference in Epidemiology', *American Journal of Public Health*, Vol 95, No. S1.

Sandison, P. (2006) 'The Utilisation of Evaluations' in ALNAP (ed.) ALNAP Review of Humanitarian Action 2005. London: ODI ([www.alnap.org/publications/RHA2005/rha05\\_Ch3.pdf](http://www.alnap.org/publications/RHA2005/rha05_Ch3.pdf)).

Saunders, M. (2011) Capturing effects of interventions, policies and programmes in the European context: a social practice perspective in *Evaluation*, Vol. 17, No 1, pp 89-103.

Schiefer, U. (2008) Integrated evaluation of change: A different perspective for evaluation in multiple intervention environments. *Evaluation*, 14(1)

Scriven, MI. (1976). Maximizing the power of causal investigations: The modus operandi method. In Gene V. Glass (ed.) *Evaluation studies review annual*, Volume 1 (101-118). Beverly Hills, CA: Sage Publications.

Shadish, W. Cook, T. and Campbell, D (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Houghton Mifflin).

Shadish, W, Cook T. and Leviton, L. ( 1991) *Foundations of Program Evaluation*, (Sage) (the final chapter contains an excellent summary of evaluation theory in relation to program design, evaluation practice, and theory of use)

Spector, P. (1981) *Research Designs*, (Sage) (a basic introduction)

Stame, N, (2010), 'What Doesn't Work? Three Failures, Many Answers', *Evaluation*, 16, 4, 371-387. (useful review of current debates on impact evaluation methodology)

Stein, J.G. (2008) 'Humanitarian Organisations: Accountable – Why, to Whom, for What, and How?' in M. Barnett and T.G. Weiss (eds), *Humanitarianism in Question: Politics, Power, Ethics*. Ithaca and London: Cornell University Press.

Stern, E. (2009), Evaluation policy in the European Union and its institutions. *New Directions for Evaluation*, 2009: 67–85. doi: 10.1002/ev.306

Stryczynski K. (2008) *Rigorous impact evaluation: the magic bullet for evaluation of Cohesion Policy?* (paper delivered at the Lisbon EES International Conference, 2008)

The Lancet (2004) 'The World Bank is finally embracing science'. Editorial, 28 August.

Treasury Board of Canada, no date, Program Evaluation Methods: Measurement and Attribution of Program Results, (useful overview). It is available free on the net at:

Trochim, W. (1984), *Research designs for program evaluation: The regression discontinuity approach*, Sage.

Trochim, W. (1989) 'Outcome Pattern Matching and Program Theory'. *Evaluation and Program Planning*. Vol. 12:355-366.

Weick, K. (1995) *Sense making in Organisations* (Sage, New York)

Weyrauch, V. and Langou, G. D. (2011) *Sound Expectations: From Impact Evaluations to Policy Change*. 3ie Working Paper 12. London: 3ie ([www.3ieimpact.org/3ie\\_working\\_papers.html](http://www.3ieimpact.org/3ie_working_papers.html)).

White, H. (2011) '*Five challenges facing impact evaluation on NONIE*' (<http://nonie2011.org/?q=content/post-2>).

Williams, R. and Imam, I (eds) (2007). *Systems concepts in evaluation: An expert anthology*. AEA Monograph Series. Pont Reyes, CA: (Edge Press of Inverness)

World Bank (Independent Evaluation Group) 2006, *Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints*, author. (this text is a highly summarized version of Bamberger's book).

Yin, R. (2000) 'Rival Explanations as an Alternative to Reforms as Experiments', in Bickman (ed) *Validity and Social Experimentation*, (Sage) (good review of how to identify and test rival explanations when evaluating reforms or complex social change)

Yin, R. (2003) *Applications of Case Study Research*,( Sage) (very good reference, includes advice on undertaking causal analysis using case studies)

Young, J. and Mendizabal, E. (2009) '*Helping researchers become policy entrepreneurs*', ODI Briefing Papers 53. London: ODI [www.odi.org.uk/resources/download/1127.pdf](http://www.odi.org.uk/resources/download/1127.pdf)).



IE4D (2011) Impact Evaluation for Development: principles for Action

[www.keystoneaccountability.org/sites/default/files/IE4DFinal%20Jan%202017.pdf](http://www.keystoneaccountability.org/sites/default/files/IE4DFinal%20Jan%202017.pdf)