Regular Article

International Journal of Qualitative Methods
Volume 22: 1–10

DOI: 10.1177/16094069231182634
journals.sagepub.com/home/ijq
Sage

Qualitative Comparative Analysis: Search Target, Reflection on the Top-Down Approach, and Introduction of the Bottom-Up Approach

Haien Ding[1]

Abstract
Based on the INUS theory of causality, the search target of qualitative comparative analysis (QCA) is to find all the minimally sufficient conditions for the outcome's occurrence in a data set, where the condition's *sufficiency*, the *necessity* of the condition's components, and the *completeness* of the solution are three core requirements. However, QCA's current top-down approach, which relies on a truth table and Boolean minimization, cannot meet the main objective of QCA. Conditions generated by the top-down approach can be insufficient for the outcome or contain unnecessary components that can be removed. We found evidence supporting our arguments by examining the correctness of top-down QCA in Study 1. Then, we show that QCA can also proceed with a "bottom-up" search strategy in sufficiency analysis, similar to coincidence analysis (CNA). We contrast solutions of the top-down and bottom-up QCA approaches by analyzing a simulated crisp-set data set in Study 2 and a real-world fuzzy-set data set in Study 3. Both results show that only the bottom-up approach can produce all the minimally sufficient conditions. We contribute to the ongoing debate pertain QCA solution types and QCA algorithms by critically evaluating the limitations of QCA's top-down approach and introducing a bottom-up approach for QCA.

Keywords
qualitative comparative analysis (QCA), coincidence analysis (CNA), configurational comparative methods, boolean minimization, Quine–McCluskey (QMC) algorithm, consistency cubes (CCubes) algorithm, causal complexity

As a methodology different from regression analytical methods, qualitative comparative analysis (QCA) has gained considerable attention and quickly spread to all social science fields (Ragin, 1987; Ragin et al., 1984). In recent years, sociology, political science, public health, business, and management scholars have extensively employed QCA to make causal inferences and develop configurational theories (Greckhamer et al., 2018; Hanckel et al., 2021; Misangyi et al., 2017; Woodside, 2013; 2014). Accompanied by QCA's quick diffusion and large applications is a growing number of ongoing debates over its methodology (Haesebrouck & Thomann, 2021; Krogslund et al., 2015; Schneider, 2018). A critical debate is about which QCA solution type is closest to the truth or can be causally interpreted (Haesebrouck, 2022; Haesebrouck & Thomann, 2021). A prerequisite of making this debate legitimate is that among three solution types, QCA can generate a correct solution (the question is which one). However, many QCA researchers ignore one possibility that the traditional QCA's top-down approach cannot guarantee to meet the requirements of a correct QCA solution, thus, making all three solution types prone to specific causal fallacies. For example, QCA might generate insufficient configurations that cannot guarantee the occurrence of the outcome. Configurations generated by QCA might also contain unnecessary components

[1]Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Oslo, Norway

Corresponding Author:
Haien Ding, Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Nydalsveien 37, Oslo 0484, Norway.
Email: haien.ding@bi.no

that can be removed without affecting the formers' status as sufficient configurations. In this article, we aim to discuss this possibility and critically evaluate the misfit between QCA's current top-down approach and its search target.

The remainder of the article is structured as follows: First, based on the INUS theory of causality, we clarify that the search target of QCA is to find all the minimally sufficient conditions of the outcome in a given data set because clarification on the QCA's main objective is a precondition for discussing QCA's correctness. Three criteria for the QCA solution's correctness are the condition's sufficiency, the necessity of the condition's components, and the completeness of the solution. Second, we illustrate how the top-down approach of QCA relies on a truth table and Boolean minimization to produce solutions. We explain why the top-down approach fails to achieve these requirements in generating solutions. We found evidence supporting our arguments in Study 1. Third, inspired by the algorithm of coincidence analysis (CNA; Baumgartner & Ambühl, 2020b), we introduce a new bottom-up search strategy of QCA that is specifically designed to generate solutions satisfying QCA's three criteria. Fourth, we provide two empirical demonstrations to show the differences in the derived QCA solutions between the traditional top-down and the new bottom-up approaches in Studies 2 and 3. Contrasts between the two algorithms show that the bottom-up approach is a better choice for generating all the minimally sufficient conditions in the sufficiency analysis of QCA.

The Search Target of QCA

QCA is often used to analyze causal complexity, which firmly rests on Mackie (1965, 1974)'s INUS theory of causation. INUS stands for "Insufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result" (Mackie, 1974, pp. 59–87). INUS theory of causality is the backbone of QCA, and the analytical search target of QCA should be to identify INUS conditions, which is explicitly agreed upon by most QCA researchers (Baumgartner, 2015; Mahoney & Acosta, 2021; Ragin, 2008; Schneider & Wagemann, 2012; Thiem, 2022). For example, in an illustrative example of what leads to a generous welfare state, Ragin (2008) claimed that four conditions combined to produce generous welfare states, and therefore, "each of the four causal conditions is an 'INUS' condition" (p. 154). Schneider and Wagemann (2012) argued that set-theoretical methods are very suitable for identifying INUS conditions, and they also clearly expressed that "QCA solution formulas are full of INUS conditions" (p. 79). Similarly, when introducing the methodology of QCA to the information management field, Pappas and Woodside (2021) also noted that QCA enables researchers to capture INUS conditions.

