Chapter 7
Testing Joint Sufficiency Twice: Explanatory Qualitative Comparative Analysis

Alessia Damonte

Abstract Standard Qualitative Comparative Analysis (QCA) applies an eliminative cross-case algorithm to identify which combinations of factors are logically associated with an outcome in a population. As such, it suits the purpose of pinpointing the conditions under which an outcome occurs or fails. However, the explanatory import of its findings only follows if the algorithm identifies theoretically *interpretable*, logically *valid*, and empirically *plausible* causal compounds.

The chapter provides an essential guide to designing an explanatory QCA that meets the three credibility requirements at once. Section 7.2 addresses how to develop starting hypotheses consistent with the assumptions of complex causation to preserve theoretical interpretability. Section 7.3 introduces the Boolean algebra required to model a hypothesis and find which part supports the explanatory claim in the cases at hand. Section 7.4 addresses the issue of gauging conditions to ensure the empirical plausibility of the analysis. Last, Sect. 7.5 summarizes the protocol, illustrated by the replicable example in the online R file.

Learning Objectives
After studying this chapter, you will be able to:

- Understand causation in terms of individual necessity and joint sufficiency of many factors.
- Develop a configurational hypothesis.
- Apply Boolean algebra to formalize configurational hypotheses and establish criteria of fit.
- Gauge factors as sets that are suitable to logical formalizations.
- Identify and discuss credible configurational solutions.

A. Damonte (✉)
University of Milan, Milan, Italy
e-mail: alessia.damonte@unimi.it

A. Damonte, F. Negri (eds.), *Causality in Policy Studies*, Texts in Quantitative Political Analysis, https://doi.org/10.1007/978-3-031-12982-7_7

7.1 Introduction

Qualitative Comparative Analysis (QCA: Ragin, 1987/2014, 2000, 2008; Duşa, 2019; Oana et al., 2021; Mello, 2021) stands amid the suite of causal techniques for three main reasons that drive as many questions.

First, QCA moves from the default assumption that causation lies in compounds or teams of conditions. Its solutions entail that things happen when all the "right" conditions are given together, like in a chemical reaction (Mackie, 1965, 1974; Cartwright & Hardie, 2012). The first question of explanatory QCA asks how to ensure that results are interpretable "recipes" for the outcome.

Second, QCA originally revolves around a pruning algorithm. It compares configurations that meet regularity requirements of association with an outcome to drop irrelevant conditions, along the lines of a most-dissimilar case design (e.g., De Meur & Berg-Schlosser, 1994), albeit run twice. The second question asks how the technique can be geared toward pinpointing valid causal compounds despite the shortcomings of such a design (e.g., Geddes, 1990; Most & Starr, 2015; Krogslund et al., 2015).

Third, QCA's solutions hold at the levels of both the population and individual cases. Such a peculiarity is based on gauging operations that preserve quantitative and qualitative information. These operations are an integral part of the analysis and bind findings to analytic units. The third question asks how these operations affect the tenability of solutions.

These three questions are addressed in Sects. 7.2, 7.3, and 7.4, respectively. Section 7.5 summarizes the protocol illustrated by the online R file.

7.2 Interpretability

The recognized hallmark of QCA lies in its assumptions that causation is an asymmetric, conjunctural, and equifinal phenomenon (Ragin, 2008; see also Rosenberg et al., 2017). *Asymmetric* means that causation has a direction and proceeds from "causes" to "effects" as a relationship of dependence or conditionality ahead of temporal considerations. *Conjunctural* refers to the first reason for asymmetry: the actual cause is a compound and consists of a team, bundle, or package of contributing factors. *Equifinal* recalls the second reason for asymmetry: different compounds can yield the same outcome. These assumptions chime with mechanistic considerations on the ultimate shape of causation (e.g., Befani, 2013; Mahoney, 2021; Chap. 2).

7.2.1 Mechanisms and Machines

QCA assumes that the factors responsible for an outcome are many and related to each other as the constituting parts are to their whole. Moreover, it allows factors have substitutes without loss of effectiveness for the causal compound (Mackie, 1966; Cheng, 1997; Cartwright & Hardie, 2012).

The textbook illustration of such a parts-to-whole relationship offers heat, oxygen, fuel, and defective or no sprinklers as the compound accounting for fire. These circumstances provide the complete set of relevant conditions under which the process of combustion must initiate (Salmon, 2020). Thus, they form a causal team based on the process that they explain.

The process also clarifies the general relationship between components, teams, and outcomes. In the textbook example, combustion results in a fire when the whole team of circumstances is given in the same place and the right state—present heat, fuel, oxygen; absent or defective sprinklers. The surefire or *sufficient cause* of the outcome is the right bundle. However, the right circumstances can take many actual shapes. For instance, a lightning bolt, a short circuit, or a lit match can all be equivalent sources of heat. Any actual bundle, then, is *unnecessary* as such. Besides, the process fails when any circumstance is given in the wrong state—poor oxygen, no fuel, or no heat all prevent combustion, while a working fire system suffocates it. Any element of the compounds, then, is a counterfactually vital—and hence, *necessary*—component of the team, despite it alone being insufficient to yield the outcome. The elements of the compound are "partial causes" or "*inus* conditions"—*inus* being the acronym of the *Insufficient* but *Necessary* part of an *Unnecessary* but *Sufficient* team.

