Avoiding Common Errors in QCA: A Short Guide for New Practitioners

Claude Rubinson
University of Houston–Downtown

Lasse Gerrits
Otto–Friedrich–Universität Bamberg

Roel Rutten
Tilburg University

Thomas Greckhamer
Louisiana State University

July 22, 2019

QCA is increasingly being adopted by researchers from a variety of disciplines and industries. As is only natural when learning a new methodological technique and approach, not all of these applications are free of errors. This is a list of common errors that we have seen when reviewing QCA projects. We provide this list as a "cheat sheet" in the hopes of helping researchers improve their projects and avoid some of the most common mistakes encountered in the wild.

For each item we briefly explain what the error is and why it is considered a problem. We also describe possible remedies whenever applicable. The order of items follows the logic of the research cycle. These discussions are by no means comprehensive and we strongly recommend that the reader review the following texts for more information:

- Ragin, C.C. (1987). *The Comparative Method*
- Ragin, C.C. (2008). *Redesigning Social Inquiry*
- Rihoux, B. and C.C Ragin. (2009). *Configurational Comparative Methods*
- Byrne, D. and C.C. Ragin. (2009). *Sage Handbook of Case-Based Methods*
- Schneider, C.Q. and C. Wagemann. (2012). *Set-Theoretic Methods for the Social Sciences*

Also, for discussions of recommendations for standards of good practice in QCA research, readers are referred to the following articles:

- Schneider, C. Q., and Wagemann, C. (2010). Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comparative Sociology*, 9, 397–418.

- Greckhamer, T., Furnari, S., Fiss, P., and Aguilera, R. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization*, 16, 482–495.
- Rubinson, C. (2019). Presenting qualitative comparative analysis: Notation, tabular layout, and visualization. *Methodological Innovations*, May–August, 1–22, DOI:10.1177/2059799119862110.

A List of Common Errors in QCA

Approaching QCA

1. Substituting QCA for small-N statistical analysis

   An important advantage of QCA is that it can be used to analyze small- and medium-N data (as well as large-N). However, this does not mean that it is always appropriate for small- and medium-N analysis. Keep in mind that conventional statistical analysis and QCA are designed to answer different types of research questions.

2. Using variable-oriented language

   Causality in QCA is expressed as conditions being sufficient or necessary for the outcome, not in terms of independent variables having an effect on a dependent variable. That is, QCA focuses on the causes of effects, rather than the effects of causes. This is particularly important when verbalizing the sufficiency of a conjunction of multiple conditions. From a configurational perspective, it is not that having friends and having a job have independent, positive effects on happiness. Rather, it is the combination of having both friends and a job that is sufficient for being happy. (Note also, that singular tense is used here because there is a single combination, comprised of multiple conditions.)

3. Asserting causation without identifying mechanisms

   QCA is a descriptive, not inferential, technique. If you wish to establish causation, you need to identify the underlying mechanism(s) at work, drawing upon substantive and/or theoretical knowledge. (This is why QCA is frequently applied in conjunction with in-depth case studies.)

Setting up research

4. Including only cases for which the outcome is present

   QCA is a method for studying diversity. This means that your data set must include observations where the outcome is absent. It also means that your truth table must include rows where the outcome is False (i.e., consistency is less than your specified threshold). If all of your (non-remainder) truth table rows are consistent with the presence of the outcome, reconsider your calibration strategy. It may be that you have miscalibrated one or more conditions, or that you need to recalibrate to a more specific target set (e.g., from "rich school district" to "very rich school district").

5. Using nouns to name conditions instead of adjectival phrases

   Nouns refer to variables; a condition must have an adjective attached. For example, "GDP" is a variable while "Developed country" is a condition. "Income" is a variable; "Very rich individual" is a condition. "Years of education" is a variable; "Highly educated person" is a condition. In each of these examples, only the adjectival phrase refers to a set in which observations may have membership. Clarifying the adjectival phrase for each of your conditions will help with calibration.

Calibration

6. Using symmetric calibrations

   It is important to remember that calibration is asymmetric. That is: the negation of "rich" is not "poor" but, rather, "not rich." When calibrating your conditions, you need to carefully consider what is meant by "fully in" (fuzzy score=1.0) and "fully out" (fuzzy score=0.0). "Fully out" is only the opposite of "fully in" in the case of true dichotomies, which is rare (e.g., "biologically male" is equal to "not biologically female" if and only if we ignore intersex conditions).

   Other examples:

   (a) the negation of "developed country" is not "under-developed country" but, rather, "not-developed country"

   (b) the negation of "large company" is not "small company" but, rather, "not-large company"

   (c) the negation of "happy family" is not "sad family" but, rather, "not-happy family"

7. Calibrating to 0.5

   When an observation is assigned a fuzzy-set score of 0.5 for a given condition, that observation will have equal membership in all truth table rows but will not have maximum membership in any truth table row. It is unclear what this score means substantively because the crossover point is the point of maximum ambiguity, where the observation is neither in nor out of the target set. The case will therefore appear to drop out of the analysis, as it doesn't belong to any corner of the vector space. If you have many observations scoring at (or close to) 0.5, it often means that there is something wrong with your calibration strategy. (This may be a consequence of using symmetric calibrations or conceiving of your conditions as variables rather than sets, discussed above).

8. Failing to explain calibrations

   It is crucial that you explain what your calibrations mean substantively. It is not sufficient to simply report the values that correspond to, e.g., 0.0, 0.5, and 1.0. You also need to explain why, e.g., 14+ years of education corresponds to "fully in the set of educated people."

