

Comparative Case Studies

Delwyn Goodrick

UNICEF OFFICE OF RESEARCH

The Office of Research is UNICEF's dedicated research arm. Its prime objectives are to improve international understanding of issues relating to children's rights and to help facilitate full implementation of the Convention on the Rights of the Child across the world. The Office of Research aims to set out a comprehensive framework for research and knowledge within the organization, in support of UNICEF's global programmes and policies, and works with partners to make policies for children evidence-based. Publications produced by the Office are contributions to a global debate on children and child rights issues and include a wide range of opinions.

The views expressed are those of the authors and/or editors and are published in order to stimulate further dialogue on impact evaluation methods. They do not necessarily reflect the policies or views of UNICEF.

OFFICE OF RESEARCH METHODOLOGICAL BRIEFS

UNICEF Office of Research Methodological Briefs are intended to share contemporary research practice, methods, designs, and recommendations from renowned researchers and evaluators. The primary audience is UNICEF staff who conduct, commission or interpret research and evaluation findings to make decisions about programming, policy and advocacy.

This brief has undergone an internal peer review.

The text has not been edited to official publication standards and UNICEF accepts no responsibility for errors.

Extracts from this publication may be freely reproduced with due acknowledgement. Requests to utilize larger portions or the full publication should be addressed to the Communication Unit at florence@unicef.org

To consult and download the Methodological Briefs, please visit <http://www.unicef-irc.org/KM/IE/>

For readers wishing to cite this document we suggest the following form:

Goodrick, D. (2014). Comparative Case Studies, *Methodological Briefs: Impact Evaluation 9*, UNICEF Office of Research, Florence.

Acknowledgements: This brief benefited from the guidance of many individuals. The author and the Office of Research wish to thank everyone who contributed and in particular the following:

Contributors: Patricia Rogers

Reviewers: Nikola Balvin, Roberto Benes, Joaquin Gonzales, Fiachra McAsey

© 2014 United Nations Children's Fund (UNICEF)
September 2014

UNICEF Office of Research - Innocenti
Piazza SS. Annunziata, 12
50122 Florence, Italy
Tel: (+39) 055 20 330
Fax: (+39) 055 2033 220
florence@unicef.org
www.unicef-irc.org

1. COMPARATIVE CASE STUDIES: A BRIEF DESCRIPTION

A case study is an in-depth examination, often undertaken over time, of a single case – such as a policy, programme, intervention site, implementation process or participant. Comparative case studies cover two or more cases in a way that produces more generalizable knowledge about [causal questions](#) – how and why particular programmes or policies work or fail to work.

Comparative case studies are undertaken over time and emphasize comparison within and across contexts. Comparative case studies may be selected when it is not feasible to undertake an experimental design and/or when there is a need to understand and explain how features within the context influence the success of programme or policy initiatives. This information is valuable in tailoring interventions to support the achievement of intended [outcomes](#).

Comparative case studies involve the analysis and synthesis of the similarities, differences and patterns across two or more cases that share a common focus or goal. To be able to do this well, the specific features of each case should be described in depth at the beginning of the study. The rationale for selecting the specific cases is directly linked to the [key evaluation questions](#) (KEQs) and, thus, to what needs to be investigated. An understanding of each case is important in establishing the foundation for the analytic framework that will be used in the cross-case comparison.

Comparative case studies often incorporate both qualitative and quantitative data. Given the focus on generating a good understanding of the cases and case context, methods such as [fieldwork visits](#), [observation](#), [interviews](#) and [document analysis](#) often dominate among the various data collection methods employed.

While the strategies used in data collection for single and comparative case studies are similar, comparative case studies require more extensive conceptual, analytic and synthesizing work. The synthesis across cases extends beyond the comparison of similarities and differences to using these similarities and differences to support or refute propositions as to why an intervention succeeds or fails. While some degree of comparison is at the heart of any design in which multiple cases are used, the distinguishing feature of comparative case studies is the emphasis on examining [causality](#) (i.e., the extent to which the intervention caused the results, particularly outcomes and [impacts](#)). As is the case for [experimental](#) and [quasi-experimental designs](#) (see Brief No. 7, Randomized Controlled Trials and Brief No.8, Quasi-experimental Design and Methods), comparative case studies are time and resource-intensive. This is mostly due to the inclusion of iterations between propositions, evidence collection and synthesis.

Main points

1. Comparative case studies can be used to answer questions about causal [attribution](#) and [contribution](#) when it is not feasible or desirable to create a [comparison group](#) or [control group](#).
2. Comparative case studies usually utilize both qualitative and quantitative methods.
3. Comparative case studies are particularly useful for understanding and explaining how context influences the success of an intervention and how better to tailor the intervention to the specific context to achieve intended outcomes.

2. WHEN IS IT APPROPRIATE TO USE THIS METHOD?

As a design option, comparative case studies are suitable in the following circumstances:

- When 'how' and 'why' questions are being posed about the processes or outcomes of an intervention.
- When one or more interventions are being implemented across multiple contexts, and there is little or no opportunity to manipulate or control the way in which the interventions are being implemented.
- When there is an opportunity for iterative data collection and analysis over the time frame of the intervention.
- When an understanding of the context is seen as being important in understanding the success or failure of the intervention.
- When experimental and/or quasi-experimental designs are unfeasible for practical or ethical reasons, or to supplement evidence from such evaluation designs.