Bundles of *inus* conditions seldom capture a generative process directly (see Chaps. 8, 9, and 10). Instead, they can capture the set of right circumstances as "nomological machines"—that is, as "sufficiently stable" arrangements of triggering, enabling, sustaining, and shielding conditions underlying the generative process (Cartwright, 1999: 49, 2017). A nomological machine is such that its components together make other factors irrelevant before the same type of outcome across time and space. Therefore, a nomological machine is the specified explanation of a regular behavior independent of the remaining context (Craver & Kaplan, 2020). Moreover, it provides the theoretical construct that affords counterfactual evidence about the contribution of single components across cases.

7.2.2 Operationalizing Typological Theories

Typological theories provide a renowned starting point for developing configurational explanations (e.g., Elman, 2005). Such theories prove especially fruitful as they enable modeling of the alternative causal bundles as different settings of the same factors.

Some theories are consistent "explications" of a driving concept. For instance, Pahl-Wostl (2008) takes "regimes" as the driving concept. She defines water management regimes as the alignment of governance style, type of sectoral integration, scale of analysis and operation, information management, plus finance and risk management. Huntjens et al. (2011) operationalize the setting of these structural dimensions for two polar types of regimes—the "market-based" and the "integrated adaptive"—then run a QCA to establish the features that account for the diversity in the policy-learning capacity of water management systems when faced with climate change challenges. In a similar vein, Colby (1991) builds on the concept of "policy paradigms." He stipulates that the compatibility of environmental and economic policy goals depends on the alignment of policy ideas and policy tools. Thus, "frontier economics" and "deep ecology" establish the trade-off between economic growth and environmental preservation, while "environmental protection," "resource management," and "eco-development" make room for their coexistence and integration. Damonte (2013) operationalizes these alternative paradigms as different settings of the same bundle of policy tools and identifies the configurations that account for the green decoupling of economic growth from pollution.

Other configurational hypotheses integrate heterogeneous streams of literature into a consistent explanatory whole. For instance, Sabatier and Mazmanian (1980) reason that the many accounts of the success and failure of policy implementation can be reduced to the consistent interplay of three dimensions: problem tractability, administrative effectiveness, and political support. Hinterleintner et al. (2016) operationalize the components of each dimension and run a QCA that explains the differences in the IMF's evaluation of austerity programs as differences in the credibility of national implementations. Theoretical integration can also be purposefully operated within the study. As an example, Lauri et al. (2020) integrate theories linking the defamiliarization of care work and gender equality with theories on the gender division of labor as embedded in different types of welfare systems. On this basis, they provide a thorough operationalization of childcare policies as bundles of tools that enforce different gender norms. QCA is applied to identify which tools, linked to the norms of which type of welfare system, yield high gender equality and which endanger the goal instead.

7.2.3 Assembling Configurational Hypotheses

A configurational hypothesis can also be crafted after a reasoned selection and integration of statistical "determinants." Surveys of scholars' practices (Amenta & Poulsen, 1994; Berg Schlosser & De Meur, 2009) pinpointed four selection strategies. The "comprehensive approach" includes all the factors from all the relevant theories; the "perspective approach" selects single variables that represent major theories; the "significance approach" only focuses on statistically significant variables; the "second look" approach mixes statistically significant variables with theoretically meaningful factors that did not survive those same tests.

However, none of these strategies is proven to yield proper configurational hypotheses unless the selected factors can be related to the unfolding of a generative process as actors' constraints and opportunities. To witness, Stiller (2017) explains governments' success in adopting major welfare reforms as the interplay of policymakers' strategies—identified in ideational leadership, concession making, and blame avoidance—with key background features that make these strategies adequate—namely, the stage of the election cycle and the government's position toward the national welfare system. Similarly, Ansell et al. (2020) account for stakeholders' participation in collaborative governance as the result of motivations—that is, perceived incentives, interdependence, trust, and purpose—and governance's support of motivations—through leadership services, opportunities to build relationships, and structures for pooling information.

A configurational hypothesis may also follow from problematizing correlational theories. Kogut and Ragin (2006) focus on the theory linking high economic development, thriving financial markets, and common law institutions. The configurational hypothesis develops from the consideration that the causal chain is underspecified. National economies, they reason, may still thrive despite poor financial markets if legality is ensured. Moreover, the effectiveness of common law institutions beyond their original contexts depends on their interplay with existing legal traditions. Thus, they run two QCAs that employ common law, features of the institutional "transplant," and commitment to the rule of law to account for differences in GDP per capita and, separately, in the dimension of the domestic financial markets, to check whether the two explanations overlap.

In short, the fundamental criterion for selecting an interpretable candidate *inus* factor is functional. It consists of whether one can develop *directional expectations* about the factor's contribution to the setting that compels and protects some causal process of interest. The expectation should support the claim that, were the factor given in the right state and in the right team, the process to the outcome would certainly follow. As we will see in Sect. 7.3.2, these directional expectations play a crucial role in the analysis as they establish the plausibility of counterfactual assumptions.

