



**EES Statement: The importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions<sup>1</sup>**

*December 2007*

The European Evaluation Society (EES), consistent with its mission to promote the 'theory, practice and utilisation of high quality evaluation', notes the current interest in improving impact evaluation and assessment (IE) with respect to development and development aid. EES however deplores one perspective currently being strongly advocated: that the best or only rigorous and scientific way of doing so is through randomised controlled trials (RCTs).

In contrast, the EES supports multi-method approaches to IE and does not consider any single method such as RCTs as first choice or as the 'gold standard':

- 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.<sup>2</sup>
- IE is complex, particularly of multi-dimensional interventions such as many forms of development (e.g. capacity building, Global Budget Support, sectoral 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 Europe and other parts of the world do not favour a specific approach or group of approaches – although they may require that the evaluator give reasons for selecting a particular evaluation design or combination.

RCTs represent one possible approach for establishing impact, that may be suitable in some situations, e.g.:

- With simple interventions where a linear relationship can be established between the intervention and an expected outcome that can be clearly defined;
- Where it is possible and where it makes sense to 'control' for context and other intervening factors (e.g. where contexts are sufficiently comparable) ;
- When it can be anticipated that programmes under both experimental and control conditions can be expected to remain static (e.g. not attempt to make changes or improvements) often for a considerable period of time;
- Where it is possible and ethically appropriate to engage in randomisation and to ensure the integrity of the differences between the experimental and control conditions.

Even in these circumstances it would be 'good practice' not to rely on one method but rather combine RCTs with other methods – and to triangulate the results obtained.

As with any other method, an RCT approach also has considerable limitations that may limit its applicability and ability to contribute to policy, e.g.:

- RCT designs are acknowledged even by many of its proponents to be weak in external validity (or generalisability), as well as in identifying the actual mechanisms that may be responsible for differences in outcomes between the experimental and control situations;

**EES SECRETARIAT:**

Motivation Office Support P.O. Box 1058, 3860 BB Nijkerk, The Netherlands

T: +31 (0) 33 247 3488

F: +31 (0) 33 246 04 70

E: ees@mos-net.nl

I: www.europeanevaluation.org

- “Scaling up”, across-the-board implementation based upon the results of a limited and closely controlled pilot situation, can be appropriate for those interventions (e.g. drug trials) where the conditions of implementation would be the same as in the trial, but this is rarely the case for most socio-economic interventions where policy or programme ‘fidelity’ cannot be taken for granted;
- An RCT approach is rarely appropriate in complex situations where an outcome arises from interaction of multiple factors and interventions, and where it makes little sense to “control” for these other factors. In a development context, as for most complex policy interventions, outcomes are the result of multiple factors interacting simultaneously, rather than of a single ‘cause’;
- RCTs are limited in their ability to deal with emergent and/or unintended and unanticipated outcomes as is increasingly recognised in complexity and systems research - many positive benefits of development interventions will often be related rather than identical to those anticipated at the policy/programme design stage;
- RCTs generally are less suited than other approaches in identifying what works for whom and under what circumstances. Identifying what mechanisms lead to an identified change is particularly important given the varying contexts under which development typically takes place and is essential for making evidence-based improvements.

We also note that RCTs are based upon a successionist (sometimes referred to as ‘factual’) model of causality that neglects the links between intervention and impact and ignores other well understood scientific means of establishing causality, e.g.:

- Both the natural and social sciences (e.g. physics, astronomy, economics) recognise other forms of causality, such as generative (sometimes referred to as ‘physical’) causality that involve identifying the underlying processes that lead to a change. An important variant of generative causality is known as the *modus operandi* that involves tracing the ‘signature’, where one can trace an observable chain of events that links to the impact.
- Other forms of causality recognise simultaneous and/or alternative causal strands, e.g. acknowledging that some factors may be necessary but not sufficient to bring about a given result, or that an intervention could work through one or more causal paths. In non-linear relationships, sometimes a small additional effort can serve as a ‘tipping point’ and have a disproportionately large effect.
- Some research literature questions whether simple ‘causality’ (vs. ‘contribution’ or ‘reasonable attribution’) is always the right approach, given the complexity of factors that necessarily interact in contemporary policy – many of them in specific contexts.

EES also notes that in the context of the Paris Declaration, it is appropriate for the international evaluation community to work together in supporting the enhancement of development partner capacity to undertake IE. Mandating a specific approach could undermine the spirit of the Paris Declaration and as the literature on evaluation utilisation has demonstrated, limit buy-in and support for evaluation and for subsequent action.

In conclusion, EES welcomes the increased attention and funding for improving IE, provided that this takes 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 participating in this endeavour.

## Notes

---

<sup>1</sup> This statement was drafted on behalf of the EES Board of Directors by an ad hoc work-group consisting of: Burt Perrin, Chair and EES Secretary; Sandra Speer, EES board member; Murray Saunders, EES President (as of 1 January 2008); Nicoletta Stame, immediate Past President of EES; and Elliot Stern, EES Past President and editor of *Evaluation*, and Zenda Ofir, EES member and Past President (and current board member) of the African Evaluation Association (AfrEA). All EES members had an opportunity to comment upon an earlier draft of this statement; all comments received supported the statement, which following minor revisions was then approved by the EES Board of Directors.

<sup>2</sup> It is beyond the scope of this brief statement to include a review of the literature, or a discussion of alternative IE frameworks and methodologies. Following however are examples of some publications that consider a number of alternative approaches for establishing impact:

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.

Campbell, Donald 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.

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

Cook, Thomas 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.

Cook, Thomas 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.

Cracknell, Basil 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.

House, Ernest R. (2006). Democracy and evaluation. *Evaluation*. 12(1), 119-127.

Iverson Alex (2003). *Attribution and aid evaluation in international development: A literature review*. Ottawa: International Development Research Centre.

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

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

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

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

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

---

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

Pawson, Ray and Tilley, Nick (1997). *Realistic evaluation*. London: Sage Publications.

Perrin, Burt. (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.

Picciotto, Robert (2007). The new environment for development evaluation. *American Journal of Evaluation*, 28(4), 509-521.

Rogers, Patricia. (2008, in press). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, in press.

Scriven, Michael. (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.

Van den Berg, Rob D. (2005). Results evaluation and impact assessment in development co-operation. *Evaluation*, 11 (1), 27-36.

Williams, Robert and Imam, Iraj eds. (2007). *Systems concepts in evaluation: An expert anthology*. AEA Monograph Series. Pont Reyes, CA: EdgePress of Inverness.