Number 13 January 2016

Centre for Development Impact
PRACTICE PAPER
Innovation and learning in impact evaluation

Qualitative Comparative Analysis: A Valuable Approach to Add to the Evaluator's Toolbox? Lessons from Recent Applications

Abstract A heightened focus on demonstrating development results has increased the stakes for evaluating impact (Stern 2015), while the more complex objectives and designs of international aid programmes make it ever more challenging to attribute effects to a particular intervention (Befani, Barnett and Stern 2014). Qualitative Comparative Analysis (QCA) is part of a new generation of approaches that go beyond the standard counterfactual logic in assessing causality and impact. Based on the lessons from three diverse applications of QCA, this CDI Practice Paper by Florian Schatz and Katharina Welle reflects on the potential of this approach for the impact evaluation toolbox.[1]

1 Introduction

Qualitative Comparative Analysis (QCA) is a social science research method that applies a systematic comparison to case study research. Its purpose is to refine and extend knowledge of the determinants of outcomes by looking at the similarities and differences of cases in terms of the causal factors and outcomes obtained (Cress and Snow 2000). In the impact evaluation field, QCA helps to explore why some interventions were successful in achieving a particular outcome while others were not. It is suitable if an impact evaluation aims to use learning from existing cases to improve future interventions.

QCA is rooted in political science and sociology. The approach was originally developed from the late 1980s onwards to undertake complex comparisons of countries or societies (Ragin 1987; Rihoux and Ragin 2009). One famous early QCA research study explores why a breakdown of democracy occurred in some countries but not others in interwar Europe (De Meur and Berg-Schlosser 1994).

QCA is located between the qualitative (case-oriented) and quantitative (variable-oriented) approaches (Rihoux and Ragin 2009), illustrated by its name in French referring to 'qualitative-quantitative analysis' (De Meur, Rihoux and Yamasaki 2002). QCA requires in-depth qualitative knowledge of each case, which is obtained using a range of qualitative research methods such as ethnography, semi-structured interviews, observation, or literature reviews. Outcome achievement and causal factors are translated into a numerical format to carry out a systematic analysis of patterns across the data.

QCA is a theory-based approach. The evaluator looks for connections between influencing factors (also called 'conditions' in QCA) and outcomes. The potentially influencing conditions are derived from existing social science theory or a programme theory of change. They are tested for their relative influence through a systematic comparison among a number of cases that aim to achieve the same outcome, some successfully and others unsuccessfully. QCA helps to filter out the more important factors from those that are less likely to make a difference among the cases that are investigated in relation to the same outcome. An important element in this analysis is the identification of 'sufficient' and 'necessary' conditions that occur in conjunction with an outcome (Schneider and Wagemann 2012). At the end of the process, the evaluator will usually find a number of different configurations of conditions, or multiple causal patterns that lead to outcome attainment

Table 1 **Different logics in establishing causality**

Correlational/counterfactual logic (statistical approaches)	Multiple-conjunctural logic (QCA, decision tree algorithms)	Generative logic (process tracing, contribution analysis)
To what extent did the intervention make a difference?	*Did the intervention make a difference and through which patterns?*	*How did the intervention make a difference?*
Correlational: establishing the frequency of associations between an intervention and an outcome, holding all other factors constant	Examining which conditions combine and interact to produce an outcome, which also depends on the context	Investigating in depth the different mechanisms at play in a particular case to produce an outcome
Counterfactual: establishing the difference between the case and another case where all other factors except the intervention were the same		

Source: Based on Befani (2012).

in different contexts. A real-world example for this logic is the relationship between heavy smoking and lung cancer: heavy smoking is neither a necessary nor a sufficient cause for lung cancer but, together with other factors, it can contribute to it (Mayne 2008). In applying QCA, the evaluator would compare a number of cases to explore which key factors combine to produce lung cancer in different contexts.

When using QCA as an approach to assess impact, the evaluator follows a particular causal logic: QCA is based on an understanding of causation as 'multiple' or 'conjunctural' (Ragin 1987). This differs from other understandings of causality (see Table 1) in that it assumes that multiple conditions can lead to the same outcome, and that a single condition does not usually produce an effect on its own. Rather, conditions interact and combine to produce an effect, and this effect is also contingent on the wider context (Cress and Snow 2000; Schneider and Wagemann 2012). This understanding is consistent with complexity theory and a realist view of the world (Olson n.d.). In the evaluation literature, this world view is most clearly represented in 'realist evaluation' (Pawson and Tilley 1997). Table 1 compares the multiple-conjunctural logic of QCA with two other common approaches to causality.