7.3 Validity

The validity of inferences about *inus* hypotheses depends on the algebra deployed to make them testable. Such a suitable algebra should allow factors to

- Have observable states, such as presence and absence;
- Form compounds as configurations of states;
- Have equifinal alternatives;
- Establish relationships of dependence.

Boolean algebras can easily render these states and relationships. Introduced as primary devices to analyze human reasoning about the world (De Morgan, 1847; Boole, 1853), their structures support a twofold reading (Stone, 1936)—logical, and set-theoretical.

7.3.1 QCA's Algebra

Like any other, QCA's algebra is a language of literals and operators suitable to render complex relationships according to fundamental rules.

7.3.1.1 Literals

Boolean algebras use “literal symbols” to indicate factors as attributes or states of a unit of observation. A literal stands for a name or an adjective denoting “either a thing or some quality or circumstance belonging to it” (Boole, 1853:27). QCA borrows the convention and indicates a state with an uppercase letter. Thus, *A* reads ⌜*A* present⌝ or ⌜*A* positive⌝ or the predicate ⌜is *A*⌝. The literal provides an empty placeholder for whatever attribute we consider as the candidate *inus* condition—such as “inflammable” referred to a material; “hierarchical” to a governance structure; “affluent” to a society; “independent” to a voter.

Once defined, a literal establishes the similarity of any units of observation u_i to which it applies. In Boole’s original proposal, and all the basic operations of QCA, such a recognition raises a class, that is, an *idempotent* collection of units. Idempotency means that, in contrast to probabilistic samples, classes satisfy the logical rule dubbed *dictum de omni*: that which can be said of the whole, it also holds for each of its parts. Boole renders idempotency as in Eq. (7.1):

$$A^2 := A \tag{7.1}$$

where := indicates a stipulation and reads ⌜is by definition equal to⌝. As the only two numerical values that satisfy the stipulation are 1 and 0, Boole’s literals can only take these two values—and the basic operations in QCA share this bivalent assumption, too.

These values convey two separate readings of the relationship between a unit and a literal:

- When the literal is understood as a *predicate*, 1 and 0 are the *truth values* that a literal can take in the actual unit u_i from the *universe of discourse* $\mathbb{U} = \{u_1, \ldots, u_N\}$. 1 reads ⌜true⌝ for ⌜it is the case that⌝, while 0 reads ⌜false⌝ for ⌜it is not the case that⌝.
- When the literal is understood as a *class*, 1 and 0 are read as *membership values*. Thus, $A_i = 1$ means that the *i*-th unit belongs to class *A*, while $A_i = 0$ indicates that the same unit does not belong to it.

The logical understanding captures the literal as the *intension* or quality of a unit. In contrast, the set-theoretical understanding captures the literal as the *extension* of the quality across the units in a universe. Operationally, the intension is decided by gauging rules—for instance, on defining which manifestations and intensity make it true that a unit ⌜is *A*⌝. Extension, on the other hand, is decided by counting—for instance, the number of units in the universe that ⌜are *A*⌝, which corresponds to the *cardinality* of class *A*. In bivalent Boolean algebra, the two readings overlap, making logical inferences especially straightforward.

7.3.1.2 Operators

The Boolean operators relevant to *inus* hypotheses correspond to the logical connectives ⌜*not*⌝, ⌜*and*⌝, ⌜*or*⌝, ⌜*only if*⌝, ⌜*if*⌝ and the set-theoretical relationships of *difference*, *intersection*, *union*, and *superset/subset*.

Negation

The connective ⌜*not*⌝ denies the literal. The Boolean notation renders it with a bar above the uppercase literal to which it applies; in QCA, also common is the tilde before the uppercase literal, or the use of the lowercase literal. Thus, $\overline{A}, \sim A, a$ all read ⌜is not-*A*⌝.

The logical negation transforms a unit's truth value into its opposite, calculated as in Eq. (7.2). The set-theoretical reading establishes the negation of a set is the collection of units that are excluded from that set. Therefore, the negated set $\overline{A}$ corresponds to the difference (indicated by the backslash \) between the universe U and set *A*, as in Eq. (7.3):

$$\overline{A}_i := 1 - A_i \tag{7.2}$$

$$\overline{A} := \mathbb{U} \setminus A \tag{7.3}$$

Equations (7.2) and (7.3) indicate that, by definition, a literal and its negation are mutual *complements*. The enforcement of this definition depends on gauging operations—an issue addressed in Sect. 7.4.

Joint Occurrence

These correspond to bundles of literals connected by the ⌜*and*⌝ operator. In logic, the operator is a wedge (∧); in set theory, it is a cap (∩). In QCA, the operator is a dot (●) or a star (∗) although the connecting symbol may be omitted.

Two implications are worth noting. Permutation and grouping are irrelevant to ⌜*and*⌝ bundles: $AB\overline{C}$ means the same as $A\overline{C}B$ and $A\left(B\overline{C}\right)$ as the resulting class clusters the same units. In short, the Boolean ⌜*and*⌝ supports the commutative and the associative rule. Therefore, bundles are blind to the time dimension of sequences; instead, they emphasize the joint occurrence or interaction of attributes in a unit.