As with other evaluation designs and methods, it is essential that the evaluation team undertaking comparative case studies has the necessary knowledge and skills to implement a comparative approach. Skills are required in qualitative and quantitative methods, concept development and theory testing, and synthesis, and in particular the team must be able to systematically investigate causal questions using techniques such as [qualitative comparative analysis](#) and [process tracing](#).

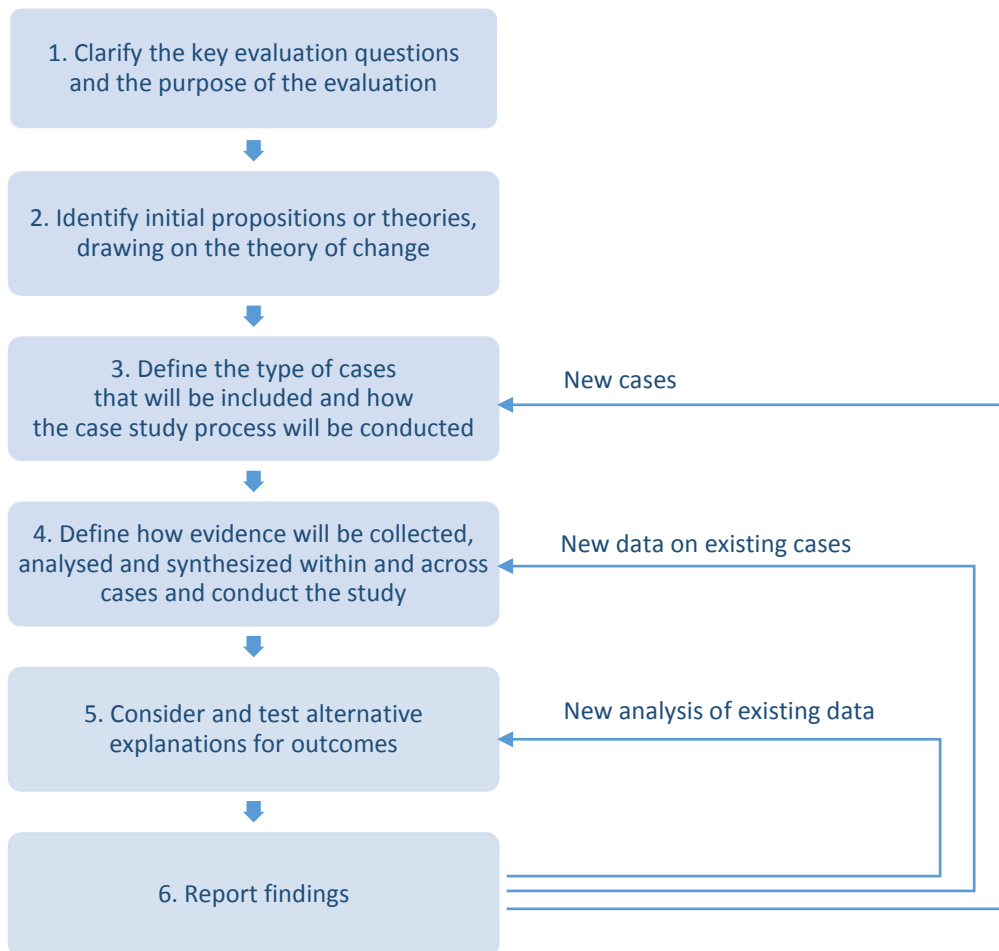
3. HOW TO CONDUCT COMPARATIVE CASE STUDIES

Comparative case studies essentially involve six steps, which should ideally be undertaken in the order shown below (see also figure 1). The programme's [theory of change](#) (see Brief No. 2, Theory of Change) and the KEQs guide the focus and selection of the cases and the dimensions that will be studied. Steps 4, 5 and 6 are likely to involve several iterations and further data retrieval or analysis.

1. Clarify the KEQs and purpose of the evaluation to determine whether the use of comparative case studies is an appropriate design.
2. Identify initial propositions or theories (see below) to focus the comparative case studies, drawing on the programme's theory of change.
3. Define the type of cases that will be included and how the case study process will be conducted.
4. Identify how evidence will be collected, analysed and synthesized within and across cases, and implement the study.
5. Consider and test alternative explanations for the outcomes.
6. Report findings.

The sequence of these steps allows for explanatory evidence to be collected and tested iteratively (a major difference that sets comparative case studies apart from experimental and quasi-experimental designs for understanding causality). The cases included in the study can be conducted concurrently or sequentially, depending on the how the programme is being implemented (e.g., linking the cases to staggered programme implementation) and the evaluation time frame and budget.

