What is Qualitative Comparative Analysis (QCA)?

Charles C. Ragin
Department of Sociology and
Department of Political Science
University of Arizona
Tucson, AZ 85721 USA

www.fsqca.com
www.compasss.org
www.u.arizona.edu/~cragin
cragin@u.arizona.edu

Background

QCA's home base is *comparative sociology/comparative politics*, where there is a strong tradition of case-oriented work alongside an extensive and growing body of quantitative cross-national research.

The case-oriented tradition is much older and is populated largely by *area and country experts*. In contrast to the situation of qualitative researchers in most social scientific subdisciplines, these case oriented researchers have *high status,* primarily because their case knowledge is useful to the state (e.g., in its effort to maintain or enhance national security) and other corporate actors.

Case-oriented researchers are often critical of quantitative cross-national researchers for ignoring the gap between the results of quantitative research and what is known about specific cases. They also have little interest in the abstract, high-level concepts that often characterize this type of research and the wide analytic gulf separating these concepts from case-level events and processes.

QCA, plain and simple, attempts to bridge these two worlds. This attempt has spawned methodological tools which are useful to social scientists in general.

Four (relatively abstract) answers to the question, “What is QCA?”

1. QCA is a method that bridges qualitative and quantitative analysis:

Most aspects of QCA require **familiarity with cases**, which in turn demands in-depth knowledge. At the same time, QCA is capable of pinpointing decisive cross-case patterns, the usual domain of quantitative analysis. QCA’s examination of cross-case patterns respects the **diversity** of cases and their **heterogeneity** with regard to their different causally relevant conditions and contexts by comparing cases as configurations.

2. QCA provides powerful tools for the analysis of causal complexity:

With QCA, it is possible to study “**INUS**” conditions—causal conditions that are insufficient but necessary parts of causal recipes which are themselves unnecessary but sufficient. In other words, using QCA it is possible to assess causation that is very complex, involving **different combinations of causal conditions** capable of generating the same outcome. This emphasis contrasts strongly with the “net effects” thinking that dominates conventional quantitative social science. QCA also facilitates a form of counterfactual analysis that is grounded in case-oriented research practices.

3. QCA is ideal for small-to-intermediate-N research designs:

QCA can be usefully applied to research designs involving small and intermediate-size *N*s (e.g., 5-50). In this range, there are often **too many cases** for researchers to keep all the case knowledge "in their heads," but **too few cases** for most conventional statistical techniques.

4. QCA brings set-theoretic methods to social inquiry:

QCA is grounded in the analysis of set relations, not correlations. Because **social theory is largely verbal** and verbal formulations are largely set theoretic in nature, QCA provides a closer link to theory than is possible using conventional quantitative methods. (Most conventional quantitative methods simply parse matrices of bivariate correlations.) Note also that important causal relations, **necessity and sufficiency**, are indicated when certain set relations exist: With necessity, the outcome is a subset of the causal condition; with sufficiency, the causal condition is a subset of the outcome. With INUS conditions, cases with a specific combination of causal conditions form a subset of the cases with the outcome. Only set theoretical methods are well suited for the analysis of causal complexity.

The bare-bones basics of crisp-set QCA

Phase 1: Identify relevant cases and causal conditions

1-1. Identify the **outcome** that you are interested in and the cases that exemplify this outcome. Learn as much as you can about these “**positive**” cases.

1-2. Based on #1, identify **negative** cases—those that might seem to be candidates for the outcome but nevertheless failed to display it (“negative” cases). Together #1 and #2 constitute the **set of cases relevant to the analysis**.

1-3. Again based on #1, and relevant theoretical and substantive knowledge, identify the major **causal conditions** relevant to the outcome. Often, it is useful to think in terms of different causal “recipes”—the various combinations of conditions that might generate the outcome.

1-4. Try to **streamline** the causal conditions as much as possible. For example, combine two conditions into one when they seem “substitutable.”

Example:

1. Identify positive instances of mass protest against austerity measures mandated by the International Monetary Fund (IMF) as conditions for debt renegotiation (“conditionality”). Peru, Argentina, Tunisia, . . .

2. Identify negative cases: for example, debtor countries that were also subject to IMF conditionality, but nevertheless did not experience mass protest. Mexico, Costa Rica, . . .

3. Identify relevant causal conditions: severity of austerity measures, degree of debt, living conditions, consumer prices, prior levels of political mobilization, government corruption, union strength, trade dependence, investment dependence, urbanization and other structural conditions relevant to protest mobilization. One recipe might be severe austerity measures combined with government corruption, rapid consumer price increases and high levels of prior political mobilization.

