Abstract

Many evaluations of infrastructure projects rely on methods that ignore the complexity of the projects. Although case studies are attentive to project complexity, it is difficult to identify general patterns that would apply to a larger sample of projects. Qualitative comparative analysis is a method that preserves the complexity of projects and generates insights across cases. In this contribution, we discuss our experiences with using qualitative comparative analysis for the evaluation of the planning and implementation of complex infrastructure projects. We will provide a short introduction into the main properties of the method (complex causality, systematic comparison) as well as describe some of the main operations (calibration, truth table analysis, interpretation). This will serve to demonstrate why qualitative comparative analysis is a fitting evaluation method in project development and implementation. Next, we will show how we used the method in a research project that aimed to find out under what conditions unplanned events in the implementation of infrastructure projects were dealt with satisfactorily, that is, what it took to respond to these events in an apt manner. Based on our experiences, we will summarize main lessons learned for conducting qualitative comparative analysis proper and provide suggestions for further reading.

Learning Outcomes

By the end of this case, students should be able to

- Explain the added value of using qualitative comparative analysis in evaluation research compared with large-\(n\) variable-oriented methods and single/small-\(n\) case-based methods
- Explain the main properties of qualitative comparative analysis (complex causality, systematic comparison)
- Apply the main operations of qualitative comparative analysis (calibration, truth table analysis, interpretation)

Project Overview and Context

Our project started with the observation that evaluations of the performance of infrastructure projects often rely on methods and research techniques that are relatively ignorant of the complexity of such projects (Gerrits & Verweij, 2012, 2013). An example is provided in the well-known work of Flyvbjerg, Bruzelius, and Rothengatter (2003). Using a large and quantitative database, they were able to identify patterns such as that rail projects have a stronger tendency to lead to cost overruns than road projects, or that European projects are more likely to experience cost overruns than projects in the United States. Given their large database and focus on the relationships between (quantitative) variables, it is difficult to unearth the complexity of individual infrastructure projects as they are being planned and implemented. Large-\(n\) approaches, although useful in
many ways, cannot explain the influence of unique and situational factors at play in specific cases. Cost overrun in one project may be caused in entirely different ways than in another project. Single in-depth case studies are better at showing this complexity. In a single case study, the researcher can focus on the many variables, relationships, perspectives, and events that are at play—holistically and systemically—in such projects. The downside of a case study, however, is that everything starts to appear to be unique. After all, no two roads, waterways, or train tracks are built in exactly the same way under the exact same circumstances. As such, it is hard to turn the case-specific findings into generalized statements that would apply to a large sample of projects. On top of that, case studies are often criticized for lacking transparency and rigor.

In our research project, we investigated whether qualitative comparative analysis (QCA) (Ragin, 1987, 2000, 2008) could be used to bridge the gap between large-n variable-oriented methods and single/small-n case-based methods in research on complex infrastructure project planning and implementation. Based on our investigation, we concluded that QCA is indeed helpful in preserving the complexity of individual projects while also generating insights across cases (Gerrits & Verweij, 2013; Verweij & Gerrits, 2012, 2013)

These findings encouraged us to further deploy QCA in empirical research. QCA requires the researcher to identify the conditions that are believed to be possible explanations for a defined outcome of interest. This is somewhat different from selecting variables. The conditions in QCA are part of the case (i.e., they constitute the case), whereas variables are instead external to the case (Gerrits & Verweij, 2018). In one study, for example, we selected the conditions “project management,” “public-private cooperation,” “project scope,” “project size,” and “contract type” to understand if and how (combinations of) these conditions could explain whether people were satisfied with a project (Verweij, 2015b). In QCA, the combinations of conditions form configurations that explain the outcome. Each case is assigned to the logically possible configuration that it represents. The comparative procedure then centers on the pairwise comparison of those configurations that agree on the outcome (based on the cases that it is covered by) and differ in but one of their conditions. For instance, if two configurations both had a high satisfaction, and both had the same project management style, the same public–private cooperation style, the same project scope, and the same project size, but they had different contract types, then we could infer that the “contract type” is irrelevant for explaining project satisfaction. Given the ways in which the method works, it is important to keep the number of cases and the number of the conditions limited (Rihoux & Ragin, 2009). Having too many cases will prohibit the researcher from obtaining in-depth knowledge of the individual cases. Having too many conditions will likely result in rather complex cross-case patterns, thus prohibiting the researcher from saying anything meaningful.

