

# Evaluation

<http://evi.sagepub.com/>

---

## Editorial

*Evaluation* 2013 19: 213

DOI: 10.1177/1356389013497113

The online version of this article can be found at:  
<http://evi.sagepub.com/content/19/3/213.citation>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



The Tavistock Institute

**Additional services and information for *Evaluation* can be found at:**

**Email Alerts:** <http://evi.sagepub.com/cgi/alerts>

**Subscriptions:** <http://evi.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Jul 10, 2013

[What is This?](#)



## Editorial

*Evaluation*  
19(3) 213–216  
© The Author(s) 2013  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1356389013497113  
evi.sagepub.com



This special issue of the journal is devoted to case studies: what they can and can't do; what methods are appropriate when; and what can be learned from contemporary 'state of the art' thinking not only in evaluation but from across the social sciences.

The issue builds on an international workshop that took place in Copenhagen in May 2012. It was organized by the Evaluation Department of DANIDA – Danish Ministry of Foreign Affairs; the University of Copenhagen; and the journal *Evaluation*. Six of the seven articles that follow were written by participants of the 2012 Workshop. The seventh article is an invited piece prepared by Robert Yin who has for decades been a leading advocate for strengthening the design and implementation of case studies in evaluation.

The Copenhagen workshop<sup>1</sup> was specifically focused on case studies in international development: a field where single case studies are commonly relied on for a deeper understanding of the needs of poor communities; and where multiple cross-country case studies have become the method of choice for policy-oriented evaluations. However, the challenges faced in international development are not unique. The workshop asked questions such as: How can we be confident in methods used? How far can one generalize from case studies? and How can one best synthesize and learn lessons from many different but overlapping cases? These concerns for 'validity', 'rigour' and 'generalization' are universal and we are confident that all readers of this journal with an interest in case studies will be able to connect with the ideas, methods and theories that are discussed here.

Case studies are a taken-for-granted part of the evaluator's toolkit. They can take many forms – ethnographic, comparative, quantitative and qualitative; and address many purposes – descriptive, explanatory, normative and exploratory. Probably the strongest tradition in evaluation is for qualitative and descriptive case studies that are able to unpick complex bounded systems in a holistic way, often giving a voice to stakeholders and the presumed beneficiaries of programmes or policies. Because they can 'drill down' deep, case studies have also been seen as good at explaining – addressing the 'how' and 'why' questions in an evaluation, at least in a particular context, as illustrated by several contributions to this special issue.

One reason that insights from the Copenhagen workshop are likely to be widely applicable is that policy-maker pressure for 'results', 'impact' and 'effectiveness' is now a reality across all policy domains. It is these pressures that have given urgency to questions about external validity and causal analysis. Those who accept the power of case studies to accurately represent and describe 'this case, here and now', often express doubts about whether case studies can ever provide a basis for generalization and causal inference. This special issue directly addresses these kinds of doubts.

In the last decade or so there has been a tidal wave of methodological innovation in the application of case studies across many disciplines. Some, but not all, of these innovations have begun to permeate evaluation theory and practice. This has been supported by serious re-thinking of the philosophical and theoretical assumptions that underpin a wide spectrum of research designs and

methods. Among the core ideas that have come to the fore are: the centrality of contexts – things are not the same across all contexts; the interconnectedness of networks and large systems – why we have to look at wholes as well as parts; the agency (or power) of human actors; the nature of complexity; how to take into account the influence of time given that socio-economic phenomena are rarely static; what constitutes a sound basis and logic for causal inference; the importance of theory in explanation; demonstrating the contribution of multiple causes (rather than hunting for the elusive ‘silver bullet’); how far and under what circumstances we can generalize from a single or even a few cases; and asking whether the distinction between quantitative and qualitative methods is as clear cut as has been traditionally argued. Together these preoccupations have helped reshape thinking about case studies – as with many other methods – and it is these ideas that populate the articles in this issue.

David Byrne begins with a trenchant methodological manifesto for renewing the way we think about case-study research and evaluation. For Byrne, in a complex world, causality cannot be adduced from the reductionist analysis of parts of a whole; and the whole must be seen in context. Instead of a deterministic worldview, he argues that causality is probabilistic. Byrne takes on one of the key questions that surfaced in Copenhagen: What is a case? He sees a case-study approach as necessarily ‘configurational’ or ‘set-theoretic’, rather than made up of decontextualized variables. Byrne argues for ‘an explicit rejection of the notion that it is meaningful to assign partial causal powers to any element in the configuration’. This is an understanding also consistent with ‘realist’ thinking and with QCA (Qualitative Comparative Analysis) as an approach and method that Byrne advocates. One implication of this perspective is that we need to pay as much attention to within-case analysis as to cross-case analysis. This is a repeated theme in most of the articles in this issue.

Michael Woolcock addresses many similar issues as David Byrne from the perspective of a seasoned development practitioner and researcher. He is especially concerned with ‘external validity’ – the ‘basis for generalizing any claims about impact across time, contexts, groups and scales of operation’. Woolcock reviews examples, often RCT or quasi-experimentally based, drawn from health, bio-medicine, education and criminal justice to demonstrate that external validity ‘is a fraught exercise’. Citing Nancy Cartwright and Jeremy Hardie, he argues that we need to understand what other ‘key facts’ must be in place if we want to be able to generalize. Woolcock goes on to suggest an ‘integrated framework’ made up of three elements: ‘causal density’ or complexity; ‘implementation capability’; and a ‘reasoned expectation’ based on ‘grounded theory’ about what change is achievable. Within such a framework Woolcock argues that analytic case studies have particular strengths – able to bridge the ‘qualitative/quantitative divide’; ‘elicit causal claims and generate testable hypotheses’; and ‘getting inside the black box’ to identify causal mechanisms.