4. Streamlining: Based on case knowledge, the researcher might surmise that high levels of trade dependence and high levels of investment dependence are substitutable manifestations of international economic dependence and therefore create a single condition from these two, using logical “or.”

Phase 2: Construct the truth table and resolve contradictions

2-1. Construct a “truth table” based on the causal conditions specified in phase 1 or some reasonable subset of these conditions (e.g., using a recipe that seems especially promising). A truth table sorts cases by the combinations of causal conditions they exhibit. All logically possible combinations of conditions are considered, even those without empirical instances.

2-2. Assess the consistency of the cases in each row with respect to the outcome: Do they agree in displaying (or not displaying) the outcome? A simple measure of consistency for crisp sets is the percentage of cases in each row displaying the outcome. Consistency scores of either 1 or 0 indicate perfect consistency for a given row. A score of 0.50 indicates perfect inconsistency.

2-3. Identify contradictory rows. Technically, a contradictory row is any row with a consistency score that is not equal to 1 or 0. However, it is sometimes reasonable to relax this standard, for example, if an inconsistent case in a given row can be explained by its specific circumstances.

2-4. Compare cases within each contradictory rows. If possible, identify decisive differences between positive and negative cases, and then revise the truth table accordingly.

Example, using one possible recipe as a starting point:

Row#	Prior mobiliz.?	Severe austerity?	Gov't corrupt?	Rapid price rise?	Cases w/ protest?	Cases w/o protest	Consis-tency
1	0 (no)	0 (no)	0 (no)	0 (no)	0	0	??
2	0 (no)	0 (no)	0 (no)	1 (yes)	0	0	??
3	0 (no)	0 (no)	1 (yes)	0 (no)	0	4	0.0
4	0 (no)	0 (no)	1 (yes)	1 (yes)	1	5	0.167
5	0 (no)	1 (yes)	0 (no)	0 (no)	0	0	??
6	0 (no)	1 (yes)	0 (no)	1 (yes)	4	0	1.0
7	0 (no)	1 (yes)	1 (yes)	0 (no)	0	0	??
8	0 (no)	1 (yes)	1 (yes)	1 (yes)	5	0	1.0
9	1 (yes)	0 (no)	0 (no)	0 (no)	0	3	0.0
10	1 (yes)	0 (no)	0 (no)	1 (yes)	1	7	0.125
11	1 (yes)	0 (no)	1 (yes)	0 (no)	0	10	0.0
12	1 (yes)	0 (no)	1 (yes)	1 (yes)	0	0	??
13	1 (yes)	1 (yes)	0 (no)	0 (no)	1	5	0.167
14	1 (yes)	1 (yes)	0 (no)	1 (yes)	6	0	1.0
15	1 (yes)	1 (yes)	1 (yes)	0 (no)	6	2	0.75
16	1 (yes)	1 (yes)	1 (yes)	1 (yes)	8	0	1.0

Table Notes:

a. This table has five rows without cases (1, 2, 5, 7, 12). In QCA, these rows are known as “remainders.” Having remainders is known as “limited diversity.”

b. There are seven noncontradictory rows, three that are uniform in not displaying the outcome (consistency = 0.0; rows 3, 9, 11) and four that are uniform in displaying the outcome (consistency = 1.0; rows 6, 8, 14, 16).

c. The remaining four rows are contradictory. Three are close to 0.0 (rows 4, 10, 13), and one is close to 1.0 (row 15).

d. Suppose that the three (unexpected) positive cases (one each in rows 4, 10, 13) are all cases of contagion—a neighboring country with severe IMF protest spawned sympathy protest in these countries. These contradictory cases can be explained using case knowledge, showing that these instances are irrelevant to the recipe in question. Thus, these three cases can be safely set aside.

e. Suppose the comparison of the positive and negative cases in row 15, reveals that the (unexpected) negative cases all had severely repressive regimes. This pattern suggests that having a not-severely-repressive regime is part of the recipe and that the recipe has five key conditions, not four. The revised truth table follows. (To simplify the presentation, only rows with cases are shown.)