However, we still lack clear definitions of an INUS condition and causal relevance in QCA solutions, making it hard to evaluate whether QCA achieves its intended purpose. Although there are four adjectives in Mackie's theory of INUS causation, only two (i.e. Sufficient and Necessary) are crucial to define causal relevance (Haesebrouck, 2022). The I (i.e. Insufficient) and the U (i.e. Unnecessary) are not needed to describe what Mackie considers causally relevant conditions. The causally relevant conditions in the sufficiency analysis are "Necessary parts of Sufficient conditions (Haesebrouck, 2022, p. 10)."

Clarity on the concept of causal relevance enables us to propose three requirements to define a correct QCA solution. First, an identified condition (or configuration) should be *sufficient* for or guarantee the occurrence of the outcome (Duşa, 2019a): Its sufficiency score should pass the desired threshold (e.g. $\geq$0.8 is usually adopted sufficiency threshold). Second, any component of a sufficient condition (or configuration) should be *necessary* or indispensable. A condition (or component) is considered redundant or not necessary when it can be removed from a condition (or configuration) without changing the latter's status as a sufficient condition (Baumgartner, 2015; Haesebrouck, 2022). Unnecessary components cannot be causally interpretable because they do not make a difference in the outcome (Baumgartner & Thiem, 2017; Haesebrouck & Thomann, 2021). A condition conforming to the two requirements is minimally sufficient: It is sufficient for the outcome, while no proper part of it is sufficient (Baumgartner, 2015). Any component of a minimally sufficient condition cannot be removed without changing the latter's status as a sufficient condition. Third, a correct QCA solution should also achieve *completeness*, enabling us to get all the minimally sufficient conditions in a given data set (Mahoney & Acosta, 2021). As Mahoney and Acosta (2021) pointed out, all sufficient conditions must be identified if a causal interpretation can be applied in the solution set.

Although some scholars argue that an ideal QCA solution should be "minimally necessary disjunctions of minimally sufficient conditions (Baumgartner, 2008, p. 13; Graßhoff & May, 2001; Mahoney & Acosta, 2021, p. 12)." Under this standard, the whole solution set is both necessary and sufficient for the outcome (Mahoney & Acosta, 2021). However, finding a solution set that is both complete and necessary is extremely difficult in social science practice for at least two reasons (Mahoney & Acosta, 2021; Ragin, 2000). First, complex social phenomena have multiple causes that are hard to identify. Second, the real world only provides limited information for studying these complex phenomena (Mahoney & Acosta, 2021). Therefore, a more realistic analytical objective of QCA is to find all the minimally sufficient conditions for the outcome in a given data set (Baumgartner, 2015; Duşa, 2019a; Haesebrouck & Thomann, 2021), where sufficiency, necessity, and completeness are three core requirements.

Evaluating the Traditional Top-Down Approach

QCA adopts a top-down search strategy to generate solutions. It begins with constructing a truth table consisting of all logically possible combinations of the presence and absence of causal conditions. Therefore, a truth table contains 2^k rows, where k is the number of included conditions. Each unique row is often called a configuration or a complex condition. Commonly, some logically possible rows do not have corresponding empirical data, a phenomenon often referred to as limited diversity (Cooper & Glaesser, 2016; Schneider & Wagemann, 2012, p. 151). These rows that lack empirical cases (i.e. the case frequency is 0) are called logical remainders. Besides, an outcome column shows each configuration's outcome value, measuring set-theoretic consistency (i.e. sufficiency) between the configuration and the outcome (Ragin, 2008, p. 125; Thiem & Duşa, 2013). Based on a consistency threshold (usually at least 0.8), complex combinations can be coded as *positive* (1) if they pass the threshold or *negative* (0) if they do not. The subsequent Boolean minimization procedure will include only positive (or sufficient) configurations.

Then, QCA relies on logical minimization to remove redundancies from primitive sufficient conditions as much as possible. We can use an uppercase letter (e.g. A) to indicate the presence of the condition and a lowercase letter to indicate the absence of the condition (e.g. a) to illustrate the principle guiding logical minimization. If two sufficient configurations differ in only one condition-it is present in one configuration but absent in another configuration (e.g. "ABC" and "AbC"), then the condition "B" is considered redundant or unnecessary and should be removed because its presence or absence is irrelevant (or unnecessary) for the outcome (Schneider & Wagemann, 2012, p. 105). Finally, we can get a simpler but logically equivalent sufficient condition (i.e. "AC"). The simplified conditions after logical minimization are also called prime implicants (Duşa, 2019b, p. 203). Different coping strategies on logical remainders result in different QCA solution types: parsimonious solution (QCA-PS), intermediate solution (QCA-IS), and conservative/complex solution (QCA-CS). In summary, the objective of the truth table aims to guarantee the condition's sufficiency, while the purpose of logical minimization attempts to ensure the necessity of the condition's components.

However, using the three requirements to revisit the top-down approach of QCA, we find that the traditional top-down approach is not specifically designed to fulfill the requirements of sufficiency, necessity, and completeness. First, the QCA top-down approach only imposes the consistency cutoffs on the primitive complex conditions; it cannot guarantee that the sufficiency of a generated condition is also above the desired threshold after logical minimization (Baumgartner & Ambühl, 2020a). Accordingly, it is not surprising to find that QCA-IS and QAC-PS might generate conditions that are insufficient for the outcome in empirical research (Castaño et al., 2016). One potential reason is that QCA-IS and QAC-PS involve counterfactual analyses, thus, increasing the risk of getting insufficient conditions. Some scholars have admitted such a risk. For instance, Duşa (2019a) pointed out that "QCA-PS presents an insufficient solution as sufficient" (p. 7). Baumgartner (2021) also noted that in the redundancy-free (RF) approach, the QCA solution does not "guarantee the outcome's occurrence (p. 2)" and "does not mind if its models do not identify substantively sufficient conditions" (p. 2).