How is QCA applied?

QCA employs an iterative approach to data collection and analysis based on systematic procedures, involving the following key steps.[2] First, the evaluator selects a number of cases that share the same outcome ambition. Based on a theory of change (which can be from social science theory, a programme theory of change, personal experience, grey literature on the topic, etc.), the evaluator then unpacks the contextual and programme-related conditions that are likely to influence the outcome. Next, the evaluator iterates between collecting and analysing data against these conditions – refining, adding or taking away conditions and cases, depending on new insights that emerge during the analysis. Because the evaluator goes back and forth between the sampling, data collection and analysis stage, this approach is also called 'retroductive' (as opposed to inductive or deductive) (Olson n.d.). As part of the analysis, the evaluator scores all cases on their outcome achievement, and all conditions on their presence or absence for each case. By doing so, the evaluator reduces a rich data set to a numerical summary that can now be systematically compared using Boolean algebra.

To execute this step, QCA-specific software is needed (such software is usually available free of charge). The analysis reveals sufficient and necessary conditions and combinations of conditions in the data set as a way of highlighting important patterns in the data. Based on that, the number of conditions are reduced to the most crucial patterns leading to outcome achievement or otherwise. A Venn diagram can be used to visualise these patterns. The evaluator interprets the causal patterns based on the rich qualitative data that led to the scoring of presence or absence of conditions and outcomes. This interpretation is crucial in understanding the causal patterns and communicating the results to a wider audience.

Next, we turn to three diverse applications of QCA that illustrate the kinds of insights the approach can reveal, and some key lessons learned: (1) a study on how information and communication technology (ICT)-based reporting can improve water supply sustainability (Welle *et al.* 2015); (2) an evaluation of the Medicines Transparency Alliance (Stedman-Bryce *et al.* forthcoming); and (3) a macro evaluation of the portfolio of the Department for International Development (DFID) in the area of social accountability (Holland and Richardson 2015).[3]

2 Lessons from recent applications of QCA

Case 1: Using QCA to better understand how ICT-based reporting can contribute to water supply sustainability

While great progress has been made in providing basic access to water supply all over the world – covering 89 per cent of the world's population in 2012 (World Health Organization/UNICEF 2014) – sustaining water supply services over time continues to be a problem, particularly

for rural water schemes in sub-Saharan Africa. Experts estimate that approximately a third of all rural water schemes on the continent are non-functional at a given point in time (Foster 2013; Rural Water Supply Network 2009).

As a way to improve sustainability, the water sector has witnessed a surge in the use of ICT to improve the reporting and repairs of breakdowns. The most common ICT initiatives rely on text messages relayed through mobile phone technology providing information on scheme functionality status. While some initiatives use crowdsourcing – i.e. users reporting water scheme breakdowns – other initiatives created mobile apps that support regular reporting mechanisms by government or non-governmental organisation (NGO) staff. A third type of initiative relies on automatic reports generated by mobile phone chips built into hand pump handles. The aim of the QCA research was to carry out a systematic analysis of the factors affecting the success of different ICT initiatives in rendering water services sustainable. Below, we expand on a number of the key lessons emanating from applying QCA and outline some of the insights that we obtained from the results.

Lesson 1: Expect a potential trade-off between developing a concise outcome definition and the number of applicable cases

At the start, we revisited the initially identified outcome to make it more operational for QCA. To do so, we split the overall outcome of 'greater sustainability of water services through ICT-based monitoring' into three sub-outcomes, each representing the successful completion of a sub-process, namely: (1) successful ICT-based reporting; (2) successful processing of ICT reports; and (3) successful water scheme repairs. However, when selecting our cases, we found ourselves with a drastically reduced number of applicable cases. Out of an original data set of 46 cases, only 13 broadly shared the sub-outcomes we had identified. Because of other constraints, we ended up with sufficient data for only eight case studies. Furthermore, we had to exclude two cases from the analysis of our third sub-outcome because they focused on improving sector budgeting and planning rather than specific scheme repairs. These challenges notwithstanding, developing concise outcome definitions added rigour to the comparison in that it forced us to focus only on cases that shared all three sub-outcomes.