Figure 1. The logic of comparative case studies



Source: Diagram based on the discussion of case study designs in Yin, Robert K., Case study research: Design and methods, fifth edition, Sage, Los Angeles, 2014

1. Clarify the KEQs and purpose of the study to determine whether a comparative case studies design is appropriate

KEQs will guide decisions about whether or not a comparative case studies design is an appropriate evaluation design (see also Brief No. 2, Theory of Change and Brief No. 3, Evaluative Criteria). It is important that evaluation managers are clear about the purpose of using comparative case studies and define whether they are primarily interested in:

- a **description** of the similarities and differences found between cases to generate a holistic view of how the programme worked
- an **interpretation of the implications** of those similarities and differences across sites or programmes to guide further implementation, and/or

- comparisons that can identify and/or test certain **explanatory** propositions about how and why a programme worked in particular contexts.

In an [impact evaluation](#), the focus will be on explaining how the programme contributed to observed outcomes, and this will necessarily involve iterative description, interpretation and explanation. Hence, there will be a strong emphasis in an impact evaluation on synthesizing similarities and differences within and across cases and contexts.

2. Identify initial propositions drawing on the programme's theory of change

Comparative case studies are most powerful when they are informed by a good programme theory of change – an explanation of how the activities are understood to contribute to a chain of results that will produce the intended impacts (see Brief No. 2, Theory of Change). Evaluation managers and evaluators should draw on the theory of change to guide the selection of cases and to identify which properties and dimensions of the cases will be explored.

Propositions may also be generated from previous research and theoretical development around the intervention or from existing policy or programme documentation. While these 'initial' propositions (i.e., they still need to be tested) are important in identifying which cases may be of interest to explore in more depth, the evaluation must remain open to new lines of inquiry that may transpire from the iterative analysis process, which occurs as data are analysed and evidence synthesized within and across cases.

3. Define the type of cases that will be included and how the case study process will be conducted

The selection of cases requires careful consideration as the decisions made at this stage have implications for how well causality can be addressed in the data analysis and synthesis process.

It is important to base the selection of cases on clear objectives, rather than select a case just because it seems readily accessible or convenient to study. If cases are nested within other designs (e.g., an experiment), the cases may represent [intervention](#) and control groups, or be used to explain variations in outcomes. In this scenario, the cases may be more readily identifiable at the start of the evaluation study.

Some 'initial' cases may be selected up front (usually with some knowledge of the wider context), but cases may also be selected at a later stage in the study, as evidence becomes available. The evaluation manager should ensure that a clear rationale for the selection of cases is provided and also that the data collection strategy and unit of analysis (i.e., the 'case') will address the KEQs.

Given the importance of conceptual clarity in the selection of cases, it is helpful to develop a case study protocol to outline the processes of data collection and information retrieval that will occur across all selected cases. This case study protocol ensures that systematic procedures will be put in place to gather and retrieve data to aid later comparisons within and across cases.

The number of cases included is usually limited since a deep understanding of each case is needed and this requires intensive data collection and analysis. Comparative case studies may involve as few as two cases; when a large number of cases is included, trade-offs in terms of depth and detail will need to be made. In large, multi-site interventions, [purposive](#) and [stratified sampling](#) is often used to select a smaller number of cases for in-depth study, with a more limited description and comparison process conducted across an extended set of cases. As a general rule of thumb, the larger the number of cases examined, the less depth can be achieved when describing the patterns within and across cases (unless evaluation timings and funding pose no limits).

For example, consider an evaluation of the implementation of a programme across several communities, which seeks to understand how and why the programme worked or didn't work. It would be important to

identify those cases that will provide the best test of causal propositions. An initial description might show that the programme had been more successful where there was strong leadership, an existing funding base and strong community ownership. It would be premature, however, to claim that these attributes had produced these results. It would be useful to select examples that did not fit the pattern – a case that had these characteristics but which was unsuccessful, and a case that did not have these characteristics but was successful. This analysis would provide a test of the initial propositions and help to identify any other features of the context that may explain the achievement (or not) of successful outcomes.

Cases may represent different levels of social organization – macro, meso or micro. For example, an impact evaluation of a programme could compare cases at the micro level, comparing different villages where interventions have been implemented; at the meso level, comparing different approaches used for service delivery or capacity development; or at the macro level, comparing different national programmes.

The unit of analysis defines what the case represents. The selection of a unit of analysis should follow on from reflection on the KEQs. Good questions to ask include: Which group's response to the intervention is of interest? How will an understanding of this group's perspective assist in helping to understand what worked, how and why? Which dimensions of the cases require study to answer the KEQs?

Decisions about the unit of analysis have implications for resources, time frames and data collection options. Some case studies have embedded units of analysis, for example, an examination of responses of teachers (embedded cases) within the schools (overall case). The cases to compare can include individuals, programmes, groups, decisions and/or implementation processes. The specific characteristics of the case will determine the data to be collected, so gaining agreement early on in the evaluation process is important. For example, the way in which a *school district* implements a new policy reform may be quite different to how *schoolteachers* adapt and implement the policy at the school level.

The selection of units of analysis will also have implications for the staging and sequencing of data collection, analysis and synthesis steps in the case study. Some cases can be studied concurrently, with the evaluator gathering evidence from multiple cases at the same time. In other instances, for example, in the staggered implementation of a programme across sites, it makes more sense to study the cases in sequence. The latter is particularly useful for theory testing and elaboration of causal conditions necessary or sufficient for programme success, as the evaluator can select particular cases with or without these conditions, or select cases on the basis of outcome success or failure. For example, it may be that strong leadership of a community intervention is present in every site that has achieved the intended outcomes. Further examination of cases may reveal, however, that strong leadership on its own is insufficient to generate success. Rather, it is strong leadership combined with high levels of community ownership that together produce the conditions necessary to achieve outcomes.