Second, although logical minimization appears able to remove unnecessary components, its procedure does not conform to the definition of necessity. According to Baumgartner (2015)'s definition, a component is considered a necessary part of a condition when it cannot be removed without changing the latter's status as a sufficient condition (Haesebrouck, 2022). This definition requires us to compare two conditions (i.e., one including a specific component and another removing that component) to decide whether a component is necessary or not. Continuing with the above example, we need to compare "ABC" and "AC" or "AbC" and "AC" to decide whether "B" or "b" is unnecessary. If "AC" is sufficient, then we can say that "ABC" and "AbC" contain unnecessary components because "AC" can guarantee the outcome's occurrence without "B" or "b." Therefore, using a procedure that does not conform to the definition of necessity, Boolean minimization cannot guarantee the removal of unnecessary components from sufficient conditions. For example, when "ABC" is sufficient but "AbC" is insufficient, Boolean minimization generates "ABC" as a minimally sufficient condition, ignoring the possibility that "AC" might also be sufficient. However, it is possible that subsets (e.g. "AbC") of a sufficient condition (e.g. "AC") might not be sufficient. In a real-world research context, researchers often operationalize a sufficiency relationship in an imperfect (but more realistic) way (e.g. a consistency score $\geq$0.8) because perfect sufficiency relationships rarely happen in the real world (Greckhamer et al., 2018; Ragin, 2008). However, an imperfect sufficiency relationship means a sufficient condition does not guarantee the outcome's occurrence in some situations. Therefore, a sufficient condition (e.g. "AC") might be insufficient when combined with other factors (e.g. "AbC"). In the traditional top-down approach, subsets of a sufficient condition could be insufficient and excluded from logical minimization. Therefore, a generated configuration might contain unnecessary components that can be removed.

Third, it is unclear how QCA can achieve the completeness of the solution set because the top-down approach is not specifically designed to achieve this requirement. QCA does not include a technique to examine whether the solution set identifies all the minimally sufficient conditions in a given data set.

The conditions generated by the traditional QCA are prime implicants instead of minimally sufficient conditions. According to Duşa (2019b), a prime implicant is "the simplest possible, non-redundant, fully consistent superset of any positive output configuration" (p. 203). "Fully consistent" means it is not a superset of an observed negative output configuration (i.e. insufficient configuration). This search goal seems meaningless for at least two reasons. First, once we accept imperfect consistency (e.g. a consistency score ≥0.8), the subset of a sufficient condition could be insufficient, as we illustrated before. Therefore, there is no need to require a sufficient condition to be "fully consistent" in its observed subsets. For example, when "AC" is minimally sufficient, there is no need to examine whether its subsets (e.g. "ABC" or "AbC") are sufficient because they contain unnecessary components. Second, the conclusion about prime implicants is context-specific: A prime implicant might not be a fully consistent superset of any positive output configuration under another data context, such as when adding more conditions included in the research. Such a risk is real because one QCA study cannot include all the variables causally relevant to the outcome (i.e. the omitted variables problem; Radaelli & Wagemann, 2019). It is reasonable to expect that a prime implicant might be a superset of an observed negative output configuration when additional conditions are added to the truth table.

Study 1: Examining the Correctness of the Top-Down QCA

We have discussed the misfit between QCA's top-down approach and its search target in a deductive way. We aim to support the above conclusions by empirically examining whether QCA's top-down approach can meet the sufficiency, necessity, and completeness requirements in Study 1. The data sets used in Study 1 are real-world data from empirical QCA studies. Rohlfing et al. (2021) conducted a reproduction analysis of 106 empirical QCA articles, and they found 28 articles are fully reproducible. Therefore, we use these 28 data sets in Study 1. The main analytical objective of QCA is to find all the minimally sufficient conditions in a given data set, where sufficiency, necessity, and completeness are three core requirements. Satisfying a former requirement is a precondition of testing a latter requirement. Therefore, we check the three requirements in turn.

First, we check the consistency of generated conditions. We do not set a uniform threshold because researchers chose different cutoffs in their studies. The threshold level we choose is the actual consistency score set by researchers in building the truth table. For example, a researcher might use 0.8 as a threshold for consistency to distinguish between sufficient and insufficient configurations in the truth table analysis. The generated conditions are considered sufficient if they pass this threshold.

Second, we check whether any component of a sufficient condition is necessary or non-redundant. Testing one condition's component's necessity makes sense only when this condition is sufficient. A component is necessary if it cannot be removed from a condition without changing the latter's status as a sufficient condition (Haesebrouck, 2022, p. 9)." Therefore, a sufficient condition's components are necessary if it cannot reach the consistency threshold anymore when we remove any component or if any supersets of a sufficient condition are insufficient. For example, "AB" is minimally sufficient if (a) it is sufficient and (b) all its supersets (i.e. "A" and "B") are not sufficient.