Lesson 2: QCA requires high data consistency across cases, which can lead to analytical 'blind spots'

A particular feature of QCA is that a condition can only be included in the analysis if sufficient data can be obtained for all cases. If there are no data for one case, either the case or the condition needs to be dropped. This was the case for one of our conditions, which centred on the trust relationship between users, service providers and policymakers that might prevent people from sending mobile-based reports. We had evidence from some initiatives that water-fetchers were reluctant to use the ICT mechanism because they feared negative repercussions. In other cases, user expectations that the water service provider would improve services were so low that they saw reporting service problems as superfluous. Even though there was some evidence that trust mattered for successful ICT-based reporting, we had to drop our related condition from the analysis because we were unable to obtain information on this for all cases. While we brought trust relationships into our interpretation of the data, it was not formally part of the analysis and might therefore represent a blind spot – a situation where a potentially important influencing factor is omitted from the formal analysis.

Lesson 3: Allocate sufficient time for QCA

A crucial element of QCA is the iteration between the sampling of cases, data collection and analysis. This is necessary to refine the analysis and increase its robustness. For example, in our case, it emerged that various conditions overlapped and needed consolidating, while other conditions needed splitting or reformulating to give interviewees a better idea of what we wanted to find out. This, in turn, sparked new rounds of data collection, as all data have to be collected for all conditions and all cases. The importance of obtaining accurate and sufficient data has implications for the time and level of effort required for carrying out a robust QCA – an element that is often hard to negotiate in evaluations.

Lesson 4: Scoring can sometimes require complex judgements, and there is a danger of losing these in the analysis

In order to allow for QCA, the evaluator scores all cases for their outcome attainment and all conditions on whether they are present or absent. This process leads to a numerical representation of all outcomes and all conditions, which are then analysed via QCA software to arrive at a reduced number of causal patterns that lead to outcome achievement. In crisp set[4] QCA, the scoring is dichotomous: either 0 or 1. In reality, however, it can be difficult to allocate a binary score. For example, one of our conditions was: 'does the ICT-report require human interaction or is it automatic?' One of the cases that we examined used both reporting modes – an automatic and an interactive one. We overcame this issue by focusing on the predominant reporting mode used in this case. There is a danger of omitting such complex judgements in the final write-up of the QCA findings. Therefore, the interpretation of QCA results is a crucial final step. Converting the synthesised and abstract QCA findings into meaningful interpretations can be challenging. It is important that the evaluator has a very good understanding of the analysis and makes use of the rich qualitative data set in his/her interpretation and write-up of QCA results.

What did QCA reveal?

In our study, QCA did not throw up any surprising results. Rather, it added an understanding about which of the known factors were most influential in determining the outcomes among the eight cases, and it provided additional rigour. For example, the analysis showed that initiatives that relied on crowdsourcing were less successful

Figure 1 **MeTA's theory with sub-theories**

Source: Authors' own, based on Stedman-Bryce *et al.* (forthcoming).

than those where reporting was led by the service provider or government. This pattern was consistent across the three sub-outcomes (ICT reporting, report processing and carrying out related repairs): the most successful initiatives were those where all three processes were led by the same organisation. The results confirmed that initiatives aiming to hold a service provider or government to account via citizen action are less likely to be successful than those where there is high responsiveness from the service provider to start with. The added value of the QCA approach was that it confirmed the role of factors already highlighted in the literature via a systematic and methodologically rigorous approach (Welle *et al.* 2015).

Case 2: Evaluation of the Medicines Transparency Alliance (MeTA): understanding how evidence-based policymaking in the medicines sector occurs

Around two billion people in low- and middle-income countries still face challenges in accessing affordable, quality essential medicines (Lu *et al.* 2011). Globally, medicines account for over a quarter of total health expenditures, with some countries spending up to 67 per cent of their total health expenditures on pharmaceuticals (*ibid.*). Inefficient public and private markets, with poorly functioning supply chains and procurement processes, underlie and exacerbate the challenges poor people face in accessing medicines.