In what follows, we will explain how we used QCA in our study on megaprojects (Verweij, 2015a; Verweij & Gerrits, 2014, 2015; Verweij, Teisman, & Gerrits, 2017). The study we explain focused on the occurrence of unplanned events in a single project, the management responses to them, and the resulting outcomes in terms of the managers’ satisfaction with how the event was dealt with.

---

**Research Questions and Design**
In this research, we aimed to see which responses to unplanned events would bring about the most favorable outcome in terms of managers’ satisfaction. The specific research question was (see Verweij, 2015a): How did the project managers in the implementation of the highway megaproject “A15 Maasvlakte-Vaanplein” respond to events, and which management responses produced satisfactory outcomes?

The events, and the responses to them by the managers, tend to be very particular to that event’s situation. QCA allowed us to do justice to the complexity of those situations and at the same time evaluate, in a very transparent and systematic manner, the configurations of conditions that explained the satisfaction (or lack thereof) in the project. We will explain how we conducted our study and present our experiences, in particular regarding the case study design and analysis.

Although there are multiple examples of QCA using large-n quantitative data, the method was originally conceived as a way of dealing with a small or medium number of cases in a systematic manner (Pattyn, Gerrits, & Verweij, 2015; Rihoux, 2013). Using qualitative data as the basis will allow the researcher to preserve more of the complexity of the case. The comparative part of the method requires the researcher to transform the (qualitative) data into scores between 0 and 1. This is referred to as calibration: the process of assigning numerical values to chunks of case-based data for each condition and the outcome. In our research project, the outcome was the satisfaction with how the event was dealt with. We calibrated three conditions—“nature of the event,” “project management style in response to the event,” and “public–private cooperation style in response to the event”—and of course the outcome “satisfaction”—and of course the outcome “satisfaction” into numerical values. Informed by existing literature, we calibrated our conditions using the calibration scheme as shown in Table 1. Note that other types of calibration are also possible, most notably “fuzzy-set calibration” where cases can receive finer-grained scores between 0 and 1, for example, 0.00, 0.33, 0.67, and 1.00 (Ragin, 2008).

### Table 1. Calibration rules in the study.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Abbreviation</th>
<th>Calibration rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of the event</td>
<td>EVENT</td>
<td>0 = physical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = social</td>
</tr>
<tr>
<td>Response</td>
<td>MAN</td>
<td>0 = internally oriented</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = externally oriented</td>
</tr>
<tr>
<td>Public–private cooperation</td>
<td>COOP</td>
<td>0 = private contractor acts autonomously</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = public client cooperates in the response</td>
</tr>
<tr>
<td>Satisfaction (the outcome)</td>
<td>SATIS</td>
<td>0 = dissatisfaction prevails</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = satisfaction prevails</td>
</tr>
</tbody>
</table>
Research Practicalities and Data Collection

The qualitative data were collected through interviews with project managers in the project. Interviews were first held with project managers from the public client “Rijkswaterstaat” (Ministry of Infrastructure and Water Management), because we already established a relationship with them through pilot interviews that we conducted prior to the study. In the pilot interviews, a relationship of trust was established between the researchers and the respondents, where the latter learned that we intended no harm to the project and that our interest was in learning about the project’s implementation process. The thus established relationship gave us an entry point to the A15 Maasvlakte-Vaanplein project. Before the interviews, the respondents from Rijkswaterstaat were a bit hesitant to help us in also getting interviews with the project managers from the private contractor A-Lanes A15. The reason was that Rijkswaterstaat and A-Lanes A15 were at the time in a rather tense relationship with each other because of implementation and collaboration difficulties in the project. However, trust and rapport between us and the project were established over time with each interview conducted with the Rijkswaterstaat respondents; the respondents were offered the opportunity to review the transcripts of the interviews, to identify parts of the transcripts that they did not want to be quoted, and the anonymity of their statements was guaranteed. After some interviews were held and the respondents saw that we intended no harm, the project director of Rijkswaterstaat contacted his compeer at A-Lanes A15 to ask for A-Lanes A15’s cooperation with our study. Hence, we got access to A-Lanes A15 for conducting interviews there as well. We conducted a total of 20 interviews between May 2012 and January 2013 (Verweij, 2015a).