Third, we check a generated solution's completeness. A complete solution should contain all the minimally sufficient conditions in a given data set. Checking the completeness of a solution makes sense only when the generated solution is composed of minimally sufficient conditions. Therefore, the solution is complete and correct if we do not find other minimally sufficient conditions. We check the completeness of a solution in two ways. First, we run a sufficiency analysis using a subset of included conditions and see if we can identify any new minimally sufficient conditions. If it did not work, we then performed a more thorough search by running a necessity analysis and checking the necessity of all possible combinations of included conditions. We use the coverage score to determine whether a condition is minimally sufficient because one condition's coverage score in the necessary analysis is identical to its consistency score in the sufficiency analysis. We can conclude that a solution is correct if we do not find new minimally sufficient conditions by either method.

Based on 28 real-world data sets, we conducted 43 separate truth table sufficiency analyses using the R package *QCA* (Duşa, 2019b) because some researchers conduct sufficiency analyses for both positive and negative outcomes. Each sufficiency analysis allows us to get QCA-CS and QCA-PS, but only in studies where researchers make directional expectations can we get QCA-IS. Researchers made directional expectations in 26 analyses, enabling us to generate 26 QCA-IS. In total, we got 43 QCA-CS and 42 QCA-PS, and 26 QCA-IS. Table 1 shows the results of sufficiency analyses. A list of 28 articles, reproduction materials of all the analyses, and details for each sufficiency analysis can be found via an anonymous link on OSF: https://osf.io/fcmez/?view_only=86010a5736e241299a5014e00b429be5.

Regarding the correctness of identified conditions, none of the three solution types can guarantee to generate only minimally sufficient conditions. QCA-CS generated 127 conditions, 126 of which passed the desired consistency thresholds, achieving a 99.21% correct rate in terms of sufficiency. However, only 3.97% of generated conditions were minimally sufficient, which means a large part of QCA-CS contains unnecessary factors. QCA-PS produced 112 conditions, of which only 82.14% were sufficient, meaning a fair number of conditions did not meet the desired consistency cutoffs. Nevertheless, QCA-PS outperformed QCA-CS in eliminating redundancies because 55.36% of identified conditions were minimally sufficient. QCA-IS outperformed QCA-PS in ensuring sufficiency because 93.75% of generated conditions passed the desired thresholds; however, it underperformed QCA-PS in eliminating redundant factors since only 14.06% of conditions were minimally sufficient. Therefore, the results of Study 1 support our argument that QCA cannot achieve its main analytical objective of finding

Table 1. Results of Sufficiency Analyses From 28 Real-World Data Sets.

Criteria	Solution Type		
	Conservative	Parsimonious	Intermediate
Information about conditions			
Number of generated conditions	127	112	64
Sufficient conditions	126 (99.21%)	92 (82.14%)	60 (93.75%)
Minimally sufficient conditions	5 (3.97%)	62 (55.36%)	9 (14.06%)
Information about solutions			
Number of solutions (analyses)	43	42	26
Solutions comprised of only minimally sufficient conditions	2 (4.65%)	16 (38.10%)	1 (3.85%)
Solutions without finding other minimally sufficient conditions	1 (2.33%)	4 (9.3%)	0

Note. One analysis did not generate a parsimonious solution because all conditions were minimized, thereby reducing the number of analyses from 43 to 42.

all the minimally sufficient conditions, regardless of the solution type. The generated conditions could be insufficient for the outcome and thus do not lead to the outcome's occurrence; they may also contain redundant components that can be removed without affecting their status as sufficient conditions.

Introducing the New Bottom-Up Approach

The two most eminent configurational comparative methods are QCA (Ragin, 1987; 2008) and CNA (Baumgartner, 2009; Baumgartner & Ambühl, 2020a). Both QCA and CNA agree with the INUS theory of causation and set theory (Baumgartner & Ambühl, 2020b; Swiatczak, 2021) and seek to analyze causal complexity (Haesebrouck & Thomann, 2021). However, they have different algorithms and search strategies (Swiatczak, 2021). Unlike QCA's top-down approach, CNA adopts a bottom-up search strategy by first testing whether the single condition is sufficient for the occurrence of the outcome (Baumgartner & Ambühl, 2020a). Once a configuration fulfills the consistency score, it is automatically a redundancy-free or minimally sufficient condition (Baumgartner & Ambühl, 2020a). Unlike QCA, which only requires the lowest bound for consistency, CNA also imposes coverage a threshold to guarantee that the conjunctions of identified sufficient conditions are also necessary for the given outcome. Such a search goal (i.e. redundancy-free sufficient *and* necessary conditions) means CNA will subsequently test conditions' coverage after sufficiency requirements have been fulfilled. Swiatczak (2021) notes that, perhaps because CNA is still a rather new tool kit, QCA scholars have not explicitly discussed whether QCA can also adopt a bottom-up search strategy for sufficiency analysis in a similar way to CNA.

We show that QCA can also proceed with a bottom-up search strategy in the sufficiency analysis. First, the new bottom-up approach requires one sufficiency criterion for filtering conditions. The consistency benchmark should be ≥0.8, in line with the widely adopted QCA practice (Greckhamer et al., 2018; Ragin, 2008, p. 201). Second, researchers often consider a case frequency cutoff, which measures how many cases with a membership score in the condition are greater than 0.5. Higher case frequency means identified conditions can represent more cases and have higher coverages. The minimum frequency threshold is typically 1 in QCA, although researchers can change it in large N-studies if they wish the identified conditions could cover more cases and have more important practical significance (e.g., case frequency ≥10% of the sample). They could also alter the other standards depending on the purposes of their studies.