Development interventions have intended to strengthen systems involved in the medicines supply chain, but a lack of transparency and accountability has limited what could be achieved. The Medicines Transparency Alliance (MeTA) was established in 2008 in seven pilot countries (Ghana, Jordan, Kyrgyzstan, Peru, the Philippines, Uganda and Zambia), with the aim of improving evidence-based policymaking in the medicines sector. MeTA supports the collection and analysis of pharmaceutical data, in particular around availability and prices of essential medicines. These data are then used to inform national multi-stakeholder platforms engaged in access to medicines policymaking. The platforms, called MeTA councils, include key stakeholders from government, the private sector and civil society.

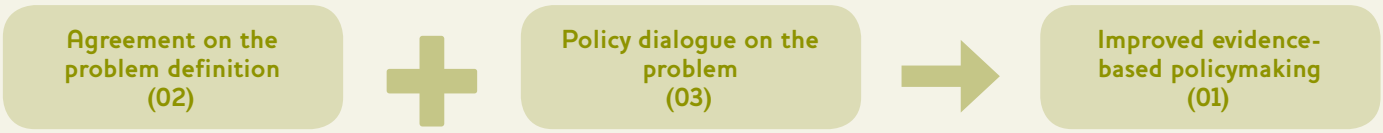
The evaluation team used QCA to test whether focusing on transparency and multi-stakeholder dialogue is effective in improving evidence-based policymaking in the medicines sector and, ultimately, access to medicines. The following section presents some of the main challenges encountered and solutions found when applying QCA for this evaluation.[5]

Lesson 5: There is a trade-off between the number of cases and conditions – the 'problem of limited diversity'

In the case of the MeTA evaluation, one important challenge was the high number of conditions (14) identified through the theory compared to the low number of cases (seven MeTA countries). If there are a large number of conditions and a small number of cases, there are high chances that each case will display a unique combination of conditions. In fact, using crisp set QCA and 14 cases, there are 2^14 = 16384 possible configurations, making it highly unlikely that two cases will show similar configurations of conditions. The analysis then results in an individual description of each case rather than succinct findings on key patterns leading to the outcome across several cases.

The first important measure to address this challenge was to limit the number of conditions to a small set of key conditions. These were identified based on social science theory on evidence-based policymaking – most importantly Kingdon's agenda-setting theory (Kingdon and Thurber 1984) – and adapted to the MeTA context. Second, the evaluation team used discrete sub-theories/models for separate analyses. Three intermediate outcomes were defined in line with Kingdon's agenda-setting theory: (1) agreement on the problem definition; (2) policy dialogue on the problem; and (3) political will to address the problem. For each intermediate outcome, a separate set of influencing conditions was identified. This allowed the evaluation team to include a maximum of five conditions (rather than 14) in each

Figure 2 **Example configuration from the MeTA evaluation**



Source: Authors' own, based on Stedman-Bryce *et al.* (forthcoming).[6]

analysis, mitigating the problem of limited diversity. Thereafter, an overarching analysis was undertaken to draw the findings together and assess what drives evidence-based policymaking in the medicines sector. Figure 1 illustrates how the sub-theories/models and the overall theory/model fit together.

Lesson 6: QCA forces evaluators to employ very tight definitions and therefore enhances replicability

Initially, the evaluation team used a condition labelled 'cohesive policy communities'. However, this was found to be too vague and had too many sub-dimensions. Different evaluation team members understood and scored this definition differently. The evaluation team therefore redefined the condition to ensure consistent and objective coding. The new condition focused on consistent multi-stakeholder engagement; presence of the condition was tightly defined as 'participation by civil society, private sector and government representatives in at least 50 per cent of multi-stakeholder meetings'. This definition was precise enough so that all team members assigned the same scores. The application of QCA forced the evaluation team to use such tight definitions to ensure rigorous analysis. This also had wider benefits, such as an increased level of objectivity and replicability of our analysis.

Lesson 7: QCA findings are very technical and abstract, and need to be prioritised and translated for the end user

QCA findings tend to be very technical, and it can be challenging to present findings in a concise way without glossing over the analysis. One approach to make QCA findings more accessible was the use of visuals to illustrate different causal packages identified by QCA, as illustrated in Figure 2. The visual shows that a combination of two conditions – for example, 'agreement on the problem definition' and 'policy dialogue on the problem' – led to the outcome of 'improved evidence-based policymaking'.