Logically, the ⌜*and*⌝ operator raises a *conjunction*. The underlying rule establishes a conjunction as true when each of its conjuncts is true. The rule is also known as "*the weakest link*": the conjunct with the lowest truth value defines the truth value of the compound.

Applied to a single predicate and its negation, the rule renders the logical *principle of non-contradiction*. As summarized by Eq. (7.4), the principle states that a predicate and its negation cannot be true of the same unit at the same time in the same sense. Set-theoretically, the principle is met when the intersection of a set and its negation is empty (∅), as in Eq. (7.5). The principle offers the first criterion of validity: it commits to rejecting inferences that build on, or lead to, *contradictions*.

$$A \wedge \overline{A}_i := 0 \tag{7.4}$$

$$A \cap \overline{A} := \varnothing \tag{7.5}$$

More generally, the weakest link of the *i*-th unit can be calculated as the minimum of its truth values in any of the $1 \leq j \leq K$ conjuncts, as in Eq. (7.6):

$$\bigwedge A_{ij} = \min\left(A_{i1}, \ldots, A_{iK}\right) \tag{7.6}$$

Therefore, in a universe of N units, the cardinality of the intersection of the k literals of interest corresponds to the sum of the $1 \leq i \leq N$ units' weakest links as in (7.7):

$$\bigcap A_j = \sum_{i=1}^{N} \min\left(A_{i1}, \ldots, A_{iK}\right) \tag{7.7}$$

Alternatives

These arise when literals are connected by the operator ⌜*or*⌝. In QCA, the operator is a plus symbol (+) and never omitted. Logic indicates it with a vee (∨); set theory with a cup (∪). Class idempotency makes permutation and grouping irrelevant to alternatives, too.

Logically, the ⌜*or*⌝ operator raises a *disjunction*. The underlying rule establishes the disjunction as true when at least one of its disjuncts is true. The rule can be dubbed "*the strongest link*": the disjunct with the highest truth value defines the truth value of the whole compound.

Applied to a single predicate and its negation, the rule renders the logical *principle of the excluded middle*. As summarized by Eq. (7.8), the principle states that, necessarily, either a predicate or its negation is true in a unit, so that the disjunction of the two raises a non-informative tautology. Set-theoretically, the principle is met when the union of the set and its negation returns the universe, as in Eq. (7.9).

$$A_i \vee \overline{A}_i := 1 \tag{7.8}$$

$$A \cup \overline{A} := \mathbb{U} \tag{7.9}$$

More generally, the strongest link of the i-th unit can be calculated as the maximum of the truth values of any of the $1 \leq j \leq K$ disjuncts, as in (7.10):

$$\bigvee A_{ij} = \max\left(A_{i1}, \dots, A_{iK}\right) \tag{7.10}$$

Therefore, in a universe of N units, the cardinality of the union of the K literals of interest corresponds to the sum of the $1 \leq i \leq N$ units' strongest links, as in (7.11):

$$\bigcup A_j = \sum_{i=1}^{N} \max\left(A_{i1}, \dots, A_{iK}\right) \tag{7.11}$$

Necessity and Sufficiency

The reliance of QCA on the assumptions of *inus* causation gives center stage to the concepts of necessity and sufficiency.

Mackie (1974) illustrates them with the different behavior of coin-operated vending machines. A "sufficiency machine" always drops a snack for a coin, and sometimes it drops one without apparent reason, too. A "necessity machine" never drops a snack without a coin, and sometimes the coin fails. Last, one and only one snack for each coin is the behavior of the perfect "necessity-and-sufficiency machine." These intuitions capture both set-theoretical and logical relationships between an observed input, or antecedent (the coin), and an observed output, or consequent (the snack), connected by an unobserved—but possibly observable—mechanism.

As for notation, QCA indicates necessity with an arrow running from the outcome to the cause and sufficiency with an arrow running from the cause to the outcome. Thus, $A \rightarrow B$ reads ⌜A is sufficient to B⌝; $\overline{A} \leftarrow \overline{B}$ reads ⌜not-A is necessary to not-B⌝.

Set-theoretically, the *necessity* of A to B corresponds to A being a *superset* of B, indicated as $B \subset A$. The relationship is satisfied when *all the B are also A* although there can be instances of A in the universe that do not display B. This corresponds to the logical situation in which being B *implies* being A or, more compactly, ⌜B, only if A⌝. The hallmark of necessity is the impossibility of the outcome in the absence of the factor, as in (7.12). Set-theoretically, it means that the proof of the necessity of A to B in the universe comes from the empty intersection in (7.13).

$$\overline{A}_i \wedge B_i = 0 \tag{7.12}$$

$$\overline{A} \cap B = \varnothing \tag{7.13}$$

Set-theoretically, the *sufficiency* of A to B corresponds to A being a *subset* of B, indicated as $A \subset B$. The relationship is satisfied when *all the A are also B*. In short, sufficiency renders the intuition of A as the constant antecedent condition of

B. Logically speaking, it corresponds to saying that, for any u_i, ⌜B, if A⌝ without exceptions. The hallmark of sufficiency coincides with the impossibility that the outcome fails when the factor is present, summarized by requirement (7.14) and its set-theoretical translation (7.15):

$$\overline{B}_i \wedge A_i = 0 \tag{7.14}$$

$$\overline{B} \cap A = \varnothing \tag{7.15}$$

7.3.1.3 Truth Tables

Stipulations and rules construe valid logical inferences as the calculus of truth values, visualized with the aid of a *truth table*. These tables clarify the possibilities that the selected literals make available ahead of observation. Logic sees it as the exhaustive catalog of the combinations of the literals' truth-values (Wittgenstein, 1922). Probabilistic theories dub such a structure "*sample space*" and understand it as the list of the potential events from random trials (e.g., Clarke, 2020). In any case, this structure reports the maximum diversity that units can display given specific literals and gauges.