The bottom-up approach first conducts an analysis of sufficiency for single conditions. If (a) consistency score is ≥0.8; (b) case frequency ≥1, then the single condition should be considered as a minimally sufficient condition for the outcome and constitutes a part of the solution. Following the logic of redundancy-free, any additional components are redundant. Therefore, subsets of the identified solutions should be directly excluded from the subsequent sufficiency analysis. However, if any single condition cannot meet the two thresholds, researchers can conclude that single conditions alone cannot lead to the outcome. Therefore, we need to test the conjunctions of two conditions. If some conjunctions of two conditions are subsets of the minimally sufficient conditions identified in the previous steps, they will be excluded from the test list. Only the conjunctions of two conditions meeting the two criteria can be incorporated into the solution formula of QCA. Then, it continues checking conjunctions of three conditions until the most complex combination of conditions has been tested.

After finishing the search process, this bottom-up approach should also generate the model fit parameters such as consistency and coverage. Unlike CNA, which defines both coverage and consistency thresholds *ex-ante*, QCA only imposes a consistency cutoff while accepting a low coverage score *ex-post*. QCA accepts models that can only explain a part of the outcome's membership score as long as the sufficiency of the identified conditions passes the desired thresholds (Pappas & Woodside, 2021; Thiem, 2017). The new approach of QCA also follows this convention.

One thing easily confused with the bottom-up QCA approach is the consistency cubes (CCubes) algorithm, a new QCA algorithm developed by Duşa (2019). CCubes also proceeds with a bottom-up search strategy when generating solutions. However, the search target of CCubes is identical to the current top-down QCA: finding all the prime implicants instead of all the minimally sufficient conditions in a given data set (Duşa, 2019b, p. 203). According to Duşa (2019)'s description, three main QCA algorithms (i.e., QMC, eQMC, and CCubes) produce "exactly the same output (p. 365) ", and their main difference is the speed when driving solutions. In the user manual of the R package *QCA*, Duşa (2021) also clearly states that "all algorithms [the classical Quine-McCluskey, the enhanced Quine-McCluskey, and the latest Consistency Cubes algorithm] return the same, exact solutions (p. 23)."

The bottom-up QCA has many advantages, automatically bypassing several problems associated with the top-down approach. First, the new algorithm does not rely on a dichotomous truth table, making set membership of greatest ambiguity at 0.5 not a problem in the entire analysis process. Second, researchers do not have to make easy or difficult counterfactual assumptions, as the new algorithm only uses observed or empirical information. It does not require researchers' prior knowledge during the data analysis process, thus, avoiding the risks of making false counterfactual assumptions. Third, increasing the number of conditions does not necessarily increase the solutions' complexity. By contrast, when the number of conditions increases, solutions generated by the traditional top-down approach usually become too complex to make a meaningful interpretation. Fourth, omitted variables or conditions will not cause problems for QCA solutions because adding new conditions to the analysis does not affect the correctness of previously identified conditions. It might only result in identifying new configurations associated with the newly added conditions. Fifth, the bottom-up approach will only generate one solution type instead of three types. QCA researchers do not have to struggle with which solution to choose.

However, the bottom-up QCA also has limitations. First, the false positive problem (i.e. Type I error) may affect solutions generated by the bottom-up QCA: The configurations uncovered by QCA may result from random chance instead of reflecting meaningful set-theoretical relationships (Braumoeller, 2015; Gibson & Vann Jr, 2016; Krogslund et al., 2015). Therefore, researchers should use some tools (e.g. R packages "braQCA" and "QCAfalsepositive") to reveal the possibility that generated configurations are spurious. Second, the bottom-up approach only improves QCA's sufficiency analysis and does not address the weakness of QCA's necessity analysis in fuzzy-set data. For example, QCA can only make necessity statements qualitatively, ignoring variations in degree (Vis & Dul, 2018). QCA researchers have not achieved a consensus on how to integrate necessity analysis with sufficiency analysis results. A recommended approach to combine sufficiency and necessity results in QCA is to use necessary condition analysis (NCA) to compute levels of conditions that are necessary for the outcome >0.5 because the presence of the outcome is defined as a set membership score >0.5 in fuzzy-set data (Dul, 2020). Then, researchers can meaningfully combine sufficiency and necessity analyses by integrating NCA's necessity results into QCA's sufficiency analysis (Ding, 2022; Dul, 2020; 2021).

Study 2: Empirical Demonstration Using a Simulated Data Set

In Study 2, we aim to provide an empirical demonstration of QCA's bottom-up approach, showing data analysis steps in detail. We use a simulated data set from Swiatczak (2021)'s study. This data set has the feature of limited diversity, noise, and varying frequency. The main reason for choosing the data set is that it is close to the real-world data because the real-world data used by empirical QCA researchers are far from perfect (Baumgartner, 2021; Rohlfing, 2016; Schneider, 2018). Table 2 displays the data set with three causal conditions (i.e. A, B, and C), one outcome (i.e. Y), and the number of each primitive complex configuration (i.e. n).

Using the R package *QCA* (Duşa, 2019b), we conducted a bottom-up search strategy following the previous procedures in the sufficiency analysis. For comparison with the results of the different approaches, the same value of 0.75 for the raw consistency threshold was specified. A case frequency of 1 for the bottom-up approach was specified as the criteria. Only conditions that meet the two criteria can be considered part of the solution formula. This new approach first checks whether any single condition is sufficient for the outcome. Table 3 shows that when A, B, and C are present, they each can guarantee the outcome's presence. Therefore, the presence of the three single conditions should be considered minimally sufficient conditions.