Moreover, the evaluation team also focused on a small number of key findings to strike a balance between communicating those clearly without losing analytical depth. To do so, findings with high levels of consistency and coverage were prioritised – e.g. conditions or combinations of conditions that were found to influence the outcome in most cases and without contradiction. This step was essential, because QCA tends to generate a large number of findings and it can be difficult to see the forest for the trees.

What did QCA reveal?

Applying QCA helped the evaluation team to generate a number of interesting and surprising findings. For example, the team found that political support for addressing access to medicines issues at high levels was not necessary to achieve evidence-based policymaking in the sector. This somewhat counterintuitive finding relates to the MeTA approach of working through multi-stakeholder dialogue. By bringing civil society and the private sector into the policymaking process, MeTA established a working platform at the technical level of government, where high-level political support was less important. In the three countries where policies were improved without consistent support from high levels of government (Jordan, Peru and Zambia), MeTA managed to develop close working relationships with senior civil servants.

Similarly, QCA revealed that public pressure to highlight access to medicines issues was not a contributing factor to success, but sometimes a barrier. In some cases, notably Ghana and Uganda, public pressure generated a response from high-ranking political figures; however, this was sporadic and did not lead to continuous support. In other countries such as Kyrgyzstan, the Philippines and Zambia, civil society adopted a much more cooperative approach, using the multi-stakeholder platform to directly access policymakers instead of creating 'civil society noise'.

Finally, the evaluation team also found that transparency in terms of data collection and analysis was not important by itself. However, data played a key role when informing multi-stakeholder dialogue. The multi-stakeholder dialogue supported by MeTA was found to be a key driver of success. QCA therefore helped validate the MeTA approach of developing multi-stakeholder dialogue as a means of improving evidence-based policymaking in the medicines sector.

Case 3: Macro evaluation of DFID's portfolio in social accountability: understanding what works, for whom, in what contexts and why

DFID has a large and diverse portfolio of around 168 projects[7] in the area of social accountability. Social accountability bounds interventions that seek to influence the 'short route' of accountability, through increased engagement between service users ('demand side') and service providers ('supply side'), on the premise that voice, choice and accountability in service delivery will improve the quality, accessibility and reliability of services, and secure longer-term improvements

in wellbeing. Social accountability can take many different forms, including citizen score cards, expenditure tracking, media campaigns, and building deliberative platforms between service providers and users. DFID's social accountability projects use different intervention strategies and modalities, target different outcomes, and operate in very different contexts. Many of the outcomes targeted are also focused on changes in empowerment and accountability, which are difficult to measure.

Drawing from learning from this complex portfolio, the macro evaluation aims to generate evidence of what works, for whom, in what contexts and why. The evaluation is still ongoing, but we have conducted a pilot to test our methodology from which the lessons below are taken. In the pilot, we used QCA to carry out a systematic analysis of 15 projects and identified key patterns of conditions influencing outcomes. Using an approach based on realist evaluation principles, we systematically included context factors in our QCA conditions. While the pilot combined QCA with narrative analysis, this paper focuses on the lessons learned from applying QCA.

Lesson 8: A mix of 'successful' and 'unsuccessful' cases is needed to generate insightful findings

In the case of the macro evaluation, the team also observed that most projects were classified as 'successful' and as having achieved the outcome. This makes it difficult to contrast 'successful' with 'unsuccessful' cases and provide insights into causality as opposed to just describing 'success'. If all cases included in the QCA display the outcome, there is no internal point of comparison and learning is limited to learning from success stories.

This is a particular challenge to evaluations, given the positive bias often observed in intervention reporting (Camfield, Duvendack and Palmer-Jones 2014; Independent Commission for Aid Impact 2014; Lloyd, Poate and Villanger 2014). In the macro evaluation, we used two approaches to address this challenge. First, we used intermediate outcomes such as 'increased citizen engagement', which showed more diversity in terms of 'success' and 'failure'. Second, 'success' was defined more tightly and we increased the threshold for outcomes to be regarded as achieved, where it was possible to justify this from a theoretical perspective.