Decisions about the staging of the cases will be informed by the nature of the programme, accessibility to information, and evaluation resources. For example, if the programme is approaching its conclusion at the time the evaluation is initiated, the evaluator will be unable to compare evidence collected over time. It may, however, be possible to draw on existing data to supplement new data collection.

4. Identify how evidence will be collected, analysed and synthesized within and across cases, and conduct the study

Comparative case studies use mixed methods, integrating qualitative and quantitative data. Whereas some mixed method evaluations use qualitative and quantitative data to answer different evaluation questions, in comparative case studies these data are analysed together with the intention of gaining an in-depth understanding of the cases. In addition, the use of mixed methods is purposefully adopted to understand and test causal propositions.

Comparison within the case and between cases is at the heart of the comparative approach. The dimensions to be compared are defined by the programme's theory of change, but other areas for comparison may emerge during the study, including:

- Comparison of key dimensions of different programmes as they operate over time – across cases and contexts.
- Comparison of anticipated outcomes, identified from programme theory or through prior research, with actual outcomes.
- Comparison of the responses of stakeholders (e.g., children or mothers/fathers/caregivers) over time within a single programme.
- Comparison of the responses of stakeholders (e.g., children or mothers/fathers/caregivers) to a programme over time in different contexts.

Cross-case analyses will likely involve some type of pattern matching logic.¹ Pattern matching involves comparing two or more patterns between the cases to see if they are similar or different as a step in explaining observed processes or behaviours.

Comparison spurs a range of new questions about the nature of the similarities and differences between different cases. These may include:

- What are the key patterns that occur over time in each of the cases?
- What might be responsible for these patterns?
- What is surprising about these patterns?
- How can these similar or different patterns be explained?

5. Consider and test alternative explanations for the outcomes

The explanatory phase of comparative case studies identifies new propositions derived from the observed patterns and relationships, and seeks out additional evidence to elaborate or test these propositions. For example, if an analysis of the cases reveals that the success of a UNICEF intervention seems to be related to the degree of collaboration with the local community, it may be useful to seek out a case in which collaboration resulted in less successful outcomes as a test of that proposition. Even one case can be a critical test of a proposition.

A single case can also be the source of new insights about critical success factors. For example, it may become apparent that one immunization programme has succeeded in one context but failed in others. Therefore, this case is selected and the attributes and features that may be responsible for the differences explored in more depth. Rather than settle with merely a description of the attributes, an explanation is built that can be tested through [triangulation](#) with other data sources. Triangulation is often used in evaluation to check answers to [descriptive questions](#) about the way things are, for example, by confirming interview statements with direct observations. In comparative case studies, triangulation can also be used to check and strengthen answers to causal questions, for example, by identifying and ruling out alternative explanations, or by identifying and explaining exceptions to the main pattern observed.

A further example serves to illustrate the iterations between case selection, proposition building and proposition testing. There may be an interest in examining the properties (e.g., leadership, community ownership, funding base) that seem most strongly associated with the successful implementation of a programme. Examination of the cases reveals that the programme was more likely to be successful in

¹ Yin, Robert K., *Case study research: Design and methods*, fifth edition, Sage, Los Angeles, 2014.

those communities with the highest level of community ownership. The level of community ownership (weak to strong) and also the strength of leadership (low to high) are dimensions of interest in examining the case. The goal is to understand and explain the properties, dimensions and outcomes of a specific case or cases and the ways in which these are influenced by the context. In this example, it may be that several factors are associated with successful programmes, but it is the examination of the dimensions (level, extent or strength) that provides insights into the critical elements of successful programmes. This understanding supports causal.

One promising analytic approach that supports causal attribution is qualitative comparative analysis. This approach offers strategies that enable rigorous examination of the categorical attributes (e.g., leadership, level of community ownership, community stability) that may be associated with particular programme outcomes (success or failure). Like other forms of qualitative analysis, it requires an in-depth understanding of the cases, strong theory and critical thinking, *and* testing of propositions. While qualitative comparative analysis can draw on both qualitative and quantitative data, the analytic approach must focus on the relationships among combinations of potential causal conditions within and across cases (i.e., case-based knowledge). In contrast, traditional forms of analysis commonly focus on specific variables and their average effects (i.e., variable-oriented knowledge).²

Qualitative comparative analysis provides a formal and clear procedure that many other approaches to qualitative analysis cannot offer. It may appeal to evaluators wishing to provide a more compelling evidence base for programme assessments and judgements. Other techniques that can be used are: process tracing (i.e., a process of backtracking from propositions to potential causal pathways to the outcome) or [multiple regression analyses](#). Methods of analysis tend to have an emphasis on *either* generating an explanation of specific outcomes in cases (i.e., comparative emphasis within and across cases) *or* on causal effects observed across a larger number of cases (i.e., traditional statistical emphasis on causal effects). There are few examples where evaluators focus on synthesizing evidence to achieve both goals.³ Therefore, evaluation managers need to understand the implications of the analytic approach and ensure that the methods adopted address the information needs of the intended users. Comparative studies tell us why and how particular outcomes were produced; statistical approaches to causal attribution tell us about the effects of particular variables in a larger number of cases, but may not pay attention to issues within or across the context of the cases themselves.

