

# Qualitative Comparative Analysis – A Rigorous Qualitative Method for Assessing Impact



Coffey How To

June 2015

## Summary

- QCA is a case based method which allows evaluators to identify different combinations of factors that are critical to a given outcome, in a given context. This allows for a more nuanced understanding of how different combinations of factors can lead to success, and the influence context can have on success.
- QCA allows evaluators to test theories of change and answer the question 'what works best, why and under what circumstances' in a way that emerges directly from the empirical analysis, that can be replicated by other researchers, and is generalizable to other contexts.
- While it isn't appropriate for use in all circumstances and has limitations, QCA also has certain unique strengths – including qualitatively assessing impact and identifying multiple pathways to achieving change which make it a valuable addition to the evaluation toolkit.

## Authors

Carrie Baptist and Barbara Befani

## What is Qualitative Comparative Analysis (QCA) and what can it do?

As evaluators, we are often asked not only to assess whether a project or intervention has worked, but why, how and under what circumstances. This contextual understanding is critical for understanding not only whether a project has been successful in a particular context, but also whether and under what conditions it can be scaled up or replicated.

The use of case-based approaches, like case studies, is a well-established method for evaluators and researchers who wish to draw out narratives and investigate a theme in-depth, probing not only what has occurred but why, how and in what context. However, these types of approaches also have weaknesses; while they provide great depth in a particular outcome in a particular context, it's unclear how generalizable that learning is to other projects or contexts. In addition, if not carried out systematically and transparently, findings from case studies are sometimes considered a 'weaker' type of evidence, as they are not replicable by external researchers (whereas quantitative analysis generates results which – in many cases – can be consistently replicated externally).

**Qualitative Comparative Analysis (QCA) is a case-based method that enables evaluators to systematically compare cases, identifying key factors which are responsible for the success of an intervention.**

As a comparative method, QCA doesn't work with a single case – it needs to compare factors at work across a number of cases in order to tease out which factors are most important for a given outcome. However, when done correctly the findings generated using QCA are generalizable: insights from one context or project can be transferred to another. QCA is also a rigorous method in that – when the analysis process is made transparent – it can be replicated by anyone and produce the same findings.

To put it a bit more technically, QCA is potentially strong on both external and internal validity while still being a qualitative method: it works with social concepts, constructs and narratives.

In this way, QCA seeks to combine the strengths of qualitative and quantitative research, by linking theory and evidence while also providing increased measurement precision<sup>1</sup>.

Additionally, while statistical methods may be more appropriate when large datasets are available, QCA does not need a high number of cases in order for the findings to be significant<sup>2</sup>.

## Testing Theories of Change and engaging with complexity

QCA provides an innovative way to test project or programme theories of change. A detailed theory of change will identify not only inputs, outputs, outcomes and impact, but also key assumptions and external dynamics which may impact success. QCA allows evaluators to test the relationships between key assumptions and external factors identified in a theory of change, and a given outcome or impact – across a set of cases. QCA will identify which, of the factors tested, are necessary or sufficient to obtain a successful outcome or impact.

Generally, a single factor isn't necessary or sufficient for success – rather a combination of factors (also known as a 'package') is needed. QCA can identify the package(s) of factors necessary or sufficient for a successful outcome, and which factors within those combinations are most critical – in which contexts<sup>3</sup>.

In other words, QCA can identify different combinations of factors that are critical to a given outcome, in a given context<sup>4</sup>. This allows for a more nuanced understanding of how different combinations of factors can lead to success, and the influence context can have on success.

It also allows evaluators to move away from a linear understanding of how change happens and incorporate multiple paths and combinations of 'ingredients', which are different but equally relevant for a successful outcome.

1. See "A Critique of the Fuzzy-set Methods in Comparative Social Policy" by Seungyoon Sophia Lee, Oxford 2008. <http://www.compassss.org/wpseries/Lee2008.pdf>

2. Although QCA is a qualitative method, both qualitative and quantitative data can be used as part of the analysis.

3. These factors are also known as "INUS causes" (see Befani 2013) or "contributory causes" (Mayne 2012)

4. These circumstances can be described by contextual conditions or by modes of implementation: for example, different levels of political stability, or different approaches to policy influence adopted by different local partners.

This contextual, 'complex' approach seems fitting for development interventions, which usually work differently in different contexts and geographical areas, and for which there is unlikely to be a single 'silver bullet' which guarantees success.