Lesson 9: Splitting outcomes into several elements can help add nuances

QCA allows both for dichotomous calibrations, in which conditions are scored as being either present or absent (a binary 0 or 1), and for fuzzy calibrations, which allow for the scoring of different degrees of presence or absence (such as 0.33 or 0.67). The former is called crisp set QCA, while the latter is labelled fuzzy set QCA. While fuzzy set QCA appears attractive due to its ability to provide more nuanced scores, it faces several challenges. First, fuzzy set QCA is much more complex, requires additional definitions of different values, and is therefore more time-consuming and resource-intensive. Second, findings are typically the same or very similar to what would have been obtained by using crisp set QCA. Finally, fuzzy set QCA requires more fine-grained data, which is often not available for evaluations. For these reasons, most evaluations to date have used crisp set QCA, which may be more applicable given real-world resource constraints.

A better solution to capturing nuances is to break up outcomes into several elements and score them separately. In the case of the macro evaluation, there were too many degrees of presence and absence to score the main outcome of improved service delivery dichotomously. In some cases, service delivery had improved at the project/local level but not at the national level, whereas in others it was a nationwide improvement. These improvements were not of the same magnitude, and the outcome was split into a lower-level outcome of improved service delivery at the local/project level, and a higher-level outcome of improved service delivery at the national level. This allowed for a more nuanced measurement while maintaining dichotomous conditions/crisp set QCA.

Lesson 10: Combining QCA with other approaches is essential to fully understand change

While splitting outcomes and other conditions into several elements can help add nuances, the findings are still relatively broad and generic. For instance, in the case of the macro evaluation, we may find that mandating the inclusion of marginalised groups in decision-making platforms is a sufficient condition for improved service delivery for such groups. Such a QCA finding would be intuitive and interesting; however, it does not tell us much about the specific mechanism at work, including the role of contextual factors. For example, it may be that the representatives of marginalised groups in decision-making platforms push reforms and changes that benefit their constituency. Or it may be that simply including them in those platforms raises attention to the issues facing marginalised people, and improves their situation without the need for reforms or changes.

Similarly, QCA might tell us that this finding is contingent on the context condition of low levels of social inequality. This could be the case because, in such a context, there are simply not many marginalised people, making it relatively easy to improve their situation. It could also mean that there are already effective policies and systems in place to limit social inequality, with a culture of inclusion that makes it much easier for people to accept a quota for marginalised groups in decision-making bodies. Only an in-depth, qualitative assessment can tell us exactly what has happened, which is why combining QCA with other approaches is essential for generating operational insights into what works, in what contexts, and how.

What did QCA reveal?

The lessons above are drawn from the pilot phase of this evaluation. The second phase, with a refined methodology and a larger number of cases, is still ongoing. The focus of the pilot was on testing the methodology, which limited the technical insights gained.

3 Conclusions

The three cases presented above show that applying QCA with a view to assessing the impact of an intervention can confirm existing findings by adding rigour via systematic comparative analysis, and that it can also lead to new insights on causal patterns. QCA is most relevant if we want to better understand the determinants of achieving an outcome when comparing a number of interventions. In such cases, QCA enables the evaluator to compare a number of cases, which is difficult to do without a systematic method. By making all assumptions and choices explicit, QCA enforces a very systematic and transparent approach.

QCA is also able to identify different complex causal patterns rather than simplistic answers, which is in line with the type of causality often observed in the real world (Rihoux and Ragin 2009). As a result, QCA has important benefits to offer in the current quest to expand the diversity of rigorous approaches to impact evaluation. However, QCA also comes with some potential caveats that need to be taken into consideration when choosing it as the main evaluation approach.

The evaluation team needs to have a thorough understanding of QCA: Applying QCA in a rigorous way requires that the evaluation team have a thorough understanding of the methodology in order to solve issues that may arise during the analysis. Examples from the above case studies include the problem of 'limited diversity', a lack of clear definitions, and solving problems related to the scoring of outcomes and conditions. Guidance from a seasoned QCA practitioner is important in maintaining the rigour of the approach. The quality of the application rests strongly on the ability of the evaluator to identify and refine the right set of cases and conditions to be tested, based on a strong theory of change. In a similar vein, translating QCA findings into policy advice or actionable recommendations is not easy. It requires the evaluator to have a good grasp of the detailed data and of the data analysis, which follows a different logic from the classic, counterfactual approach to impact evaluation. The iterative nature of QCA requires sufficient resource allocation for several rounds of data collection and analysis, including possibly for identifying additional cases.