6. Report findings

While this step is included as the final step in how to conduct comparative case studies, it is useful to discuss possible report formats with the evaluators earlier in the study process. The overall structure should ideally be framed around KEQs, but the balance of description, interpretation and explanation may differ across evaluations. Evaluators should be encouraged to share early drafts with evaluation managers, both to gain feedback on writing style and presentation of evidence and to enable the consideration of additional comparative questions.

It is generally difficult to make sense of long, unwieldy narratives that describe similarities and differences within and between cases. What is often more helpful is the development of summary tables and diagrams that capture the key dimensions of the cases. Such displays can depict raw data, data organized by themes, or theoretical insights generated through the study. Typologies, tables, diagrams or matrices can capture and summarize the information collected within and across cases and facilitate the examination of

² Ragin, Charles C., and Garrett A. Schneider, 'Case oriented theory building and theory testing', in M. Williams and W.P. Vogt (eds.), *The Sage Handbook of Innovation in Social Research Methods*, Sage, Los Angeles, 2011, pp. 150–166.

³ Mahoney, James, and P. Larkin Terrie, 'The proper relationship of comparative-historical analysis to statistical analysis: Subordination, integration, or separation?', in D.S. Byrne and C.C. Ragin (eds.), *The Sage Handbook of Case-based Methods*, Sage, Los Angeles, 2009, pp. 410–423.

similarities and differences. Rows in tables, for example, can include properties and dimensions of the cases that are hypothesized to contribute to a particular outcome.

The judicious use of summary diagrams in reports and publications can enhance readers' understanding of the richness of the information collected and enable them to quickly track how interpretations and causal propositions were elaborated.

Those commissioning, managing or reviewing comparative case studies need to be sensitive to the different ways in which generalization is understood in this design. There is often an emphasis on the transferability of the causal propositions to other contexts rather than on generalizing from one case to a wider set of cases. This is an important difference in the scope of generalization, which diverges from the traditional focus on generalizing from a sample to the population to instead have an emphasis on generalizing propositions about the characteristics that are predictive of success and failure.

4. ETHICAL ISSUES AND PRACTICAL LIMITATIONS

Ethical issues

A range of ethical issues must be addressed when undertaking comparative case studies. The most pressing issue is that the level of description required to portray the richness of the cases may mean that the cases, and the participants within the cases, are identifiable. This is not necessarily a problem, but it must be clearly discussed and negotiated with those participating. Details that are not essential to an understanding of the case (e.g., level of school education, income) may be modified and an explanatory note included in the report to indicate this has been done to protect participant identities.

Practical limitations

Comparative case studies require a range of skills and expertise. As with other designs it is important to assess the match between the required skill set and the evaluation team. Firstly, if the comparative case studies are to include quantitative and qualitative evidence, the evaluator or evaluation team must have the requisite skills in both methodologies. Secondly, the evaluator/team must possess strong synthesis skills and the capacity to integrate convergent and divergent evidence. Thirdly, the evaluator/team must be able to embrace the complexities of each case and to employ critical reasoning in making sense of the evidence and presenting a coherent argument. Good synthesis relies on good description and analysis; the evaluator/team requires skills in constructing propositions that build on these elements.

Comparative case studies have disadvantages in some contexts. One key issue is that they are often highly resource-intensive, particularly if extensive fieldwork is required. Depending on the purpose of a particular study, it may be better to purposively select a small number of cases or cluster cases of a particular type if the study budget prohibits more extended or in-depth study of a larger number of cases. Comparative case studies can be based entirely on secondary data analysis, and thus require no fieldwork or primary data collection at all. The quality of available evidence must be suitably strong for this to be an appropriate option, however.

Findings can be less reliable if there is too much of a time lag between cases in the comparison activity. The time lag may make comparisons across cases problematic due to the likely influence of other historical, social and/or programmatic factors.

5. WHICH OTHER METHODS WORK WELL WITH THIS ONE?

A comparative case studies design may be a stand-alone design or it may be nested within other designs. If nested, the comparative case studies may be used to examine a specific element or the context in more detail. For example, within an experimental design that aims to ascertain the effectiveness of an intervention, the comparative case studies may be used to explain why and how the intervention did or did not work across the intervention and control group.

The nesting of comparative case studies within other designs generates evidence about how the context has influenced patterns of outcomes. Where causal mechanisms have already been identified or proposed (e.g., from an initial programme theory), comparative case studies may test the scope, applicability and transferability of those mechanisms in other contexts.

If comparative case studies make up a single component of a larger study, it is likely that the cases will be supplementary to the core design. Cases may be used to illustrate similarities and differences in outcomes across the contexts in which a programme or policy was implemented. In this scenario, there may be a focus on the description of similarities and differences rather than on the development and testing of causal propositions.