Then, it tests whether the conjunctions of two conditions can lead to the outcome. In this step, the conjunctions that are subsets of previously identified solutions are excluded from the analysis because all of them are redundant. Therefore, in subsequent analysis, we exclude "AB," "AC," "Ab," "Ac," "aB," "aC," "Bc," "BC," and "bC." Results show that every configuration composed of two conditions cannot meet the three standards simultaneously.

Next, the sufficiency analysis for the conjunctions of three conditions shows that the combination of three conditions (i.e. "abc") is insufficient for the outcome. The most complex combinations have been examined at this step, and the sufficiency analysis ends. In total, the new approach finds that "A," "B," or "C" alone can guarantee the presence of the outcome. Therefore, the solution formula derived from the bottom-up approach should be "A + B + C → Y." Each condition is a minimally or redundancy-free sufficient condition for the result. Overall, the solution produced by the new algorithm reaches a raw consistency of 0.833 and coverage of 1.

Table 4 presents different results generated by CNA and QCA's top-down and bottom-up approaches. Such a comparison clearly shows their differences in methodological approaches and search targets. Swiatczak (2021) sets the consistency and coverage thresholds for CNA to 0.75 and the consistency cutoff for QCA to 0.75. CNA produces two models fitting equally well for the data, showing model ambiguity. Since the search target of CNA is slightly different from QCA, we mainly focus on comparing the differences in solutions derived from the top-down and bottom-up approaches of QCA.

Three requirements relevant to QCA's correctness are sufficiency, necessity, and completeness. First, all the conditions generated pass the desired consistency threshold. However, the top-down QCA fails to meet the necessity and completeness. From Table 3, it is clear that conditions "A," "B," or "C" alone can lead to the presence of the outcome because each of them can fulfill the consistency threshold independently. Only the bottom-up approach seems to guarantee the goal of QCA and generate solutions that satisfy the three requirements. In this case, both QCA-PS and QCA-CS fail to reveal this true causal structure. For QCA-CS, both conditions are redundant. The first condition (i.e. "AB") is not redundancy-free because "A" or "B" alone can lead to the outcome. The second condition (i.e. "Ac") also contains redundant elements because it is still sufficient for the outcome when condition "c" is removed from this condition. For QCA-PS, it omits the other two redundancy-free sufficient conditions, failing to achieve completeness.

The new approach of QCA can achieve sufficiency and completeness because it tests sufficiency for every single condition and every possible configuration; it can also achieve necessity as the bottom-up search strategy guarantee that identified solutions are automatically redundancy-free. The bottom-up algorithm hasn't been integrated into current QCA software, but researchers can get the same results using the function *superSubset* in R package *QCA*. When we specify the relation as "sufficiency", the function will output "all implicants which are subsets of the outcome, and similarly eliminates the redundant ones and return the surviving (minimal) subsets (Duşa, 2021, p. 41)." Therefore, researchers can use specific functions of the R package *QCA* to identify all the minimally sufficient conditions in a given data set.

Table 2. Data Set With Limited Diversity, Noise, and Varying Frequency.

A	B	C	Y	n
1	1	1	1	5
0	1	1	1	5
1	1	0	1	1
1	0	0	1	4
0	1	0	0	1
0	0	0	0	5
0	1	1	0	2

Source: Swiatczak (2021).

Study 3: Empirical Demonstration Using a Real-World Data Set

Study 3 has two main objectives. First, we aim to illustrate the applicability of the bottom-up approach in different settings. We use a real-world fuzzy-set data set, different from a simulated crisp-set data set used in Study 2. Second, we illustrate that researchers can use specific functions of R packages to derive results identical to those generated by the bottom-up approach because QCA software specifically designed to perform the bottom-up approach is currently unavailable.

The data set is from Ding (2022)'s study, in which he explores how combinations of six conditions lead to high levels of national innovation performance. For simplicity and clarity, we only included four conditions. We renamed the condition variables as "A-D" and the result variable as "Y." Using the R package *QCA*, we first ran the top-down approach (i.e, a truth table and Boolean minimization) to derive QCA-

Table 3. Procedures of Sufficiency Analysis by the Bottom-Up Approach.

Steps	Conditions	Consistency	Case Frequency
Step 1: Conducting sufficiency analysis for single conditions	A	1.000	10
	B	0.786	14
	C	0.833	12
	a	0.385	13
	b	0.444	9
	c	0.455	11
Step 2: Conducting sufficiency analysis for the conjunctions of two conditions	ab	0.000	5
	ac	0.000	6
	bc	0.444	9
Step 3: Conducting sufficiency analysis for the conjunctions of three conditions	abc	0.000	5

Note. Case frequency means the number of cases with a membership score over 0.5 in this condition.

Table 4. Solution Types of CNA, Top-Down Approach, and Bottom-Up Approach of QCA.

Solution Type	Solution	Consistency	Coverage	Threshold
CNA solution	(1) A + C ↔ Y (2) A + B ↔ Y	0.882	1	0.75
QCA-CS	AB + Ac → Y	1	0.667	0.75
QCA-PS	A → Y	1	0.667	0.75
Bottom-up QCA	A + B + C → Y	0.833	1	0.75

Note. Solutions of CNA and QCA are from Swiatczak (2021). CNA imposes a threshold of 0.75 for coverage and consistency. The two QCA approaches only impose a threshold of 0.75 for consistency. "AB" denotes "A and B"; "A + B" denotes "A or B"; "A ↔ B" denotes "A is sufficient and necessary for B"; "A → B" denotes "A is sufficient for B."