Prior mobiliz.?	Severe austerity?	Gov’t corrupt?	Rapid price rise?	Not re-pressive?	Cases w/ protest?	Cases w/o protest	Consis-tency
0 (no)	0 (no)	1 (yes)	0 (no)	0 (no)	0	4	0.0
0 (no)	0 (no)	1 (yes)	1 (yes)	0 (no)	0	5	0.0
0 (no)	1 (yes)	0 (no)	1 (yes)	0 (no)	4	0	1.0
0 (no)	1 (yes)	1 (yes)	1 (yes)	1 (yes)	5	0	1.0
1 (yes)	0 (no)	0 (no)	0 (no)	0 (no)	0	3	0.0
1 (yes)	0 (no)	0 (no)	1 (yes)	1 (yes)	0	7	0.0
1 (yes)	0 (no)	1 (yes)	0 (no)	0 (no)	0	10	0.0
1 (yes)	1 (yes)	0 (no)	0 (no)	1 (yes)	0	5	0.0
1 (yes)	1 (yes)	0 (no)	1 (yes)	0 (no)	6	0	1.0
1 (yes)	1 (yes)	1 (yes)	0 (no)	1 (yes)	6	0	1.0
1 (yes)	1 (yes)	1 (yes)	0 (no)	0 (no)	0	2	0.0
1 (yes)	1 (yes)	1 (yes)	1 (yes)	1 (yes)	8	0	1.0

Notice that there are no contradictions. The three cases of contagion have been removed, and the contradictory cases in row 15 of the previous table have been resolved.

The “full” version of this truth table would have 32 rows and 20 “remainders” (causal combinations lacking cases).

Phase 3: Analyze the truth table

fsQCA software can be used to analyze truth tables like the two just shown. The goal of the analysis is to specify the different combinations of conditions linked to the selected outcome, based on the features of the positive cases that consistently distinguish them from the negative cases.

3-1. The first part of the fsQCA algorithm compares rows of the truth table to identify matched pairs. For example, these two rows both have the outcome, but differ by ONLY ONE causal conditions, providing an experiment-like contrast:

Prior mobiliz.?	Severe austerity?	Gov't corrupt?	Rapid price rise?	Not re-pressive?	Cases w/ protest?	Cases w/o protest	Consis-tency
1 (yes)	1 (yes)	1 (yes)	0 (no)	1 (yes)	6	0	1.0
1 (yes)	1 (yes)	1 (yes)	1 (yes)	1 (yes)	8	0	1.0

This paired comparison indicates that if prior mobilization, severe IMF austerity, government corruption, and a nonrepressive regime coincide, it doesn't matter whether there are also rapid price increases; protest will still erupt. The term that differs is eliminated, and a single, simpler row replaces the two rows shown.

3-2. This process of bottom-up paired comparison continues until no further simplification is possible. Only rows with the outcome are paired; only one condition may differ in each paired comparison. The one that differs is eliminated.

3-3. Rows without cases ("remainders") also may be used to aid the process of simplifying the patterns. For example, there are no instances of the second row listed below. However, based on substantive and theoretical knowledge, it is reasonable to speculate that if such cases existed, they would be positive instances of IMF protest, just like the empirical cases they are paired with:

Prior mobiliz.?	Severe austerity?	Gov't corrupt?	Rapid price rise?	Not Re-pressive?	Cases w/ protest?	Cases w/o protest	Consis-tency
0 (no)	1 (yes)	0 (no)	1 (yes)	0 (no)	4	0	1.0
0 (yes)	1 (yes)	1 (yes)	1 (yes)	0 (no)	0	0	??

The reasoning is as follows: (1) the remainder case resembles the empirical cases above it in every respect except one; (2) the one difference (the remainder case has government corruption) involves a condition that should only make IMF protest *more likely*; (3) therefore, the remainder case, if it existed, would display IMF protest, just like the empirical cases. This pairing allows the production of a logically simpler configuration, eliminating the absence of government corruption as a possible (INUS) ingredient.

This example shows how fsQCA incorporates a form of counterfactual analysis that parallels practices in qualitative case-oriented research.

3-4. The process of paired comparisons culminates in a list of causal combinations linked to the outcome. fsQCA then selects the smallest number of these that will cover all the positive instances of the outcome.

3-5. fsQCA presents three solutions to each truth table analysis: (1) a “complex” solution that avoids using any counterfactual cases (rows without cases—“remainders”); (2) a “parsimonious” solution, which permits the use of any remainder that will yield simpler (or fewer) recipes; and (3) an “intermediate” solution, which uses only the remainders that survive counterfactual analysis based on theoretical and substantive knowledge (which is input by the user). Generally, intermediate solutions are best. For the table just presented, the three solutions are:

Complex: SA•gc•PR•nr + PM•SA•GC•NR + SA•GC•PR•NR

Parsimonious: GC•NR + SA•PR

Intermediate: PM•SA•GC•NR + SA•PR

(Multiplication indicates set intersection—combined conditions; addition indicates set union—alternate combinations.)