In this sense, QCA usefully refocuses attention back onto context, systems and institutions, and away from essentialised or contextless 'best practices'.

Put simply, QCA allows evaluators to answer the question "what works best, why and under what circumstances" quite literally, in a way that emerges directly from the empirical analysis; that can be replicated by other researchers, and is generalizable to other contexts.

## How has it been used to-date?

QCA was first developed by an American social scientist, Charles Ragin, in 1987<sup>5</sup>. To date, it has been used most extensively in the academic sphere, primarily in comparative sociology and politics, but more recently it has been applied to policy analysis<sup>6</sup> and evaluation. After a few isolated attempts (see for example Befani et al. 2007), evaluators have started using QCA more often over the past 2 -3 years, including in meta-evaluations of approaches to evaluating interventions targeting violence against women and girls<sup>7</sup>, evaluations of regional governance, empowerment and accountability programming<sup>8</sup> and evaluations of public health interventions<sup>9</sup>.

However, the use of QCA in evaluations is still new, and the advantages and pitfalls of this method for evaluators are still being established. While the potential of QCA is starting to become increasingly clear in the evaluation community, there is still no comprehensive and well-established guide on how best to use QCA for evaluation, or the strategies evaluators can employ to meet the challenges posed by QCA. While it is beyond the scope of this brief to provide

5. Ragin, Charles C. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press 1987.

6. Kaisa Korhonen-Kurki, Jenniver Sehring, Maria Brockhaus & Monica Di Gregorio (2014). Enabling factors for establishing REDD+ in a context of weak governance, *Climate Policy*, 14:2, 167-186, DOI: 10.1080/14693062.2014.852022

7. Raab, M.; Stuppert, W. Review of evaluation approaches and methods for interventions related to violence against women and girls. (2014) [http://r4d.dfid.gov.uk/pdf/outputs/misc\\_gov/61259-Raab\\_Stuppert\\_Report\\_VAWG\\_Evaluations\\_Review\\_Dfid\\_20140626.pdf](http://r4d.dfid.gov.uk/pdf/outputs/misc_gov/61259-Raab_Stuppert_Report_VAWG_Evaluations_Review_Dfid_20140626.pdf)

8. Evaluation of the Africa Regional Empowerment and Accountability Programme, 2014.

9. Applying Qualitative Comparative Analysis (QCA) in the evaluation of complex public health interventions: A case study of a health improvement service for long-term incapacity Benefit recipients. *Journal of Public Health* 36(1): 126-133, 2013.

that kind of comprehensive guide<sup>10</sup>, we can offer lessons from our recent experience using QCA to evaluate DFID's

Africa Regional Empowerment and Accountability Programme (see box below).

## The added value of QCA in the evaluation of DFID's Africa Regional Empowerment and Accountability Programme

### Context

The Africa Regional Empowerment and Accountability Programme (AREAP) is a four-year, DFID-funded £19.5m project that aims to improve access to high quality relevant data, evidence and analysis for African citizens, non-state actors and policy makers. It was implemented across 36 African countries by three Africa-based organisations, the State of the Union Coalition, Afrobarometer and the Southern Africa Trust. Like many governance, empowerment and accountability programmes it was not possible to establish a counterfactual for this programme, and a qualitative or mixed-methods approach was felt to be more appropriate than a purely quantitative one. On this basis, Coffey was contracted by DFID to provide a theory-based, independent evaluation of this programme.

The evaluation primarily consisted of a structured contribution analysis and process evaluation, however QCA was carried out as part of the mid-term evaluation of this programme, to triangulate key findings from the contribution analysis, assess impact, and add an additional level of rigour and external validity to the evaluation. QCA allowed the findings based on the fieldwork in the sampled countries to be extended to the rest of the 36 countries in which the programme operates. For this evaluation, qualitative data collection was conducted in 8 sample countries, across the three implementing organisations which then made up the cases analysed using QCA.

