



International Organisation for Cooperation in Evaluation
Organisation internationale de coopération en évaluation
Organizaciòn Internacional para la Cooperaciòn en Evaluaciòn



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.

Developing an inclusive approach to evaluating the impact or effects of policy and development interventions

Murray Saunders on behalf of IOCE

The IOCE is an alliance of national, regional and international evaluation organizations (associations, societies and networks). Its mission is to:

- build evaluation leadership and capacity, especially in developing countries;
- foster the cross-fertilization of evaluation theory and practice across borders;
- address major challenges facing evaluators worldwide and
- contribute to the identification and solution of contemporary global problems through independent and high quality evaluation.

In this context, its position on the evaluation of the effects of development interventions, programmes and policies (also called impact assessment or impact evaluation) aims to encourage a methodologically diverse approach by practitioners, commissioners and other stakeholders that acknowledges contextual requirements, available evidence and the usability of evaluation outputs. This position takes account of the stance taken by other national and regional associations and networks.

This statement

The IOCE statement outlines a broad position on the evaluation of impact. It is not an exhaustive account of the current debate. Nor does it intend to assess the strengths and

Vision

The IOCE envisions stronger evaluation theory and practice worldwide through international collaboration and partnership in evaluation which foster the cross-fertilization of ideas, high professional standards and open and global perspectives among evaluators.

Mission

To help legitimate evaluation and support evaluation societies, associations and networks so that they can better contribute to good governance, effective decision making and strengthen the role of civil society.

weaknesses of specific approaches. It simply argues for an open and inclusive approach. It is informed by what is feasible in specific and often complex environments while promoting rigorous and well designed evaluations. The statement is consistent with IOCE's mission of promoting the 'theory, practice and utilisation of high quality evaluation'.

IOCE endorses efforts to improve impact evaluation or impact assessment, with respect to project, program and policy evaluations, including those carried out in development cooperation contexts. This new emphasis on assessing the effects of interventions and policies is prompted by three drivers (Saunders, 2011). The first is geared to a better understanding of complex changes brought about by different kinds of intervention, the second aims at better economic and social stewardship and the third has to do with improved evaluation methods.

Improving understanding

The urge to 'sense make' in complex environments is an imperative when evaluations seek to build knowledge of the conditions under which interventions are designed to bring about improvements in citizen's lives. Eleanor Chelimskey's formulation based on discussions with practising evaluators speaks of 'evaluation for knowledge' (1997). The emerging requirement that development interventions or cohesion policies in Europe, for example, are shown to have positive effects of a broad-based kind, prompts the search for ways of 'knowing'. The source of this impulse is the need to know more about how complex changes are brought about by a variety of policy levers, instruments and mechanisms.

Better economic and social stewardship

The second driver is derived from an increasingly acute recognition that judicious and well managed public expenditures geared to social and economic change are needed to tackle obstacles to peace and prosperity. We are now entering an era during which public money will be under close and rigorous scrutiny. There will be less funding for the delivery of global, regional, national and local public goods.

In these circumstances, the need for transparency and equity in the distribution, use and effects of public spending has become increasingly important. It is here that a robust and legitimate set of evaluative practices can play a useful role. In this sense evaluative practice should respond to a growing need for social capital and for trust in the processes, protocols and procedures associated with public programs. To that extent, the legitimacy and authority of the principles and procedures associated with the procurement, undertaking and use of evaluation outputs moves to the centre stage. Some policy domains are slippery, ambiguous and unformed, relying on enabling networks, collaborations and partnerships (see Stern 2009). Increasingly, evaluation is being understood as part of the process by which 'policy learning' or institutional growth and development is encouraged.

This means that evaluation is mandated to restore currently low levels of social capital that result from rigid and unresponsive political systems that fail to engage citizens and can even lead to corruption. The remedy lies in enhanced citizens' 'participation' in public governance since public institutions - formal and informal - require social capital in order to function effectively. Evaluation is one means to compensate for such systemic weaknesses through engagement of the civil society and challenges to ways of making decisions, organizing

society and governing the public space. However this imperative has associated with it the greater responsibility that evaluation output is used.

Improving method

The third driver is methodological. While a recurring theme, the need for more certainty about what counts as a positive change and how we might know its attribution while elusive, is growing in intensity. This explains the renewed interest in experimental designs as a way of producing ‘certainty’, though it runs counter to the ethical constraints inherent in experimental methods (Stryczynski 2008). Appropriately the Barca report suggests that *the complex methodological issues which characterise the use of a counterfactual approach in the context of policy suggest that the effort to promote impact evaluation as one of the methodological backbones of cohesion policy must at the same time be visionary and humble. (Barca 2009 p180)*

Donors’, commissioners’ and politicians’ sensitivities about the visibility of their ‘change theories’ (e.g. for example, structural, agricultural or business development) underlies their concern to address issues of attribution or causality. This is legitimate even though the ‘logic lines’ may be difficult to establish. The overarching question is: ‘under what circumstances and why is a particular strategy rather than another likely to yield a positive effect?’