Because they are logical statements, these two recipes for IMF protest can be factored. For example, the intermediate solution can be factored to show that severe austerity (SA) is present in both:

SA•(PR + PM•GC•NR)

The expression indicates that IMF protest erupts when severe austerity (SA) is combined with either (1) rapid price increases (PR) or (2) the combination of prior mobilization (PM), government corruption (GC), and nonrepressive regime (NR).

Phase 4: Evaluate the Results

4-1. Interpret the results as causal recipes. Do the combinations make sense? What causal mechanisms do they imply or entail? How well do they relate to existing theory? Do they challenge or refine existing theory?

4-2. Identify the cases that conform to each causal recipe. Often some cases will conform to more than one recipe and sometimes there are more cases that combine two (or more) recipes than there are "pure" instances. Do the recipes group cases in a meaningful way? Do the groupings reveal aspects of cases that had not been considered before?

4-3. Conduct additional case-level analysis with an eye toward the mechanisms implied in each recipe. Causal processes can be studied only at the case level, so it is important to evaluate them at that level.

The real test of any QCA result is how well it connects to cases. In this hypothetical application of QCA, there are no cases to connect to. Still, it is worth noting that the analysis started out as an examination of a single recipe, but ended up with two, each with very different implied mechanisms. The mechanisms implied by "severe austerity combined with rapidly rising prices" are substantially different from those implied by "severe austerity combined with prior mobilization, government corruption, and a nonrepressive regime."

Crisp versus fuzzy sets

Most introductions to QCA focus on crisp sets, just as this one does. Crisp sets are simple and straightforward and thus easy to present. However, social scientists (1) are often interested in phenomena that vary by level or degree (e.g., degree of membership in the set of democratic countries), and (2) dislike dichotomizing.

Fortunately, the procedures described here can be duplicated using fuzzy sets, which allow membership scores to vary from 0.0 to 1.0. Additional issues involved in the use of fuzzy sets include:

1. Calibrating the degree of membership in sets

2. Calculating degree of membership in a configuration, conceived as an intersection of fuzzy sets

3. Analyzing fuzzy subset relations

4. Constructing a crisp truth table summarizing the results of the fuzzy set analyses

What is QCA? FREE software

For *crisp-set* analysis (dichotomies):

- **QCA 2.0** (DOS program; www.compasss.org)
- **QCA 3.0** (DOS program; www.fsqca.com)
- **TOSMANA** (MS Windows program; www.tosmana.net)
- **fsQCA** (MS Windows program; www.fsqca.com)

For *multi-value sets* (multichotomies; not ordinal)

- **TOSMANA** (MS Windows program; www.tosmana.net)

For *fuzzy-set* analysis (membership ranges from 0 to 1; ordinal, interval, ratio)

- **fsQCA** (MS Windows program; www.fsqca.com)

Related efforts:

- **QCA in R**: new module for the R programming language developed by Adrian Dusa (Romanian Social Data Archive, Bucharest); see www.compasss.org. Only for users familiar with R.
- **STATA QCA**: new module for STATA developed by Kyle Longest (University of North Carolina) and Steve Vaisey (University of California, Berkeley)—good start, but still a long way to go. It's difficult to use and includes some unfortunate default procedures.

What is QCA? An epistemic community

It’s easy to join, even peripherally.

Browse www.compasss.org and www.fsqca.com.

Attend a Compasss event (most are held in Belgium).

Attend a QCA training session or workshop. For example, Benoit Rihoux and Carsten Schneider are offering a QCA course at the ECPR Summer School in Methods and Techniques in Ljubjiana, August 4-16, 2008. I am offering a mini-course on QCA and fuzzy sets at the University of Arizona from August 26—September 18, 2008 (four weeks in the hot desert!).

Add your name to the www.compasss.org email list; go to the website and click “small-N people.”

Brief bibliography

Published Books:

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

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Smithson, Michael, & Verkuilen, Jay. 2006. *Fuzzy Set Theory. Applications in the Social Sciences*. London: Sage Publications.

Forthcoming Books:

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rihoux, Benoit & Ragin, Charles C. (eds). 2008. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. London and Thousand Oaks, CA: Sage.