In terms of data collection methods, comparative case studies are likely to use a combination of qualitative and quantitative data, including data from project documentation, performance measures, surveys, interviews and observation.

Comparative case studies might include two particular data analysis methods: qualitative comparative analysis and process tracing. Qualitative comparative analysis⁴ documents the configuration of conditions associated with each case, usually in the form of a 'truth table'. The analysis identifies the simplest set of conditions that can account for all of the observed outcomes. The method can be used when there is a single cause for an outcome and also for more complicated causal relationships, for example, where there are multiple ways of achieving an outcome. Process tracing focuses on the use of clues within a case to adjudicate between possible explanations. It uses a series of different causal tests to see if the outcomes are consistent with the theory of change and whether alternative explanations can be ruled out.

6. PRESENTATION OF RESULTS AND ANALYSIS

There are no set rules or defined requirements for the presentation of comparative case studies. Negotiation and agreement on the format and presentation of the cases should occur early on in the process, however.

Case study reports often have the advantage of being more accessible to a broader range of audiences than other types of evaluation reports. One of the dilemmas in crafting an account is what to include and what to leave out; choices about what will be communicated need to be made.

The evaluator will be synthesizing the evidence gathered in the report, so it's important to ensure that they provide sufficient information to communicate succinctly but in a transparent manner (see also Brief No. 4, Evaluative Reasoning). A good report should include examples from the data to provide evidence for claims made. Given the diverse sources and methods of data collection that typically characterize a comparative case studies approach, a technical appendix may be used to supplement the more succinct formal report.

⁴ Davies, Rick, 'Qualitative Comparative Analysis', web page, BetterEvaluation, 2013. See http://betterevaluation.org/evaluation-options/qualitative_comparative_analysis.

If the comparative case studies are part of a larger, mixed methods evaluation, or if they are nested within another design, the format and presentation of the report should reflect the intended purpose of the comparative case studies.

A comparative case studies report should, at minimum, include:

- KEQs and audiences/users
- rationale for the selection of a comparative case studies design and its value in addressing the KEQs
- discussion of case selection – a priori (generated from programme theory or programme selection) or generated over time
- type of cases examined and overall strategy – parallel or sequential cases
- description of how propositions were developed, elaborated or tested within the comparative case studies design
- case study protocol (if the comparative case studies included a fieldwork component)
- methods of data collection, data management and analysis (the reader may be directed to a technical appendix for more detail)
- brief discussion of within case and across case analysis (additional detail may be included in a technical appendix)
- limitations of the study – difficulties encountered and shortcomings of data collection and/or analysis that may have influenced the study findings.

7. EXAMPLE OF GOOD PRACTICES

The Knowledge to Policy study⁵ looked at the effects of research in international development, exploring what kind of research is most effective in informing decisions in policy councils, and the conditions under which it is most effective. The study selected multiple cases representing a series of research studies that were designated as influential in some way to decision making around policy.

Although cases were selected to represent a good cross section of research studies funded by the host agency, the International Development Research Centre (IDRC), the cases were selected as a 'positive' sample – i.e., cases where the researchers involved claimed that the study had influenced policy. The rationale was that examination of the factors that supported influence across a relatively small number of cases would shed light on how influence was produced. The identification of commonalities across a range of contexts and content areas was seen as reinforcing the influence of these factors on decision making. At the beginning of the study, 16 cases were selected, but over time key stakeholders identified additional cases as influential. A final sample of 23 cases across a range of countries and content areas (such as poverty, water management, trade and finance, and resource management) was selected for in-depth examination and comparative analysis (with most of the cases conducted concurrently).

A range of domains was identified to examine each case. For example, the study authors explored how the research studies were initiated, described the politics inherent in the context at the time, reviewed the proposal process, and examined the conduct, findings and dissemination of the research results. The study drew on both qualitative and quantitative data collection, with an emphasis on qualitative interviews and consultation approaches. Fieldworkers adopted a common data collection instrument or protocol to collect consistent data across all cases. Matrices and [typologies](#) were used to illustrate dimensions of the cases

⁵ Carden, Fred, *Knowledge to Policy, Making the Most of Development Research*, IDRC/Sage, Ottawa, 2009. See <http://www.idrc.ca/EN/Resources/Publications/Pages/IDRCBookDetails.aspx?PublicationID=70>.

that were compared and contrasted. Regional analysis workshops were conducted to build stakeholder ownership of the process and to generate insights into the features that facilitated and/or inhibited the influence of research on policy.

The findings from each of the 23 cases were described according to the common domains of interest, with learnings presented at the end of each case description. The rich, descriptive narrative indicated a thorough understanding of the cases. This is particularly impressive given that, in several instances, the fieldworkers had to rely on the recollections and reflections of the project participants. The study authors triangulated different information sources to address the inherent weaknesses of retrospective memory and self-report. The approach to analysis was iterative; the initial data collection and analysis informed subsequent phases of data collection and analysis.

The report evidences description, interpretation and judgement about the factors that support research influence in a development context. The study authors do not claim that there were single causes for research influence, but instead identified multiple, intersecting factors that support policy influence. The rigour of the methodological approach, and the level of engagement and ownership of stakeholders in both the design and analysis stages of the comparative case studies were identified by the study authors as contributors to organizational (policy) changes within IDRC (the funding agent).