We identified the events using an open interview strategy. We asked the respondents to elaborate on the daily affairs of the megaproject and asked them to give concrete examples of issues they had run into (see, for example, Weiss, 1994). In the interviews, we picked up and followed as many “markers” as we could, asking the respondents to elaborate on those markers using explicit descriptions of what actually happened. This allowed us to identify 20 unplanned events in the project’s implementation process. We then discussed the various events that had taken place with the respondents: we asked for the specific conditions under which they had taken place, which actors were involved in the events, the management actions that were deployed in response to the events, and the ex-post satisfaction with those responses. We transcribed the interviews and coded (see, for example, Boeije, 2010) them for the three conditions and the outcome. The data were then calibrated.

Data Analysis With QCA

Our experience has taught us that data analysis is not a discrete procedure. In many ways, it starts during the very first interview when the researcher reflects on the data obtained. In fact, data analysis and interpretation are iterative as the researcher moves back and forth between raw data, calibration, and potential conclusions. We also experienced that QCA helps to systematize this iterative process (Gerrits & Verweij, 2018). This
makes the research transparent and open to critical assessment by others.

The first step of the data analysis was to calibrate the coded qualitative data. Using the calibration rules as summarized in Table 1, we drafted the so-called “calibrated data matrix” depicted as Table 2. The data matrix was the result of several iterations. We started out with a different set of calibration rules, where the condition COOP had three instead of two categories. During the analysis, however, we discovered that the third category added little to our understanding of the main question, while it did produce rather complex patterns that were more difficult to interpret. This worked as follows. Because we started out with two binary conditions and one multi-value condition with three categories, the number of logically possible configurations was 12 (i.e., $2^2 \times 3^1$). We learned that the 20 events (cases) clustered together in six logically possible configurations; this means that six other logically possible configurations were not represented by any of the cases. Consequently, there were relatively few pairs of configurations that agreed on the outcome (based on the cases that they were covered by) and differed in but one of their conditions. As a result, few pairwise comparisons could be made. As such, not many conditions could be put aside (“minimized away”) as being irrelevant for explaining satisfaction. We thus did not gain much from the pairwise comparison; the results were rather complex and difficult to interpret. After we converted the multi-value condition into a binary one, the number of logically possible configurations was reduced to eight (i.e., $2^3$). Now, more comparisons could be made leading to more sensible results. These iterations and changes are allowed—encouraged even—as long as they are theoretically and/or empirically substantiated, and as long as they are documented in a log so that they are open to scrutiny. Naturally, these iterations need to improve the researcher’s understanding of the cases being compared. The iterations pressed us to look at our data from multiple angles, thereby gaining a richer understanding of the cases (events) and of the megaproject as a whole.

**Table 2. Calibrated data matrix in the study.**

<table>
<thead>
<tr>
<th>Cases (i.e., events)</th>
<th>EVENT</th>
<th>MAN</th>
<th>COOP</th>
<th>Outcome: Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAB1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CIT1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CIT2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CIT3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CIT4A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CIT4B</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>DOW</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Once the calibrated data matrix was put together, we had to develop the truth table. The truth table lists all of the logically possible configurations. Each row in the truth table represents one logically possible configuration (Ragin, 1987). In our study, the truth table had a length of eight rows (i.e., $2^3$). We then assigned each of the 20 cases to the truth table row that it belonged to. For instance, the cases DOW, GRO, and EXP were assigned to the first truth table row, as shown in Table 3. Finally, we assigned each truth table row an outcome score of either 0 or 1, on the basis of the outcome score(s) of the case(s) in the particular truth table row. A score of 0 meant that the particular configuration, according to the empirical data (i.e., the cases in that truth table row), produced dissatisfaction. A score of 1 implied that the configuration produced satisfaction. For instance, because the cases DOW, GRO, and EXP all had a low satisfaction (see Table 2), the configuration was assigned an outcome of 0. The truth table of our study is provided in Table 3.