How to draw lessons from existing case-study evaluations? This is the question that Julia Betts addresses in her synthesis of four major donor-led initiatives concerned with different aspects of governance: budget support; anti-corruption measures; public sector reform; and the adoption of ‘Paris Declaration’ principles. In all, these four evaluations produced 22 country-level case studies, which the OECD Development Aid Committee’s Evaluation Network sought to ‘synthesize’ as an input into possible policy reform. Julia Betts and colleagues adopted a ‘realist’ approach to synthesis aiming to answer the contextualized question: ‘where, how, under what circumstances and with what results’. The stages of the synthesis process allowed for the identification of programme theory; narratives and mechanisms of change; and policy relevance through validation and dialogue with stakeholders. The article concludes with thoughtful reflections on the strengths and limitations of ‘realist synthesis reviews’.

Barbara Befani takes forward one strand of David Byrne's methodological agenda in her advocacy of QCA as a method for 'case-based evaluation'. She focuses on the strengths of QCA both for causal inference and explanation; and for external validity and generalization. Befani reminds us of our continued reliance when drawing causal inference, on the logic of JS Mill, in particular his 'Method of Difference' and 'Method of Agreement'. She argues that one of QCA's strengths is its ability to go beyond Mill's 'procedures for attributing a single effect to a single cause'. Befani illustrates a number of key concepts from applications of QCA to specific evaluations. These include 'conjunctural causality' – where different causes combine; the importance of 'context' in its different forms – 'no single case makes the difference in all situations'; and ideas of 'necessity' and 'sufficiency' (following Mackie). Separating out 'necessity' and 'sufficiency' becomes increasingly important in complex programmes when multiple causality is the norm. Befani's discussion of the links between sample size and generalization is especially useful. She highlights that some forms of QCA are able to handle relatively large numbers of cases. But she also points out, following Cartwright and colleagues, that 'ability to generalize . . . depends on how many different "paths" to the outcome are covered in the sample: how many different contexts and causal influences are considered' as much as by sample size.

Sangeeta Mookherji and Anne LaFond are also concerned with generalizability and external validity. They reflect on these themes around the evaluation experience of African immunization programmes. Mookherji and LaFond distinguish the different assumptions made by different case-study traditions: statistical and qualitative. However, the authors see 'theories of change' as a common bridge between proponents of different case-study traditions. Mookherji and LaFond developed their theory of change to answer an overarching evaluation question: 'What are the drivers of routine immunization system performance?' The authors place considerable emphasis on case selection, focusing at country and district level on 'improving' and positive cases, although negative cases were also included. Synthesis Workshops in each country were used to initiate a 'generalization' phase that went beyond 'country-specific findings'. The authors end with a detailed reflection on their methodology and on lessons learned.

Sietze Vellema, Giel Ton, Nina de Roo and Jeroen van Wijk compare case studies of 'value chain partnerships' (VCPs) between companies, NGOs and producers in Rwanda and Uganda. The authors, like others in this issue, are firmly wedded to 'contextualized' and 'configuration-based' understandings. This article only covers the first step in a three-part analytic strategy, that of descriptive inference. Subsequent steps include developing impact hypotheses; and focusing on attribution and generalizability. A distinctive feature of this evaluation was that Vellema and colleagues were involved both as evaluation researchers and 'action researchers'. It was their action involvement – organizing policy dialogue and an 'R&D marketplace' – that allowed them to track the evolution of VCPs through time, thus refining their initial theories and hypotheses. The authors conclude that: 'Rather than attributing all the changes to the partnerships and its activities' their within and cross-case analysis 'underlined that partnerships are (at most) a contributory factor in a wider constellation of factors that generate change.'

Robert Yin, who did not attend the Copenhagen workshop, offers a personal view of the articles that derived from that workshop. One of his aims is to 'stimulate' future methodological developments. He focuses in particular on strengthening validity, generalization and lessons for future case-study methodology. In terms of validity, Yin emphasizes the importance of 'plausible rival explanations', 'triangulation' and 'logic models', while for generalization he concentrates on 'analytic generalization'. In the author's view there are three priorities for future methodological studies: paying attention to the implications of evaluation questions; examining what is meant by 'complexity' 'rather than relying on the use of the label alone'; and thinking more systematically

about how and when different case-study methods should be applied – and how they might be strengthened.

This special issue as a whole underlines the importance of new and more thoughtful reflection on case-study methods in evaluation. Some of the exciting developments that have been signposted are likely to be at the heart of methodological innovation in coming years.

It would be a worthy outcome of the Copenhagen Workshop if, even in a small way, it increased evaluators' awareness of new thinking about case studies. At the very least we hope to raise eyebrows and provoke debates among the readers of this journal and the evaluation community more widely.

### **Note**

1. See <http://um.dk/en/danida-en/results/eval/reference-documents/workshp12/> for programme, summary and presentations at the Copenhagen Workshop.

**Elliot Stern (Editor)**  
**Ole Winckler Andersen (DANIDA)**  
**Henrik Hansen (Copenhagen University)**