### Evaluation Design and Data Collection

The programme and projects had pre-existing theories of change, however they were not detailed enough to support a rigorous theory-based evaluation, or to support QCA. To address this deficit, the evaluation team led theory of change workshops with each implementing organisation, to draw out additional detail and specificity. These, particularly detailed, theories of change became the basis for the selection of factors to be analysed with QCA. For this evaluation, cases were defined by the work of a particular implementing organisation, in a particular sample country. High quality qualitative data was then collected through fieldwork as part of the contribution analysis process across all cases, as having relevant comparable data across the cases is a requirement for QCA. However, even though the team collected detailed qualitative data in each of the 8 sample countries, the evaluation still faced difficulties in identifying some types of national level secondary data which would be available for all of the 8 countries while also being relevant to the programme. This forced the evaluation team to reduce the number of factors considered for the QCA, as reliable data was not available for all of the relevant conditions that emerged from the theories of change workshops.

### Findings Enabled by the QCA

#### Testing project theories of change

As part of AREAP theory of change, the implementing partners put forward claims that a number of factors were all necessary to achieve specific outcomes (and articulated this as part of the causal chains within their theory of change). Using QCA, the evaluation team was able to test these claims, and ended up emphasising a smaller number of factors than those put forward by the implementing organisations as being important for success. Some of these factors were discovered to be associated with success only in some cases but not all (and thus were declared not necessary for success).

Additionally, the evaluation team was able to drill down and specifically test individual linkages in project theories of change, between outputs, intermediate and overall outcomes.

10. An extensive guide is in preparation at <http://eba.se/en/evaluating-development-interventions-with-qca-potential-and-pitfalls/>

In these cases, the evaluation team specifically tested the relationship between intermediate outcomes, and the overall outcome in the project theory of change, to see if it was really necessary to achieve one in order to achieve the other.

The intermediate outcome was not found to be necessary, suggesting that there are other routes to success, and that the project may need to re-examine their theory of change.

As part of the QCA analysis, the team was also able to show which factors made the difference for success, in which contexts and under what circumstances. For example, the ability of implementing organisations to engage with a variety of diverse audiences (donors, government, private sector) was shown to be decisive for civil society coalitions' success in focusing on key areas of policy change only in national contexts considered unfavourable for state openness and accountability. This ability to engage with diverse audiences was not found to be so relevant in national contexts which are more open to civil society engagement and where the state is more often held to account. .

### Triangulating other approaches and adding nuance to key findings

For the evaluation of AREAP, QCA was used in conjunction with contribution analysis, where contribution analysis formed the basis for the bulk of the evaluation findings, and QCA was used to triangulate and expand on a handful of findings, around relevance and effectiveness. The findings from QCA generally confirmed those emerging from the contribution analysis, but in a few cases they provided additional emphasis on a particular aspect.

For example, one part of the evaluation touched on the role of 'champions' in government, individuals that an organisation has established a relationship with, who can help orient them in government and champion shared policy goals. The presence of a champion in government emerged from the contribution analysis as an important contributor to success, but in a generic sense without greater contextual clarity. The QCA findings allowed the team to understand the role of champions more specifically than what had emerged from the Theories of Change workshops. From the QCA, champions appeared necessary to enact policy change, but not so when the aim was to "merely focus" on advocating key policy issues. Moreover, champions were found to be largely necessary for strengthening the capacity to engage with policy makers.

In conclusion, while the AREAP evaluation primarily made use of contribution analysis to test the implementing partner theories of change. QCA provided valuable triangulation of findings, validating key conclusions from the contribution analysis, as well as refining some of the emerging hypotheses, which increased the robustness and specificity of those findings. QCA helped provide additional clarity when testing the project theories of change, strengthening the quality and rigour of the evaluation.

## What does QCA require?

As the use of QCA is relatively new in evaluation, evaluators often struggle to know what the method requires in order to be applied properly. Below is a list of four essential requirements:

### 1. Clear, unusually detailed theories of change

QCA is a theory-based approach, so it can only be used when a clear theory of change is present. To develop a list of key conditions to test using QCA, the theory of change must include details on key factors which are felt to be required for the project to achieve a particular outcome or impact. Usually this will require going beyond mapping activities, outputs, outcomes and impacts, to also explore intermediary outputs and outcomes, external contextual conditions, and internal organisational conditions thought to be necessary for success.

### 2. Comparable, granular data across a set of cases

After key factors have been identified in a theory of change, in order to apply QCA comparable data is needed about the presence or absence of these factors, across the set of cases being assessed.

The information available about these factors needs to be carefully compared and benchmarked (also known as 'calibrated') across the set of cases. This involves qualitatively defining scoring criteria - benchmarks - and then using the qualitative data available for each case to rate a particular factor against the scoring criteria. The scores are usually one or zero, which mean respectively presence and absence; however there is room for more fine-grained scales with values like 0.2, 0.6 or 0.33.