**Table 3. Truth table of the study.**

<table>
<thead>
<tr>
<th>Case</th>
<th>EXP</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRO</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HBR1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>HBR2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>HBR3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LEI</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MUN1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MUN2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>PRO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>PRV1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>PRV2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>RWS1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RWS2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Adapted from Verweij (2015a).*
In the next step, the truth table had to be analyzed. This is the so-called “truth table minimization.” In this step, each configuration from the truth table is compared one-on-one with each of the other configurations in the truth table. The guiding rule of the truth table minimization is,

if two […] expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression. (Ragin, 1987, p. 93)

In practical terms, the configurations that agree on the outcome and differ in but one of their conditions are pairwise compared.

We noted that two truth table rows were devoid of cases (see Table 3). These so-called “logical remainders” can also serve a role in the analysis (see, for example, Gerrits & Verweij, 2018). However, we refrained from using them because of the highly empirical nature of the study and because of the absence of strong and unambiguous theoretical expectations about the impact of the conditions on the outcome. If you would use the logical remainders, this means that you are in fact conducting a kind of counterfactual analysis where the empty truth table rows serve as “thought experiments”—“if the configuration would have been empirically present, then we would have expected it to produce this outcome.” The inclusion of the logical remainders can potentially lead to simpler expressions (i.e., more conditions that can be “minimized away”), because more comparable pairs of configurations (agree on the outcome, differ in only one condition) are likely to be found. However, because the logical remainders will carry the same weight in the pairwise comparison as...
the empirically present configurations, it is important that strong theoretical arguments are available about the
configuration producing a certain outcome. We felt that in our study this was insufficiently the case.

We performed two comparative analyses: one for the configurations with dissatisfaction (outcome 0) and one
for the configurations with satisfaction (outcome 1). The results of the analyses are shown in Tables 4 and 5.
The lowercase conditions indicate a calibrated score of 0 and the uppercase conditions indicate a calibrated
score of 1. For instance, “EVENT*MAN” indicates that social events that are managed externally oriented
produces satisfaction (see Table 4). Conversely, when social events are managed internally oriented (i.e.,
EVENT*man), then dissatisfaction is the result (see Table 5).

**Table 4. Results from truth table analysis for satisfaction.**

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Cases</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>event<em>man</em>COOP</td>
<td>CAB1</td>
<td>1</td>
</tr>
<tr>
<td>EVENT*MAN</td>
<td>CIT1, LEI, PRO</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>CIT2, PRV2, HBR2, PRV1</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Adapted from Verweij (2015a).*

**Table 5. Results from truth table analysis for dissatisfaction.**

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Cases</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVENT*man</td>
<td>CIT3, CIT4A, MUN2, RWS1</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>CIT4B, RWS2, HBR1, HBR3, MUN1</td>
<td></td>
</tr>
<tr>
<td>man*coop</td>
<td>DOW, GRO, EXP</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>CIT3, CIT4A, MUN2, RWS1</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Adapted from Verweij (2015a).*

The results express the generalized cross-case patterns in the megaproject. At the same time, they also
show the complexity of the megaproject’s implementation. In QCA, there are in particular three aspects to this
complexity (but see also Gerrits & Verweij, 2016, 2013; Verweij & Gerrits, 2013). First, both dissatisfaction
and satisfaction are explained by **configurations**. In QCA, this is referred to as conjunctural causation, that
is, a “situation in which the effect of a single condition unfolds in combination with precisely specified other
conditions” (Schneider & Wagemann, 2012, p. 324). Second, **multiple** configurations explain the outcome. In
QCA, this is referred to as equifinality, that is, the situation where there are “different, mutually non-exclusive
[…] paths for the outcome” (Schneider & Wagemann, 2012, p. 326). Third, a condition can **play different
roles** depending on the other conditions in a configuration. This can, for example, be seen by comparing
event*man*COOP (satisfaction) with EVENT*man (dissatisfaction). For physical events (event) and when the public and private partners cooperate (COOP), an internally oriented management style (man) produced satisfaction. However, in response to social events (EVENT), an internally oriented management style (man) produces dissatisfaction. This issue is referred to in QCA as multifinality, that is, the situation where the same “condition can be causally relevant for producing both the occurrence of the outcome Y and its complement ~Y” (Schneider & Wagemann, 2012, p. 329). In this example, the condition “man” played different roles in different situations.