Aspects of present practice

Four bi-annual meta-evaluations of evaluation reports from CARE projects in many countries provide a telling story. Colleagues familiar with other agencies report proportions of evaluations without pre-test + post-test, nor counterfactual, are typically even lower than the percentages shown in the Table below (Bamberger 2006). The extent to which even rudimentary experimental designs are used is very low.

What kinds of evaluation designs are actually used in the real world (of international development)? Findings from meta-evaluations of 336 evaluation reports of an INGO.

Post-test only	59%
Before-and-after	25%
With-and-without	15%
Other counterfactual	1%

Proponents of experimental methods have acknowledged that randomised control trials (RCTs) are only appropriate for perhaps 5% of development interventions (Bamberger 2006). Examinations of evaluations in the OECD/DAC DEReC database by bilateral and multilateral organisations found only 5% used even a counterfactual design. While we recognize that experimental and quasi experimental designs have a place in the toolkit for impact evaluations and establishing the effects of interventions, this statement supports the

view that more attention needs to be paid to the roughly 95% of situations where these designs would not be possible or appropriate.

- The literature clearly documents how *all* methods and approaches have strengths and limitations and that there are a wide range of scientific, evidence-based, rigorous approaches to evaluation that have been used in varying contexts for assessing impact.
- Impact Evaluation (IE) is complex, particularly of multi-dimensional interventions such as many forms of development, and consequently requires the use of a variety of different methods that can take into account rather than dismiss this inherent complexity.
- Evaluation standards and principles from across the world are pluralistic in nature. They do not favour any specific approach or group of approaches. Some of the evaluation standards go on to specify that the evaluator should give reasons for selecting a particular evaluation approach or combination, commensurate with the characteristics of the program being evaluated, and the purposes of evaluation.

In the light of these realities, the IOCE calls for a multiple method approach to IE that does not consider any single method (such as RCTs) as first choice or as the ‘gold standard’. More importantly, we are mindful of Howard White’s (2011) exhortation, reported by Ben Ramalingham (2011) in the ODI note ‘Learning how to learn: eight lessons for impact evaluations that make a difference’. To paraphrase and adapt White’s points, we encourage the view that evaluative practice should be shaped by the following challenges:

- Identify and strengthen processes to ensure that evidence is used in policy.
- Institutionalise the evaluation of effects and outcomes.
- Improve evaluation designs to answer policy-relevant questions.
- Make progress with carefully designed small scale evaluations of impact.
- Expand knowledge and use of systematic reviews.

Recent research and literature have questioned if direct attribution or causality, in an unproblematic sense, may be the right question, given the complexity of factors that need to interact together to create effects. In other words, other concepts, like for example ‘plausible contribution’ or ‘reasonable attribution’ may be more applicable and usable.

The Future

So, in the context of the Paris Declaration and the Millennium Development Goals, it is appropriate for the international evaluation community to work together in supporting the development of local capacity to undertake the evaluation of effects and impact by acknowledging these contextual factors. The IOCE endorses the following foci for evaluations adapted from statements by OECD/DAC. Evaluations should take account of:

Relevance: the extent to which the aid activity and policy is suited to the priorities and policies of the target group, recipient and donor,

Effectiveness: the extent to which an aid or policy activity attains its objectives,

Efficiency: the outputs and outcomes in relation to the inputs,

Impact and effects: the positive and negative changes, (practices, systems, cultures and wellbeing) produced by a development intervention, directly or indirectly, intended or unintended,

Sustainability: whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable,

Inclusivity: the extent to which intended beneficiaries and other stakeholders have participated in aspects of evaluation design, implementation and use.

The IOCE encourages diverse yet rigorous methodological solutions to IE in order to strengthen the quality of evaluations of impact to expand buy-in and support for evaluation and for subsequent action.

The IOCE welcomes the increased attention and funding for improving IE. It champions a multi-method approach drawing from the rich diversity of existing frameworks and one that engages both the developed and developing world. We would be pleased to join with others in this endeavour.

References

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

Barca, F., (2009) An agenda for a reformed cohesion policy (EU: http://ec.europa.eu/regional_policy/policy/future/barca_en.html)

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

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

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.

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)

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