The truth table entails a fundamental sense-making operation (Quine, 1982); thus, in it, each combination of the literals' truth values can be dubbed a *primitive*. The number of primitives depends on the number of literals and truth values under consideration; K bivalent literals yield 2^K unique primitives. In the remaining, a truth table will be indicated as Ω and its primitives as ω.

The shape of truth tables follows conventional rules. The primitives are listed as rows: ω_1 displays all true literals; ω_{2^K}, all false ones (cfr. Duşa, 2019). Each of the remaining columns in the classical truth table is for the *truth function* of a connective, i.e., the truth values that each primitive returns when the connective's rule is applied to the states of its literals.

Table 7.1 displays a truth table of two literals (A, B) and five operators to indicate as many relationships—respectively, of conjunction (*and*), disjunction (*or*), necessity (*only if*), sufficiency (*if*), plus necessity and sufficiency (*iff*).

The values in the truth functions of each operator indicate the type of units that will (1) and will not (0) be observed if the relationship holds in the universe of reference (Sprenger, 2011). These expectations inform the discourse on the threats to the validity of inferences that are currently addressed by either design (e.g., Chap. 3) or model (e.g., Chaps. 6 and 8, Sect. 7.3.2 below).

- The *and* truth function follows from the application of the weakest link rule as in Eqs. (7.6) and (7.7) and returns a single true point in correspondence with the matching primitive (ω_1 in Table 7.1). Thus, evidence of a conjunction is only provided by the units displaying every conjunct in the right state.

Table 7.1 Truth table of two literals and five operators

Ω	*A*	*B*	*A and B*	*A or B*	*B, only if A*	*B, if A*	*B, iff A*
ω_1	*1*	*1*	1	1	1	1	1
ω_2	*1*	*0*	0	1	1	0	0
ω_3	*0*	*1*	0	1	0	1(*)	0
ω_4	*0*	*0*	0	0	1	1	1

Note: (*) observing this primitive makes the statement of sufficiency vacuously true

- The *or* truth function follows from the strongest link rule as in Eqs. (7.10) and (7.11) and always returns a single false point, corresponding to the primitive with no matching values (ω_4 in Table 7.1). It conveys that any unit displaying at least one disjunct in the right state provides evidence of a disjunction.
- The *only if* truth function has a single false point corresponding to the impossible primitive established by Eqs. (7.12) and (7.13). It shows that the relationship of necessity is only inconsistent with evidence of the consequent *B* occurring in some units where the antecedent *A* is missing (ω_3 in Table 7.1). Therefore, the logical relationship of necessity assumes the antecedent *A* is not substitutable, as is oxygen to fire.
- The *if* truth function has a single false point in the impossible primitive defined by Eqs. (7.14) and (7.15). It shows that the claim of sufficiency is only inconsistent with evidence that the consequent fails under the antecedent in some units (ω_2 in Table 7.1). The logical relationship of sufficiency is the regular connection of antecedent and consequent. When the actual cause is composite, the requirement can only be satisfied by the antecedent that comprises all the components of a compound—including the factors that shield the causal process from obstructions. Section 7.4.2 will suggest a strategy for construing suitable shielding factors.

A further note is due about the starred value of ω_3 in Table 7.1. The instances of this primitive do not contradict the claim of sufficiency after the principle that *ex falso quodlibet*—meaning that anything can follow in the units where the antecedent is missing or otherwise false. However, units of this type provide *vacuous* evidence about the relationship (e.g., Salmon, 2020), as they may

(a) point to its nonsensical nature. The evidence that Socrates is not a triangle yet is a philosopher makes the claim vacuous that "if Socrates is a triangle, then he is a philosopher."
(b) divert attention from the conditionality of interest. Evidence about salt that is not put in water is irrelevant to establish the claim that "if salt is put in water, then it dissolves."
(c) unveil some spurious relationship or incomplete explanation. The evidence that the barometer reads "storm" during a sunny day makes the claim vacuous that "if the barometer reads 'fair,' then it is a sunny day."

Although the exact meaning of a vacuous observation depends on the interpretability of the relationship of interest, it nevertheless makes the problem visible as a formal issue of validity.

- The *iff* relationship arises from the conjunction of the truth functions of necessity and of sufficiency. It indicates the identity of the two literals and the overlapping of the respective classes of units in the universe. Thus, the truth function has two false points. In Table 7.1, these correspond to ω_2 and ω_3. In short, evidence of any inconsistency in the covariation of the two states challenges the validity of the identity.

QCA does not deploy logic, truth tables, and truth functions normatively. Instead, it relies on them as modeling tools and heuristics for the analysis.