The study paid careful attention to case selection and ensured iterative processes of data collection and analysis, guided by a consistent data collection protocol that aided comparability. While the study could have been strengthened by a greater focus on searching for alternative explanations about research influence, the examples and the robust data generated through the iterative process increased confidence in the findings. The cases were positive examples where research influence was apparent (as identified by the researchers). A useful extension of this study, to strengthen causal attributions, would be to select cases where the causal features were present but where a positive outcome (influence) did not occur.

8. EXAMPLES OF CHALLENGES

A common failing in the use of comparative case studies in impact evaluation is that they describe multiple cases but do not systematically and transparently identify and test causal propositions. Another weakness in some studies that use the label of comparative case studies lies in the selection of cases. The cases are not selected in terms of how well they will be able to inform the causal analysis but in terms of providing a range of examples across various dimensions, for example, covering different regions, and different sized countries. The cases are analysed descriptively but no causal analysis is undertaken, nor are any causal propositions rigorously tested. In part, this weakness seems to be associated with a lack of clarity about the difference between case studies that are described and compared and comparative case studies that aim to address causal attribution. While causality is complex and non-linear, it is the aim of comparative case studies to address the messy complexities of real world programmes.

For example, one UNICEF impact evaluation involved multiple cases of service delivery in two countries. Three specific models of support were compared and contrasted according to strengths and weaknesses. Each model of support was treated as a case. While the study included some useful summary tables/matrices, it did not provide a coherent and theoretically informed description and analysis of the context of each site. Data were presented from the cases, but there was little explanatory synthesis of the evidence across cases. Instead the focus remained on the description of similarities and differences.

9. KEY READINGS AND LINKS

Balbach, Edith D., 'Using case studies to do program evaluation', California Department of Health Services, 1999. See : <http://www.case.edu/affil/healthpromotion/ProgramEvaluation.pdf>.

Bazeley, Patricia, *Qualitative Data Analysis: Practical Strategies*, Sage, Los Angeles, 2013.

Byrne, David S., and Charles C. Ragin, (eds.), *The Sage Handbook of Case-based Methods*, Sage, Los Angeles, 2009.

Campbell, Shelagh, 'Comparative case study', entry in A.J. Mills et al. (eds.), *Encyclopedia of Case Study Research*, 2010, pp. 174–175.

Koenig, G. K., 'Realistic evaluation and case studies: Stretching the potential', *Evaluation*, 15(1), 2009, pp. 9–30.

Miles, Matthew B., et al., *Qualitative Data Analysis: A Methods Sourcebook*, Sage, Los Angeles, 2014.

Ragin, Charles C., and Garrett A. Schneider, 'Case-oriented theory building and theory testing', in M. Williams and W.P. Vogt (eds.), *The Sage Handbook of Innovation in Social Research Methods*, Sage, Los Angeles, 2011, pp. 150–166.

Simons, H., 'The paradox of case study', *Cambridge Journal of Education*, 26(2), 1996, pp. 225–240.

Stake, Robert E., *The Art of Case Study Research*, Sage, Thousand Oaks, 1995.

Stake, Robert E., *Multiple Case Study Analysis*, Guilford Press, New York, 2006.

Stufflebeam, Daniel L., and Anthony J. Shinkfield, *Evaluation Theory, Models, and Applications*, Jossey-Bass, San Francisco, 2007.

United States General Accounting Office, Program Evaluation and Methodology Division, 'Case study evaluations', GAO/PEMD-31-10.1.9, US Government, Washington, D.C., 1990.

Vogt, W. Paul, et al., 'Innovations in program evaluation: Comparative case studies as an alternative to RCTs', in M. Williams and W.P. Vogt (eds.), *The Sage Handbook of Innovation in Social Research Methods*, Sage, Los Angeles, 2011.

Yin, Robert K., *Case Study Research: Design and Methods*, fifth edition, Sage, Los Angeles, 2014.

Yin, Robert K. and Darnella Davis, 'Adding new dimension to case study evaluations: The case of evaluating comprehensive reforms', *New Directions for Evaluation*, 113, 2007, pp. 75–93.

Yin, Robert K., *The Case Study Anthology*, Sage, Thousand Oaks, 2004.