Evaluation client expectations and concerns need to be clarified from the outset: The QCA requirement of data consistency across all cases might lead to the omission of some cases or conditions from the analysis. The potential loss of case studies may be a difficult concession to make for some evaluation clients. In addition, QCA does not measure the *net* impact of an intervention, and does not explain the nuanced mechanism at work and how it is embedded in context. Rather, it identifies packages of conditions associated with the outcome of interest. This logic may not meet the expectation of the evaluation client, and it will be important to clarify this from the outset.

A potential solution to this problem is to combine QCA with other approaches that are geared at isolating the contribution of a given intervention such as contribution analysis (Mayne 2008) or to apply other qualitative approaches that further unpack how context matters. Similarly, combining QCA with a Realist Evaluation approach (Befani, Ledermann and Sager 2007) can help unpack the role of context and the specific mechanisms at work. As with many other evaluation approaches, QCA on its own remains a narrow tool, and should ideally be combined with other approaches to ensure both breadth and depth of the analysis.

The evaluation object needs to be fit for QCA: Finally, the utility of QCA also depends very much on the nature of the evaluation object. QCA is most relevant if the evaluation object comprises a number of comparable cases, and if the causality observed fits with the QCA logic of multiple causal pathways. Moreover, QCA is most applicable if a complete data set is available; other approaches may be more suitable when a data set is likely to be patchy.

Is QCA a valuable addition to the evaluator's toolbox? We think so, with the proviso that its utility depends on the three points raised above: the evaluation team's ability to apply the approach; a clear fit with the client's expectations; and suitability of the evaluation object. Applying QCA is a fine art and the challenges in navigating through the process should not be underestimated.

Notes

1 This CDI Practice Paper builds on an earlier contribution to the series on QCA (Befani 2013), which compares the pros and cons of statistical analysis and QCA in assessing policy influence and uptake.
2 A detailed step-by-step guide to applying QCA is presented in a Coffey 'How To' note (Baptist and Befani 2015).
3 Case 1 is based on research funded by the Making All Voices Count initiative, while cases 2 and 3 draw on lessons from DFID-funded evaluations.
4 For a discussion of fuzzy set QCA, see Lesson 9 below.
5 Some of the lessons presented above also apply to this case, but are not repeated here.
6 The coverage and consistency of this configuration were both 1.000000.
7 This number was generated in September 2015 and will change over time as the evaluation database evolves. The database and information on the evaluation methodology and emerging findings can be found here: www.itad.com/knowledge-and-resources/dfids-macro-evaluations/empowerment-and-accountability-macro-evaluation/ (accessed 1 December 2015).

References

Baptist, C. and Befani, B. (2015) *Qualitative Comparative Analysis – A Rigorous Qualitative Method for Assessing Impact*, London: Coffey

Befani, B. (2013) 'Between Complexity and Generalization: Addressing Evaluation Challenges with QCA', *Evaluation* 19.3: 269–83

Befani, B. (2012) 'Models of Causality and Causal Inference', in E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (eds), *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper 38, London: Department for International Development

Befani, B.; Barnett, C. and Stern, E. (2014) 'Introduction – Rethinking Impact Evaluation for Development', *IDS Bulletin* 45.6: 1–5

Befani, B.; Ledermann, S. and Sager, F. (2007) 'Realistic Evaluation and QCA: Conceptual Parallels and an Empirical Application', *Evaluation* 13.2: 171–92

Camfield, L.; Duvendack, M. and Palmer-Jones, R. (2014) 'Things You Wanted to Know about Bias in Evaluations but Never Dared to Think', *IDS Bulletin* 45.6: 49–64

Cress, D.M. and Snow, D.A. (2000) 'The Outcomes of Homeless Mobilization: The Influence of Organization, Disruption, Political Mediation, and Framing', *American Journal of Sociology* 105.4: 1063–1104