7.3.2 Identifying Valid Inus *Hypotheses*

Logic provides scaffolding and criteria to render an *inus* hypothesis first, then decide whether it is rightly specified to the universe under analysis.

7.3.2.1 Rendering Hypotheses

Logic renders an *inus* hypothesis as a theoretically meaningful yet unwarranted claim about the sufficiency of a conjunction of *K* conditions to the occurrence of the outcome *Y*, as in (7.16)

$$\bigcap_{j=1}^{K} A_j \rightarrow Y \tag{7.16}$$

The formula means that ⌜were it the case that these *K* conditions together make an *inus* machine, then the outcome should certainly occur in an ideal instance displaying them all in the right state, and fail otherwise⌝. For it to hold, the starting hypothesis should contain the sufficient bundle to the positive and the negative outcome, which may have different specifications. QCA acknowledges this fact and addresses the positive and the negative outcomes in separate analyses. Nevertheless, the two sets of findings are related as long as both follow from the same truth table in which primitives are exclusively assigned to one outcome, and no contradiction is detected.

The value of an explanatory QCA lies in identifying the *plausible* bundle beneath the success and failure of an outcome in the population of interest, to define the tenability of the starting hypothesis and its underlying theory. Its identification procedure addresses validity issues as the underspecification or the overspecification of the starting hypothesis.

7.3.2.2 Tackling Underspecification

QCA deploys truth tables as a diagnostic device for detecting underspecification. Therefore, QCA's truth tables are partially different from those of logic.

A QCA's truth table contains as many columns as *inus* conditions in the hypothesis, plus one for the outcome and at least three additional columns for as many parameters of fit. The truth value of the outcome is the last column to be filled, depending on the researcher's decisions about the parameters, as follows:

Decision 1: Frequency Cut-Off

This parameter establishes whether a primitive is observed or realized in the universe of reference based on the minimum number of its "best instances" (Ragin, 2008). A unit is the best instance of the primitive in which it gets a membership scorehigher than 0.5 according to the weakest link rule (7.6).

Units' classification yields two kinds of primitives: *observed* or *realized*, and *unobserved* or *unrealized*. The unrealized ones are also known as *logical remainders* and constitute a common occurrence. Although the ratio of units to conditions inevitably plays a role in raising them (Marx & Duşa, 2011), their number is relatively independent of the richness of the hypothesis or the size of the universe. Instead, the logical remainders expose the *limited diversity* of the units under analysis and serve as a source of counterfactual reasoning (Ragin, 2008; see below).

The researcher's decision regarding the frequency cut-off may also increase the number of unrealized primitives. Conventionally, one best instance is enough to declare a primitive realized albeit rare. However, the frequency cut-off can be raised if the numerosity of the population and the gauging strategy suggest a risk of errors in units' classification.

Decision 2: The Consistency Threshold

The second of the researcher's decisions on the truth table for a QCA concerns the assignment of the realized primitives to either the positive or the negative outcome. In Standard QCA, the decision mainly follows considerations on consistency.

In line with consolidated axiomatizations (Hájek, 2011), QCA captures the *consistency of the sufficiency* of each primitive to an outcome (*S.cons* for short, also known as *incl* for "inclusion": Ragin, 2008; Schneider & Wagemann, 2012; Duşa, 2019) as an extensional gauge that checks for empirical violations of the impossibility requirement in (7.15) through the ratio in Eq. (7.17):

$$S.cons_{\omega_* \to Y} = \frac{|\omega_* \cap Y|}{|\omega_*|} \tag{7.17}$$

The vertical bars indicate the size of a partition. The denominator of the ratio is for any antecedent of interest—otherwise understood as the number of trials—and here corresponds to the primitive of interest. The numerator is for the number of successful trials, that is, the intersection of the primitive with the outcome. When none of the N units under analysis qualifies as an instance of the inconsistent intersection $\omega_* \overline{Y}$, the numerator overlaps the denominator, and the *S.cons* gets its highest value of 1.00, which supports the claim that ω_* is sufficient to Y. The lower the overlapping, the lower the *S.cons* parameter and the credibility of the claim of sufficiency.

The detection of critical inconsistencies justifies the dismissal of the hypothesis in the current shape as incomplete or otherwise misspecified (e.g., Rihoux & De Meur, 2009; Rohlfing, 2020). The textbook illustration comes from a configurational model applying Lipset's socioeconomic theory of democratization to account for the breakdown of democracy in Europe between the two World Wars. The model yielded a straightforward truth table with a single remarkable contradiction: the German case displayed all the socioeconomic conditions for a thriving democracy, but it experienced a clear regime breakdown. The contradiction disappeared after adding institutional conditions of government stability to the model.

The researcher's decision concerns the value of the *S.cons* below which the inconsistency is severe enough to preclude the assignment of the primitive to the outcome. An established convention suggests setting it at 0.85, although the range of *S.cons* values in the table may justify a different choice. An additional criterion considers "natural gaps"—that is, steep falls in the ordered series of the primitives' *S.cons* values. These gaps may suggest setting the consistency threshold in between clusters of primitives.