Table 5. Results of QCA's Top-Down And Bottom-Up Approaches.

Solution Type	Solution	Consistency	Coverage	Threshold
QCA-PS	AB + CD → Y	0.859	0.834	0.8
QCA-CS	AB + ACD + BCD → Y	0.862	0.832	0.8
Bottom-up QCA	D + AB + AC + BC → Y	0.748	0.942	0.8

Note. We use 0.8 as the sufficiency threshold in Study 3. The sufficiency (or consistency) scores for "D," "AB,""AC," and "BC" are 0.812, 0.897, 0.828, and 0.908, respectively.

PS and QCA-CS. Then we use the *superSubset* function in the R package *QCA* to derive all the minimally sufficient conditions. Table 5 presents the results. Reproduction materials of Study 3 can be accessed via the same link provided in Study 1.

Although the bottom-up QCA's solution set as a whole is not sufficient, each condition included in the solution is minimally sufficient for the outcome's occurrence. The bottom-up QCA generates four minimally sufficient conditions, highlighting the flaws of solutions generated by the QCA's top-down approach. QCA-PS and QCA-CS are not correct because they produce conditions that contain unnecessary components or omit some minimally sufficient conditions for the outcome. The findings indicate that the bottom-up approach is a better choice for finding redundancy-free sufficient conditions than the top-down approach.

Conclusion

Many scholars have uncritically employed QCA to analyze causality complexity despite growing debates and controversies around QCA's methodology (Haesebrouck & Thomann, 2021). The current QCA follows a top-down search strategy, which could induce many causal fallacies. We found evidence supporting this argument in Study 1. Although the similarities and differences between QCA and CNA have attracted researchers' attention, few have realized that CNA's bottom-up search strategy can also be used in QCA's sufficiency analysis (Swiatczak, 2021). We have discussed such a methodological possibility in this article. Using two empirical demonstrations, we show that the bottom-up approach is guaranteed to identify all the minimally sufficient conditions, thus making it a better choice for QCA users. Although a program specifically designed for bottom-up QCA is currently unavailable, researchers can use specific functions of R package QCA to get the same results.

Acknowledgment

BI Norwegian Business School funded the publication fee.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: BI Norwegian Business School funded the publication fee.

ORCID iD

Haien Ding https://orcid.org/0000-0002-1519-7200

References

Baumgartner, M. (2008). Regularity theories reassessed. *Philosophia*, *36*(3), 327–354. https://doi.org/10.1007/s11406-007-9114-4

Baumgartner, M. (2009). Inferring causal complexity. *Sociological Methods & Research*, *38*(1), 71–101. https://doi.org/10.1177/0049124109339369

Baumgartner, M. (2015). Parsimony and causality. *Quality & Quantity*, *49*(2), 839–856. https://doi.org/10.1007/s11135-014-0026-7

Baumgartner, M. (2022). Qualitative comparative analysis and robust sufficiency. *Quality & Quantity*, *56*(4), 1939–1963. https://doi.org/10.1007/s11135-021-01157-z

Baumgartner, M., & Ambühl, M. (2020a). Causal modeling with multi-value and fuzzy-set Coincidence Analysis. *Political Science Research and Methods*, *8*(3), 526–542. https://doi.org/10.1017/psrm.2018.45

Baumgartner, M., & Ambühl, M. (2020b). *CNA: An R package for configurational causal inference and modeling*. https://cran.r-project.org/web/packages/cna/vignettes/cna.pdf

Baumgartner, M., & Thiem, A. (2017). Model ambiguities in configurational comparative research. *Sociological Methods & Research*, *46*(4), 954–987. https://doi.org/10.1177/0049124115610351

Braumoeller, B. F. (2015). Guarding against false positives in qualitative comparative analysis. *Political Analysis*, *23*(4), 471–487. https://doi.org/10.1093/pan/mpv017

Castaño, M. S., Méndez, M. T., & Galindo, M. Á. (2016). The effect of public policies on entrepreneurial activity and economic growth. *Journal of Business Research*, *69*(11), 5280–5285. https://doi.org/10.1016/j.jbusres.2016.04.125

Cooper, B., & Glaesser, J. (2016). Qualitative comparative analysis, necessary conditions, and limited diversity: Some problematic consequences of schneider and wagemann's enhanced standard analysis. *Field Methods*, *28*(3), 300–315. https://doi.org/10.1177/1525822X15598974

Ding, H. (2022). What kinds of countries have better innovation performance?-A country-level fsQCA and NCA study. *Journal of Innovation & Knowledge*, *7*(4), 100215. https://doi.org/10.1016/j.jik.2022.100215

Dul, J. (2020). *Comparing NCA and QCA. Supplement to 'Dul, J. (2020) conducting necessary condition analysis*. Sage Publications. https://www.erim.eur.nl/necessary-condition-analysis/publications/book/complementing-nca-and-qca/

Dul, J. (2021). *Advances in necessary condition analysis*. https://bookdown.org/ncabook/advanced_nca2/

Dusa, A. (2019). Consistency cubes: A fast, efficient method for exact boolean minimization. *The R Journal*, *10*(2), 357. https://doi.org/10.32614/RJ-2018-080

Duşa, A. (2019a). Critical tension: Sufficiency and parsimony in QCA. *Sociological Methods & Research*, *51*(2), 1–25. https://doi.org/10.1177/0049124119882456

Duşa, A. (2019b). *QCA with R: A comprehensive resource*. Springer International Publishing.