De Meur, G. and Berg-Schlosser, D. (1994) 'Comparing Political Systems: Establishing Similarities and Dissimilarities', *European Journal of Political Research* 26.2: 193–219

De Meur, G.; Rihoux, B. and Yamasaki, S. (2002) 'L'analyse quali-quantitative comparée (AQQC-QCA): approche, techniques et applications en sciences humaines', https://scholar.google.co.uk/scholar?hl=en&q=analyse+quali-quantitative+compar%C3%A9e&btnG=&as_sdt=1%2C5&as_sdtp=#0

Foster, T. (2013) 'Predictors of Sustainability for Community-Managed Handpumps in Sub-Saharan Africa: Evidence from Liberia, Sierra Leone, and Uganda', *Environmental Science & Technology* 47.21: 12037–046

Holland, J. and Richardson, E. with Befani, B.; Forss, K. and Hughes, C. (2015) *Macro Evaluation of DFID's Policy Frame for Empowerment and Accountability. Empowerment and Accountability Annual Technical Report 2015: Final Draft*, ePact

Independent Commission for Aid Impact (2014) *How DFID Learns*, http://icai.independent.gov.uk/report/dfid-learns/ (accessed 18 December 2015)

Kingdon, J. and Thurber, J. (1984) 'Agendas, Alternatives, and Public Policies', http://mason.gmu.edu/~alovell/portfolio/images/Educ875_Agendas,_Alternatives_and_Public_Policies.pdf (accessed 18 December 2015)

Lloyd, R.; Poate, D. and Villanger, E. (2014) 'Results Measurement and Evaluability: A Comparative Analysis', *Journal of Development Effectiveness* 6.4: 378–91

Lu, Y.; Hernandez, P.; Abegunde, D. and Edejer, T. (2011) *The World Medicines Situation 2011. Medicine Expenditures*, 3rd ed., Geneva: World Health Organization, www.who.int/entity/health-accounts/documentation/world_medicine_situation.pdf?ua=1 (accessed 21 December 2015)

Mayne, J. (2008) *Contribution Analysis: An Approach to Exploring Cause and Effect*, ILAC Brief 16, Institutional Learning and Change Initiative

Olson, W. (n.d.) 'What is Qualitative Comparative Analysis?', PowerPoint presentation

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, London: SAGE

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

Rihoux, B. and Ragin, C. (eds) (2009) *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*, Thousand Oaks: SAGE

Rural Water Supply Network (2009) Handpump Data, selected countries in sub-Saharan Africa

Schneider, C.Q. and Wagemann, C. (2012) *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*, Cambridge: Cambridge University Press

Stedman-Bryce, G.; Balogun, C.; Hodgkin, C. and Schatz, F. (forthcoming) 'Medicines Transparency Alliance (MeTA) Evaluation: Final Report', ePact

Stern, E. (2015) *Impact Evaluation: A Guide for Commissioners and Managers*, London: Bond

Welle, K.; Williams, J.; Pearce, J. and Befani, B. (2015) *Testing the Waters: A Qualitative Comparative Analysis of the Factors Affecting Success in Rendering Water Services Sustainable Based on ICT Reporting*, Brighton: Itad

World Health Organization/UNICEF (2014) *Joint Monitoring Programme (JMP) for Water Supply and Sanitation*, www.wssinfo.org/ (accessed 21 December 2015)

Centre for Development Impact (CDI)

The Centre is a collaboration between IDS (www.ids.ac.uk), Itad (www.itad.com) and the University of East Anglia (www.uea.ac.uk).

The Centre aims to contribute to innovation and excellence in the areas of impact assessment, evaluation and learning in development. The Centre's work is presently focused on:

(1) Exploring a broader range of evaluation designs and methods, and approaches to causal inference.
(2) Designing appropriate ways to assess the impact of complex interventions in challenging contexts.
(3) Better understanding the political dynamics and other factors in the evaluation process, including the use of evaluation evidence.

This CDI Practice Paper was written by **Florian Schatz** and **Katharina Welle**.

ISSN: 2053-0536
AG Level 2 Output ID: 321

Institute of Development Studies, Brighton BN1 9RE, UK
T +44 (0) 1273 915637 **F** +44 (0) 1273 621202 **E** ids@ids.ac.uk **W** www.ids.ac.uk