This produces a dataset which looks like the example in Table 1.

Table 1 (Illustrative)

Cases	Factor: A	Factor: B	Factor: C	Outcome
	Existence of space for dialogue between state and civil society	Capacities of key civil society actors to engage with state	Horizontal coordination between key civil society actors	Stronger national and regional policy making and implementation
	0 = no or weak evidence to support	0 = no or weak evidence to support	0 = no or weak evidence to support	0 = no or weak evidence to support
	0.33 = some evidence to support	0.33 = some evidence to support	0.33 = some evidence to support	0.33 = some evidence to support
	0.66 = strong evidence to	0.66 = strong evidence to	0.66 = strong evidence to	0.66 = strong evidence to
DRC	0.66	0	0.66	0.66
Senegal	0	0.33	0.66	0.33
South Africa	0	0.66	0	0

### 3. A clear definition of what a ‘case’ is

QCA is a case based approach which requires that clear, distinct cases be present within a programme. How these cases are defined can vary: for example a case could be a country, region or city, an intervention type or implementing partner, or some combination of these. In many evaluations it is not straightforward how cases will be defined; to this end, what matters is that cases present differences in the outcome and in the factors presumed responsible for it. It is also important to define cases so as to have a relatively high number of them, preferably more than five<sup>11</sup>. For instance, in a situation where a programme is being implemented in, say, three countries, defining these three countries as cases is not optimal for QCA, as that will only produce three cases. In this type of situation, it is best to assess whether there is scope to differentiate the cases further within the same country, for example by implementing partner, administrative area or typology of implementation, to allow for more cases to be defined.

### 4. Technical capacity and understanding of QCA software

Finally, specialist skills and software are required, at least for the final part of the analysis. The synthesis procedures in QCA demand a high level of technical understanding, particularly of Boolean logic and set-theory (including fuzzy sets). Familiarity with different software platforms<sup>12</sup> is also required, as different platforms have different strengths and weaknesses and can carry out different types of analyses. A

11. This threshold does not mean that QCA cannot be applied below it, it just means that the full QCA procedure brings more value added to the design as the number of cases increase.

specialist may need to be brought in, if that experience isn't present within the evaluation team.

### Limitations of QCA

QCA is not suitable for use in all situations and has been subject to critique. As detailed in the QCA process step-by-step section below, there may be some instances where QCA is not appropriate or feasible. QCA is not put forward here as a ‘better’ method than others; rather just as another tool in the toolbox for evaluators, which can be useful – particularly in conjunction with other theory-based, comparative or case-based approaches<sup>13</sup>. Further, as QCA is still developing as a methodology, particularly ‘fuzzy-set’ QCA, issues can arise in how datasets are calibrated and the analysis is conducted which can compromise the results<sup>14</sup>. This means that it is important that QCA analysis is done carefully, and that the analysis process is documented clearly and in detail.

Finally, QCA has been critiqued as it involves synthesising rich qualitative information into scores, which removes some of the detail and nuance from the qualitative data. Synthesis itself is not uncommon in evaluations using

12. For the list of available QCA software, see the list provided at <http://www.compass.org/software.htm>.

13. See “A Critique of the Fuzzy-set Methods in Comparative Social Policy” by Seungyoon Sophia Lee, Oxford 2008. <http://www.compass.org/wpseries/Lee2008.pdf>

14. See “Paradoxes and Pitfalls in Using Fuzzy Set QCA: Illustrations from a Critical Review of a Study of Educational Inequality”, by Barry Cooper and Judith Glaesser, Durham University 2011. <http://www.socresonline.org.uk/16/3/8.html>

qualitative methods, and so this is not a unique pitfall. However, it is important to note that the findings from QCA analysis can only be developed by going back to the data to interpret the solutions produced by the analysis, thereby re-grounding the QCA in the original evidence.

## The QCA process: Step by step

It is important to note that QCA isn't a mechanistic process – presented here are key stages in the analysis, but the exact order of steps may vary from application to application and not all of the steps may apply each time. Further, in many cases the QCA process will be iterative – going back and forth between interpretation and analysis to fully unpick relationships between conditions, assess any anomalies in the cases and go back to the evidence to check the findings. Broadly, the QCA process follows these steps:

### 1. Assessing whether QCA is appropriate:

Does the evaluation need to assess what works, why and under what circumstances? Does the evaluation need to synthesise lessons learned across a set of cases? Within an evaluation, QCA may be useful for assessing some types of evaluation questions, but not others.