9. Mechanistic calibration

   Successful calibration requires one to carefully reflect upon the nature of one's measures and their meaning. Automated and statistical clustering/rescaling techniques rarely produce useful calibrations. Also avoid simply adopting another researcher's calibrations without first considering how they may need to be adjusted for your particular project.

10. Using a measure of central tendency and/or the variable's distribution to calibrate

    In general, you do not want to use the variable's distribution as the basis for your calibration. Having a higher than average income does not mean that one is rich. Instead, you need to rely upon substantive and theoretical knowledge of your domain in order to develop meaningful calibrations. If substantive and theoretical knowledge does not exist, you need to get to know your cases better. If you nevertheless use the variable's distribution as a basis for your calibration, make this clear when naming the condition: "Above average family size" or "Below average family size."

11. Using the full range of a Likert-type scale or index to calibrate

    This is not always an error but often is. When calibrating a Likert-type scale or index, you should not automatically assign the bottom value to 0.0, the middle value to 0.5, and the maximum value to 1.0. Instead, you should carefully consider the meaning of the scale. In particular, consider whether some values may need to be collapsed together. For example, with a 7-point scale, it may be the case that scores 1-3 are "fully out" of the target set, 4 is "more out than in," 5 is "more in than out" and 6-7 are "fully in." The precise calibration depends upon the what the scale means.

    You also should not automatically use the same calibration strategy for all conditions. Just because you have multiple 7-point measures does not mean that you should calibrate them all in the same way. Instead, think about what each item means substantively.

Analysis and interpretation

12. Using a sufficiency consistency threshold < 0.80

    In *Redesigning Social Inquiry*, Ragin states that the consistency threshold for sufficiency testing should not be less than 0.75 and recommends a threshold of at least 0.85, especially for macro-level data. Most QCA reviewers today anticipate a sufficiency consistency threshold of at least 0.8 and expect a substantive and/or theoretical justification for a lower threshold.

13. Ignoring very low unique coverage scores

    Very low unique coverage scores indicate substantial overlap among your sufficiency recipes. This often indicates that you really have a single recipe with slight variants due to substitutable conditions. To understand the nature of the overlap, begin by identifying the observations that belong to the overlapping recipes.

14. Not running a separate QCA for the negation of the outcome

    If you wish to explain both the presence and the absence of the outcome, as is common in QCA, you need to conduct a separate analysis for each. Recall that QCA is asymmetric and what explains the presence of the outcome may be different than what explains its absence.

15. Neglecting to conduct a necessity test

    QCA is not only about sufficiency. It is important to also test for necessity, unless you have a theoretical or methodological justification for not doing so.

16. Analyzing conditions that are always present

    If a condition is always present, for all observations in your data set, it has no explanatory power because it is necessary for the presence of the outcome and also for the absence of the outcome. There are two ways of thinking about such a condition. It may be a "trivial necessary condition" (e.g., oxygen is necessary for both war and peace) or it may be a scope condition that describes your sample. Alternatively, it could simply indicate that you have miscalibrated and need to revise your calibration of that condition.

17. Ignoring low solution coverage scores

    A low solution coverage indicates that there are many instances of the outcome that are not explained by your model. Whether that is problematic needs to be discussed. It may be that your model is poor or it may be that your model simply does not explain everything and that further research is needed. Either way, a low coverage score should not simply be ignored.

18. Ignoring equifinality because coverage is low

    One of QCA's strengths is that it can identify when there are multiple pathways to the same solution (equifinality). However, researchers sometimes ignore pathways with low coverage. Just because a pathway only explains a relatively small number of cases doesn't mean that it is unimportant. Just because most instances of lung cancer are due to smoking does not mean that we should not study other causes.

19. Failing to interpret the results

    Do not simply conclude your project by presenting your Boolean solution(s) and their consistency/coverage scores. Rather, you must explain what your results mean. A statement of necessity or sufficiency verbalizes a cross-case regularity that must then be interpreted into a causal mechanism. For example, if having both friends and a job is sufficient for being happy, this indicates that having a social network to satisfy your emotional needs and having an income to satisfy your material needs produces happiness. Return to your cases to assess the substantive meaning of your empirically observed cross-case regularity. Remember that in QCA causation is established through substantive knowledge, not empirical metrics.

Author biographies

Claude Rubinson (corresponding author; `rubinsonc@uhd.edu`) is Associate Professor of Sociology at the University of Houston–Downtown (Houston, Texas, USA) where he teaches courses on research methods, social inequality, and culture. His research program has three tracts: methodological research on developing formal methods of qualitative research, focusing on QCA; sociological research on the relationship between the global political-economy and aesthetic form; and health services research on the conditions that facilitate/hinder the success of interventions designed to improve patient outcomes.

Prof. Dr. Dr. Lasse Gerrits (`lasse.gerrits@uni-bamberg.de`) is professor in political science. His research focuses on the nexus of technology, policy and politics. He is interested in methodological issues with regard to social complexity and the ways in which this complexity can be unboxed. He has published extensively on this topic, among others in relationship to qualitative comparative analysis, critical realism, social complexity and qualitative methods.

Dr. Roel Rutten (`r.p.j.h.rutten@uvt.nl`) researches the geography of knowledge creation and the organization of knowledge creation. He works at the School of Social and Behavioural Sciences of Tilburg University (the Netherlands) and at European Regional Affairs Consultants (ERAC).

Thomas Greckhamer (`tgreck@lsu.edu`) is the Catherine Rucks Professor of Management in the E.J. Ourso College of Business at Louisiana State University. His research interests are at the intersection of organization studies, strategic management, and research methods, focusing on empirical applications of as well as theoretical and methodological contributions to set-theoretic and qualitative approaches.