The results we obtained had to be interpreted and explained. We learned that it is important not to conclude the analysis as soon as the cross-case patterns are produced. It is only once you have the patterns that perhaps the most interesting part of the analysis starts. Benoît Rihoux (2003; Rihoux & Lobe, 2009) explained this by using the metaphor that the minimized configurations act as “flashlights” that illuminate the crucial spots in the black boxes that are the cases; they highlight the most relevant parts of the cases for explaining the outcome in those cases. With the results now on the table, we went back to our raw qualitative data and interpreted the findings. That is, using the minimized configurations to focus our qualitative interpretation, we tried to make sense of the cases. As such, we arrived at the following explanations for the findings (adapted from Verweij, 2015a).

1. Two configurations are associated with satisfaction. The configuration that is represented only by CAB1 is an odd one out (see Table 4). The case CAB1 concerned an event where cables and pipelines in the project area had to be rerouted to make room for the newly planned A15 highway. In this event, Rijkswaterstaat and A-Lanes A15 decided to cooperate, although they had separate financial responsibilities for different cables and pipelines involved, simply because the cables and pipelines were physically intertwined.

2. The other configuration was covered by seven cases (see Table 4). In those cases, social events that were responded to with externally oriented management resulted in satisfaction. There, the public–private cooperation was not the decisive factor. Instead, it appeared that when the project managers actively engaged with the external stakeholder environment to relieve the stakeholders’ worries and frustrations with the project implementation (e.g., noise nuisance and interference of the megaproject with other external road planning processes by other public authorities), this resulted in satisfactory solutions. We found that this finding was in line with previous research on successful project management strategies (e.g., Edelenbos & Klijn, 2009).

3. Two configurations were associated with dissatisfaction. In nine events (see Table 5), “most of which involved meeting deadlines and getting things built in line with the project plan, encouraged by the structure of the [public-private partnership] contract, A-Lanes’ managers responded to social events with internally-oriented policies” (Verweij, 2015a, p. 196). Together with the second finding from Table 4, this showed that social events require externally oriented management responses and not internally oriented ones. The finding reiterated the importance of actively engaging with the external stakeholder environment to solve their
worries and frustrations. If the worries and frustrations of the stakeholder environment are dealt with by not involving the stakeholders and instead by “ignoring” them, this can backfire on the project implementation, causing even more challenging events for the project managers.

4. In seven events, A-Lanes A15 responded autonomously and with internally oriented management (see Table 5). This further strengthened our finding that internally oriented management is not preferred. It also pointed out that autonomous management responses often are associated with dissatisfaction. Combined with the first finding (see Table 4), this indicated that a cooperative strategy is preferred over a non-cooperative one. In the seven cases, the private contractor A-Lanes A15 tried to deal with the events by itself. As a consequence, the stakeholder management capacities of the public client Rijkswaterstaat were not put to optimal use. That is, prior to the implementation of the megaproject, Rijkswaterstaat had built relationships with the stakeholder environment, but these ties were not optimally utilized.

We have now described in a nutshell the main analytical process in QCA as well as shared our practical experiences with using the method. Next, we will summarize some final lessons learned from the application of QCA in our study and will offer suggestions for further reading.

Conclusions and Lessons Learned

We are convinced of the value of QCA in evaluation studies. We would like to stress that it should not be rashly bolted on top of a different approach or be used as an afterthought. Indeed, the whole process from the definition of the research question to the writing of the final conclusions should be thoroughly informed by the logic of QCA. The underlying motive for this recommendation is that QCA has a different understanding of how causality (set-theoretic, complexity) can be understood and researched compared with many other methods. QCA can be used very well with other analytical methods (such as regression-based methods or process tracing), and frameworks and guidelines are in fact being introduced for this (e.g., Rohlfing & Schneider, 2013; Schneider & Rohlfing, 2016), but it is important to be aware of the particular way in which QCA produces results and what the results mean, so as to avoid flawed interpretations and conclusions.

A second lesson that follows from the first is that the researcher should not worry about adjusting, for example, calibrations or the selection of conditions during the application of the method. QCA is a decidedly iterative method that encourages the researcher to make these adjustments, as long as it is done transparently (and again: QCA allows for that). In fact, it is through the iterations that the researcher can gain an increasingly better understanding of the cases being studied.