Duşa, A. (2021). *Package 'QCA.'*. CRAN. https://cran.r-project.org/web/packages/QCA/index.html

Gibson, C. B., & Vann, B. Jr. (2016). *The bootstrapped robustness assessment for qualitative comparative analysis arXiv: 1606.05000)*. arXiv https://doi.org/10.48550/arXiv.1606.05000

Graßhoff, G., & May, M. (2001). Causal regularities. In W. Spohn, M. Ledwig, & M. Esfeld (Eds.), *Current issues in causation* (pp. 85–114). Mentis Verlag.

Greckhamer, T., Furnari, S., Fiss, P. C., & Aguilera, R. V. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization*, *16*(4), 482–495. https://doi.org/10.1177/1476127018786487

Haesebrouck, T. (2022). Relevant, irrelevant, or ambiguous? Toward a new interpretation of QCA's solution types. *Sociological Methods & Research*, *in press*, 004912412110361. https://doi.org/10.1177/00491241211036153.

Haesebrouck, T., & Thomann, E. (2021). Introduction: Causation, inferences, and solution types in configurational comparative methods. *Quality & Quantity*, *56*(4), 1867–1888. https://doi.org/10.1007/s11135-021-01209-4

Hanckel, B., Petticrew, M., Thomas, J., & Green, J. (2021). The use of qualitative comparative analysis (QCA) to address causality in complex systems: A systematic review of research on public health interventions. *BMC Public Health*, *21*(1), 877. https://doi.org/10.1186/s12889-021-10926-2

Krogslund, C., Choi, D. D., & Poertner, M. (2015). Fuzzy sets on shaky ground: Parameter sensitivity and confirmation bias in fsQCA. *Political Analysis*, *23*(1), 21–41. https://doi.org/10.1093/pan/mpu016

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly*, *2*(4), 245–264. https://www.jstor.org/stable/20009173

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford University Press. https://doi.org/10.1093/0198246420.001.0001

Mahoney, J., & Acosta, L. (2021). A regularity theory of causality for the social sciences. *Quality & Quantity*, *56*(4), 1889–1911. https://doi.org/10.1007/s11135-021-01190-y

Misangyi, V. F., Greckhamer, T., Furnari, S., Fiss, P. C., Crilly, D., & Aguilera, R. (2017). Embracing causal complexity: The emergence of a neo-configurational perspective. *Journal of Management*, *43*(1), 255–282. https://doi.org/10.1177/0149206316679252

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in Information Systems and marketing. *International Journal of Information Management*, *58*, 102310. https://doi.org/10.1016/j.ijinfomgt.2021.102310

Radaelli, C. M., & Wagemann, C. (2019). What did I leave out? Omitted variables in regression and qualitative comparative analysis. *European Political Science*, *18*(2), 275–290. https://doi.org/10.1057/s41304-017-0142-7

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. University of California Press.

Ragin, C. C. (2000). *Fuzzy-set social science*. University of Chicago Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Ragin, C. C., Mayer, S. E., & Drass, K. A. (1984). Assessing discrimination: A boolean approach. *American Sociological Review*, *49*(2), 221–234. https://doi.org/10.2307/2095572

Rohlfing, I. (2016). Why simulations are appropriate for evaluating qualitative comparative analysis. *Quality & Quantity*, *50*(5), 2073–2084. https://doi.org/10.1007/s11135-015-0251-8

Rohlfing, I., Königshofen, L., Krenzer, S., Schwalbach, J., & Bekmuratovna R, A. (2021). A reproduction analysis of 106 articles using qualitative comparative analysis, 2016–2018. *PS: Political Science & Politics*, *54*(2), 292–296. https://doi.org/10.1017/S1049096520001717

Schneider, C. Q. (2018). Realists and idealists in QCA. *Political Analysis*, *26*(2), 246–254. https://doi.org/10.1017/pan.2017.45

Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge University Press.

Swiatczak, M. D. (2021). Different algorithms, different models. *Quality & Quantity*, *56*(4), 1913–1937. https://doi.org/10.1007/s11135-021-01193-9

Thiem, A. (2017). Standards of good practice and the methodology of necessary conditions in qualitative comparative analysis. *Political Analysis*, *24*(4), 478–484. https://doi.org/10.1093/pan/mpw024

Thiem, A. (2022). Beyond the facts: Limited empirical diversity and causal inference in qualitative comparative analysis. *Sociological Methods & Research*, *51*(2), 527–540. https://doi.org/10.1177/0049124119882463

Thiem, A., & Duşa, A. (2013). Boolean minimization in social science research: A review of current software for qualitative comparative analysis (QCA). *Social Science Computer Review*, *31*(4), 505–521. https://doi.org/10.1177/0894439313478999

Vis, B., & Dul, J. (2018). Analyzing relationships of necessity not just in kind but also in degree: Complementing fsQCA with NCA. *Sociological Methods & Research*, *47*(4), 872–899. https://doi.org/10.1177/0049124115626179

Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research*, *66*(4), 463–472. https://doi.org/10.1016/j.jbusres.2012.12.021

Woodside, A. G. (2014). Embrace•perform•model: Complexity theory, contrarian case analysis, and multiple realities. *Journal of Business Research*, *67*(12), 2495–2503. https://doi.org/10.1016/j.jbusres.2014.07.006