The primitives not assigned to Y cannot be automatically assigned to $\overline{Y}$. Instead, the consistency of each primitive has to be tested with both states of the outcome separately. Nevertheless, meaningful solutions can be expected when the realized primitives below the consistency cut-off to Y return high *S.cons* values to $\overline{Y}$. This suggests that the starting hypothesis can account for both the occurrence and the non-occurrence of the outcome consistently.

Decision 3: The Coverage Cut-Off

The least common and last of the possible researcher's decisions concerns the empirical import of the claim of sufficiency—how relevant the primitive is to the set of instances of the outcome of interest. The related parameter, dubbed *coverage of sufficiency* (*S.cov* for short) is calculated as in (7.18)

$$S.cov_{\omega_* \to Y} = \frac{|\omega_* \cap Y|}{|Y|} \tag{7.18}$$

When all the instances of a primitive ω_* display the outcome, the numerator in (7.18) equals the denominator, and the parameter takes its highest value of 1.00 supporting the claim that the primitive accounts for any unit with the positive outcome. But the empirical relevance of a factor to an outcome is the extensional gauge of its necessity in the cases at hand. Hence, the *S.cov* of ω_* to *Y* gauges the *consistency of necessity* (*N.cons* for short) of the primitive to the outcome. Specularly, the *S.cons* of ω_* to *Y* gauges the empirical relevance of the primitive as a necessary compound to the outcome—and hence counts as the *N.cov* of ω_* to *Y*.

A primitive's *S.cov* value decreases with the increase in the evidence that the outcome can occur without the primitive. Coverage cut-offs may be established to ensure the analysis is based on sufficient primitives that also are empirically relevant. However, decisions driven by empirical relevance may prove unwise, as even rare primitives may contribute to specify the composition of *inus* machines.

7.3.2.3 Tackling Overspecification

Overspecification depends on having included factors in the starting hypothesis that prove irrelevant to account for the units' diversity.

The issue arises as mistaking some features for an *inus* component entrenches solutions in very specific contexts and unnecessarily reduces their portability (e.g., Craver & Kaplan, 2020; Salmon, 2020; cfr. Álamos-Concha et al., 2021; Chap. 10).

The acknowledged sources of overspecification are twofold: irrelevant components, and trivial factors.

Irrelevant Components

Quine-McCluskey's *minimizations* provide the standard approach to irrelevant conditions (Ragin, 1987/2014, 2000, 2008). These minimizations identify irrelevant components in the single varying conjunct of two otherwise identical primitives. To witness, the minimization is possible of the primitives $ABCD$ and $ABC\overline{D}$ if both display high *S.cons* values to the same outcome. The formal reason is that the two allow the factorization $ABC\left(D \cup \overline{D}\right)$, where $D \cup \overline{D} := \mathbb{U}$ by Eq. (7.9). The operation highlights that the *implicant ABC* is sufficient to *Y* regardless of *D*, which can be dismissed as not *inus* a factor.

The adjudication of the *inus* nature of single components may change depending on how minimizations deal with the logical remainders. The Standard Analysis affords three alternative *counterfactual assumptions*, each leading to "solutions" at different degrees of specification, as follows:

- *Conservative or complex solutions*. These are returned under the assumption that unrealized logical remainders would have proven ambiguous had they been realized. Hence, minimizations only operate on observed primitives. With high limited diversity, the solutions could be as rich as the disjunction of any realized primitive.

- *Parsimonious solutions*. A superset—and hence, more general in scope—of the conservative solutions, the parsimonious solutions are returned under the assumption that any logical remainder could prove sufficient if matching a realized primitive except for one literal.
  The surviving factors are the *inus* components in the hypothesis that are essential to account for the difference between the instances of the successful outcome and the instance of the failed one.
  However, parsimonious minimizations can yield gappy explanations. Like the treatment variable in the Potential Outcome Framework (see Chap. 3) or the mediators in Path Analysis (see Chap. 6), the solutions from the parsimonious minimization may capture a causal channel, but certainly dismiss the information about the covariates needed to account for the effect (Damonte, 2021b). The reason is that the parsimonious minimizations drop factors regardless of the plausibility of the logical remainders that they employ.
- *Intermediate or plausible solutions.* These are returned under the assumption that only those logical remainders qualifying as *easy counterfactuals* would have proven sufficient if realized.
  To understand the difference between an easy and a hard counterfactual, imagine the following. At the outset, we include condition A in the starting hypothesis under theoretical and empirical reasons to assume that it is an *inus* factor. More specifically, we assume that the condition makes an unknown causal compound Φ sufficient to the outcome Y when given in a state, say A, while in the opposite state, say $\overline{A}$, it turns Φ into a failure machine. In short, we add A under the *directional expectations* that

  (i) $A\Phi \subset Y$; and

  (ii) $\overline{A}\Phi \subset \overline{Y}$,

  where $\subset$ indicates a subset.

After we build and populate the truth table, we find the primitive $\omega_1 = ABCD$ is observed with an *S.cons* of 1.00 to Y, while we do not observe (hence we star) the primitive $\omega_9^* = \overline{A}BCD$. According to the single difference rule, ω_1 and ω_9^* can be minimized to BCD. However, the minimization entails that ω_9^* is consistent with Y, and hence that $\overline{A}\Phi$ would yield Y if observed. This goes against our directional expectation (*ii*) and makes a *hard* or *implausible counterfactual* of ω_9^*.