### 2. Assessing whether QCA is feasible:

Is there a clear theory of change present which includes key explanatory factors (assumptions, context), or can one be developed? Are there enough distinct cases or can such cases be defined? Can consistent qualitative and / or quantitative data be collected across all cases? Is there sufficient technical capacity within the team to use the software, understanding and interpreting the different procedures available for the analysis?

### 3. Theory of change development:

Is there a sufficiently detailed theory of change? A new theory of change may need to be developed, or additional detail may need to be added to an existing theory of change.

### 4. Identification of cases and outcomes of interest:

In order to conduct QCA analysis, there must be different levels of success observed across the cases being compared. Some cases must have a positive outcome, while others have a negative outcome, so that the factors leading to success or “failure” can be compared and identified. In addition, QCA can be used to test multiple outcomes, one by one.

### 5. Developing a set of factors to test for each key outcome:

Based on the theory of change, a set of key factors must be

identified across the cases for each outcome of interest.

### 6. Data collection:

Once the factors of interest have been identified, consistent data must be collected to assess each of these factors across all cases. Cases with missing data cannot be included in the analysis. Data can be primary or secondary, quantitative or qualitative.

### 7. Re-assessing the set of factors and cases for each key outcome:

After data collection, some factors may no longer be seen as relevant or useful, or consistent data on them across all cases may not have been available. In these instances, either the factor or the case would be removed from the QCA analysis.

### 8. Benchmarking and calibration:

Scoring criteria (rubrics) need to be defined for each factor, providing a qualitative descriptor of each score. This provides clear and consistent benchmarks which the data can be assessed against. Once scoring criteria have been defined, numerical values can be associated with each criteria.

### 9. Populating the dataset:

Once scoring criteria have been set and numerical values assigned, the evidence for each factor can be assessed and rated against the criteria. This creates a dataset, most often of zeros and ones, similar to the one included on page four of this brief.

### 10. Analysing the dataset:

A range of software can be used to analyse the data, notably fs/QCA, R and/or Tosmana. To a certain extent, in the presence of a very small number of cases or conditions, relationships can also be spotted by eye. The software provides a rigorous way of analysing patterns that the eye might miss, and usually misses for more than four to five cases or conditions.

### 11. Analyse the findings in light of the original theory of change and hypotheses:

Which aspects of the project theory of change are reinforced and which are challenged? Are there any anomalies or outliers in the data?

### 12. Iteration:

If necessary, develop new hypotheses, add conditions to the dataset and repeat the analysis on the new dataset.

## Conclusions

Designing impact evaluations is increasingly seen as a complex process where different combinations of methods are seen as appropriate or not depending on the evaluation questions and the characteristics of programmes and evaluations<sup>15</sup>. This has implications for the evaluator's methodological toolkit, which needs to include a broad range of options, including methods which are uniquely suited to answer specific questions under specific circumstances.

In this context, QCA fills a specific gap, with its ability to generate robust findings about what makes the difference for success in different contexts, and generalise those findings. It can be used to provide additional rigour to evaluations in cases where experimental methods can't be used; or to triangulate and help contextualise other

qualitative or quantitative approaches. For theory-based evaluations particularly, QCA helps provide additional clarity when testing project theories of change and helps increase the generalizability and replicability of evaluation findings, strengthening the quality and rigour of the evaluation.

**QCA isn't appropriate in all circumstances – it requires a strong theory of change, clearly defined cases and cannot measure the net effects of an intervention, or provide the same level of precision in that sense as quantitative methods. However, it has certain unique strengths – including qualitatively assessing impact and identifying multiple pathways to achieving change which make it a valuable addition to the evaluation toolkit.**

---

15. See Stern et al. (2012) "Broadening the Range of Designs and Methods for Impact Evaluations", DFID Working Paper 38

## About the Authors

Carrie Baptist is an evaluation and research consultant at Coffey. She can be contacted at [Carrie.Baptist@coffey.com](mailto:Carrie.Baptist@coffey.com).

Barbara Befani is an independent consultant and QCA expert. She can be contacted at [Befani@gmail.com](mailto:Befani@gmail.com).