A third lesson relates to the fact that QCA, by being complexity-informed, can sometimes render seemingly ambiguous outcomes. This is neither a weakness of the method nor of the researcher per se; it is simply a strong expression of the complexity of actual projects. Stronger, still, there are quite a few more popular methods that generate seemingly less ambiguous conclusions but that hide or ignore the causal complexity.
underneath it, which is not informative at all. By implication, this means that results derived from a QCA are usually provisional and that better explanations may be found in future studies including different or more cases and/or other conditions. The results serve an important purpose in learning from evaluations (Verweij, 2017). We experienced that evaluators in particular like clear-cut answers, if only for reasons of accountability. However, the reality of projects is such that there is not always one variable that explains all the problems encountered in projects. QCA highlights that fact.

That being said, we also learned that the method has some drawbacks. Looking at planning and implementation processes, we saw how situations could change over time; projects render different outcomes depending on what point in time the evaluation is carried out (see, for example, Busscher et al., 2017). The strength of QCA is in its ability to uncover the causal complexity of, essentially, static systems. It has more trouble to account for time. In addition, the method will struggle with large \(n\)-studies in the sense that larger samples prohibit the researcher from understanding every individual case within the sample thoroughly. Consequently, the iterative movements will make little sense. As the method was not originally conceived to deal with time or large samples, this is not necessarily a major problem. However, we have seen research that ignores these points, which makes one question the suitability of QCA in such studies.

Deploying QCA for evaluation also implies that there is ample room for evaluators to treat the method as the proverbial box of chocolates, that is, to cherry-pick those items that match the expected outcome. It is important to remain transparent throughout the research process so that peers can follow each step thoroughly, and to communicate to the sponsors that the method should be used conscientiously (see for more lessons on QCA in practical evaluation settings, for example, Gerrits & Verweij, 2018; Pattyn, Molenveld, & Befani, 2017).

Last but not least, QCA is labor-intensive. It requires a good understanding of (considerable amounts of) data and conscientious processing of such data. This should not be underestimated. QCA is not a press-button technique. QCA-software can and should be used to avoid errors in the formal comparative parts of the method (Schneider & Wagemann, 2010), but it cannot replace the role the researcher has in the research process. QCA involves the quantification of data to allow systematic comparison and create transparency, but the phenomenon studied with the method cannot be reduced to mere numbers. The researcher plays an important role in decisions taken in the iterative course of the method, related to inter alia the calibration, dealing with contradictions or logical remainders, and the interpretation of the results.

**Further QCA Guidance**

For extensive introductions to QCA, we refer to the available QCA-textbooks. Schneider and Wagemann (2012) provide a detailed overview of and guide to the technical details of QCA (see, for a review, Verweij, 2013). Rihoux and Ragin’s (2009) textbook focuses additionally on QCA as a research approach, addressing also issues of research design including case selection and the selection of conditions. Our recent textbook (Gerrits & Verweij, 2018) introduces and explains QCA for infrastructure project research in particular. The
book aims to be accessible for researchers and evaluators with no prior knowledge of the set-theoretical foundations of QCA, or with little experience with qualitative data collection, analysis, and interpretation. It also features an overview of six strategies to incorporate the time dimension in a QCA-study and it discusses a range of practical issues that evaluators applying QCA may run into. Finally, pointers on the calibration of qualitative data specifically can be found in the recent work of inter alia Basurto and Speer (2012) and De Block and Vis (2018).

Exercises and Discussion Questions

1. The research we presented above as an example featured three constituent conditions and one outcome. We made the claim that these conditions cannot be seen as independent from one another. This is expressed in the truth table. Explain in which three ways QCA’s notion of causality is different from a linear understanding of causality.

2. Using the example, develop an answer of how it is possible that single conditions can contribute to different outcomes (i.e., satisfaction and dissatisfaction) when combined with other conditions.

3. We said that, in QCA, the conditions are part of the case, whereas variables are external to the case. Explain in your own words what this means exactly and why it matters for the types of conclusions reached.

4. The two main ways to calibrate conditions and outcomes in QCA are crisp-set QCA (where cases are scored with either a 0 or a 1) or fuzzy-set QCA (where cases are scored on a scale that goes from 0 to 1). Discuss in which instances crisp-set QCA would be preferred over fuzzy-set QCA (and the vice versa).

5. We have made the point that QCA can be used by evaluators to cherry-pick and calibrate the data in such a way that their desired conclusions are reached. Discuss what possible measures could be used to prevent this.

Further Reading


Web Resources

See www.compasss.org for an extensive overview of QCA books, articles, software, training courses, and conferences.

References