GLOSSARY

<u>Attribution (causal)</u>	<i>Ascription of a causal link between observed (or expected to be observed) changes and a specific intervention. (OECD-DAC definition, 2010)</i>
<u>Document analysis</u>	<i>Systematic analysis of the content of relevant documents for research and evaluation purposes. See: content analysis.</i>
<u>Causal question</u>	<i>Causal questions ask whether or not, and to what extent, observed changes are due to the intervention being evaluated rather than to other factors, including other programmes and/or policies.</i>
<u>Causality</u>	<i>The principle that one variable (X) produces change in another variable (Y). It is based on the assumption that events occur in a predictable, non-random way, and that one event leads to, or causes, another. To establish causation, the two variables must be associated or correlated with each other; the first variable (X) must precede the second variable (Y) in time and space; and alternative, non-causal explanations for the relationship (such as spurious ones) must be eliminated. Events in the physical and social worlds are generally too complex to be explained by any single factor. For this reason, scientists are guided by the principle of multiple causation, which states that one event occurs as a result of several factors operating or occurring in combination.</i>
<u>Comparison group</u>	<i>In a quasi-experimental research design, this is the group of research participants/subjects that, for the sake of comparison, does not receive the treatment or intervention given to the intervention group. Comparison group subjects are typically not randomly assigned to their condition, as would be true of control group subjects in an experimental design study. See: control group, intervention group.</i>
<u>Contribution</u>	<i>One of the causes of observed change. Helps to answer whether the intervention (programme of policy) helped to cause (contributed to) the observed changes. However, in some cases ‘causal attribution’ and ‘casual contribution’ are used interchangeable (see Brief 6 for details). See: attribution.</i>
<u>Control group</u>	<i>Participants in a research study/evaluation who do not receive the experimental treatment/intervention.</i>
<u>Descriptive questions</u>	<i>Descriptive questions ask how things are and what has happened, including describing the initial situation and how it has changed, the activities of the intervention and other related programmes or policies, the context in terms of participant characteristics, and the implementation environment.</i>
<u>Experimental design</u>	<i>A research or evaluation design with two or more randomly selected groups (an experimental group and control group) in which the researcher controls or introduces an intervention (such as a new programme or policy) and measures its impact on the dependent variable at least two times (pre- and post-test measurements). In particular RCTs – which originated in clinical settings and are known as the ‘gold standard’ of medical and health research – are often used for addressing evaluative research questions, which seek to assess the effectiveness of programmatic and policy interventions in developmental settings.</i>

<u>Fieldwork visit</u>	<i>A research method which involves a visit to the physical site where the intervention is taking place in order to assess and understand the case and case context. Evaluators/researchers observe but do not manipulate the study subjects. Related: field research.</i>
<u>Impact</u>	<i>Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended. (OECD-DAC definition, 2010)</i>
<u>Impact evaluation</u>	<i>An evaluation that provides information about the impacts produced by an intervention. It can be undertaken of a programme or a policy, or upstream work – such as capacity building, policy advocacy and support for an enabling environment. It goes beyond looking only at goals and objectives to also examine unintended impacts. See: impact.</i>
<u>Intervention group</u>	<i>Also called experimental or treatment group, this is a group of research participants that receives some form of treatment or intervention, i.e. they are exposed to the independent variable. See: comparison group, control group.</i>
<u>Interview</u>	<i>A method of data collection where the interviewer asks the interviewee questions. This methodology is common in qualitative research.</i>
<u>Key evaluation questions (KEQs)</u>	<i>High-level (macro level) evaluation questions about overall performance, which the evaluation should aim to answer. KEQs are derived from the purpose of the evaluation.</i>
<u>Multiple regression analysis</u>	<i>A statistical procedure for understanding the effects of two or more independent variables on a dependent variable.</i>
<u>Observation</u>	<i>Watching something and taking note of anything it does/anything that happens. Participant behaviour may be recorded in quantitative (e.g., real-time coded categories), qualitative (e.g., note-taking for case study) formats, or by special media (e.g., audio or video recordings).</i>
<u>Outcome</u>	<i>The likely or achieved short-term and medium-term effects of a programme or policy's outputs, such as a change in vaccination levels or key behaviours.</i>
<u>Process tracing</u>	<i>A case-based approach to causal inference which involves developing alternative hypotheses and then gathering evidence (clues) within a case to determine whether or not these are compatible with the available hypotheses.</i>
<u>Purposive sampling</u>	<i>A form of non-probability sampling in which units/participants are selected by the researcher based on one or more predetermined characteristics. It involves the examination of information-rich cases from a given population to make analytical inferences about the population. The sample size can be as small as one.</i>
<u>Qualitative comparative analysis</u>	<i>Involves comparing the configurations of different case studies to identify the components that appear to be most responsible for producing specific outcomes.</i>
<u>Quasi-experimental design</u>	<i>A research/evaluation design in which participants are not randomly assigned to treatment conditions, but in which comparison groups are constructed by statistical means. It differs from the (classic) controlled experiment by not having random assignment of the treatment/intervention.</i>

<u>Stratified sampling</u>	<i>A probabilistic sampling method used when sub-populations vary considerably. The researcher divides the entire population into different sub-groups (strata) according to categories relevant to the research interest and then randomly selects the final subjects/participants proportionally from the different strata.</i>
<u>Theory of change</u>	<i>Explains how activities are understood to produce a series of results that contribute to achieving the final intended impacts. It can be developed for any level of intervention – an event, a project, a programme, a policy, a strategy or an organization.</i>
<u>Triangulation</u>	<i>The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment. Note: by combining multiple data sources, methods, analyses or theories, evaluators seek to overcome the bias that comes from single informants, methods, observers or single theory studies. (OECD-DAC definition, 2010)</i>
<u>Typology</u>	<i>A classification scheme; an examination of types. Merton's 'Modes of Individual Adaptation' and Sorokin's 'Sensate and Ideational Culture' are well-known typologies.</i>