Now imagine the primitive $\omega_{13} = \overline{A}\overline{B}CD$ is realized with an *S.cons* of 1.00 to Y, while the primitive $\omega_5^* = A\overline{B}CD$ is a logical remainder. Again, according to the single difference rule, ω_{13} and ω_5^* can be minimized to $\overline{B}CD$. The minimization entails that ω_5^* is consistent with Y and that $A\Phi$ would yield the outcome if observed. This agrees with our directional expectation (*i*); hence, ω_5^* qualifies as an *easy* or *plausible counterfactual*.

Intermediate minimizations return solutions from observed primitives and easy counterfactuals only. The factors added to the parsimonious solution terms may not be essential to preserve the non-contradictoriness of the compounds. As they improve the sufficiency of the implicant, they offer a more complete account of why the outcome failed in specific units while succeeding in others (Ragin, 2008; Fiss et al., 2013; Duşa, 2019; Oana & Schneider, 2018; Damonte, 2021a; cfr. Baumgartner, 2015; Baumgartner & Thiem, 2020).

A Note on Ambiguity in Solutions

Regardless of the usage of the logical remainders, it has been emphasized that solutions in Standard QCA may encounter problems of ambiguity as the same primitives to an outcome may yield different prime implicants. To witness, the primitives $ABC, AB\overline{C}, A\overline{B}C$ can legitimately be minimized as $AB \cup A\overline{B}C$ or $AC \cup AB\overline{C}$. The information is displayed in a *Prime Implicant Chart* that shows which prime implicant covers which primitive, as displayed in Table 7.2.

Originally, the PI Chart was devised to allow the researchers making a decision on which implicants could be retained in solutions in light of their theoretical import. The practice has been deprecated, as cherry-picking implicants may build a confirmation bias into solutions (e.g., Baumgartner & Thiem, 2020; Baumgartner, 2015), and the current good practices require that alternative implicants are reported, too. Besides, the alternative minimizations may contain information of interest for discussion. For instance, in the example above, the two solutions indicate that *A* is always required—it can be an enabling condition—but, in the cases at hand, it obtains in team with *B* or *C*—which can play as triggering conditions. The richer implicants $A\overline{B}C, AB\overline{C}$ add that the one trigger can compensate for the absence of the other. These two richer implicants are currently left implicit by the reporting conventions that reward lean solutions. Under these rules, privileged prime implicants are those terms that, together, maximize the coverage of primitives—as are *AB*, *AC* in Table 7.2. Indeed, the conclusion that the union $AB \cup AC$ obtains the outcome does justice to alternative minimizations while logically entailing the richer implicants. Still, the information in the PI Chart deserves some attention, for it may suggest more accurate causal interpretations.

Table 7.2 Example of Prime Implicant Chart

Primitives *Implicants*	ABC	$A\overline{B}C$	$AB\overline{C}$
AB	x		x
AC	x	x	
$AB\overline{C}$			x
$A\overline{B}C$		x	

Dealing with Trivial Factors

Trivial factors are degenerate necessary conditions, that is, limiting cases of supersets. These arise when all or almost all the units in the universe of reference make the same state of the condition true—in short, when their distribution is skewed or constant.

Trivial factors can be detected by plugging the size of one condition in the place of the primitive in the formulas of the *N.cons* as in (7.18). When all the instances of the tested condition display the outcome, the numerator equals the denominator, and the parameter takes its highest value of 1.00, supporting the claim that the condition is necessary to the outcome. Conditions with a score of *N.cons* higher than 0.95 can be tested for skewness through a further parameter dubbed *Relevance of Necessity* (*RoN*: Schneider & Wagemann, 2012) and calculated as in (7.19) below:

$$RoN_{A \leftarrow Y} = \frac{|1 - A|}{|1 - A \cap Y|} \tag{7.19}$$

The parameter takes its lowest scores when the distribution of the condition by the outcome of reference proves trivial—when the size of $1 - A$ is remarkably smaller than the size of $1 - A \cap Y$, indicating the instances of the negative outcome raise independently of the absence of the condition. The standard recommendation is to consider dropping the factors with *N.cons* close to 1.00 and low *RoN* from the hypothesis. Thus, such "analysis of necessity" is a recommended step to be performed ahead of constructing the truth table (Schneider & Wagemann, 2012).

The original expected advantage was of pinpointing those constant conditions that double the number of primitives in the truth table while leaving almost half of them unobserved and lowering the consistency of every solution. However, the dismissal of a quasi-constant may prove unwise if the model requires it to prevent contradictory primitives (Rohlfing, 2020). The essentiality of the contribution can be easily ascertained by verifying whether a change in the consistencies of the primitives occurs after the seemingly trivial condition is dropped from the hypothesis (Damonte, 2021a). Nevertheless, the calculation of the parameters of fit on individual conditions remains a crucial source of information, as their values can support directional expectations or suggest reconsidering them.

7.4 Soundness

The actual link between sets, predicates, and the real world is decided by how truth values are assigned to literals—that is, by gauging.

The standard assumption in representation measurement theory maintains real-world properties depend on some units' deep structure that we can know indirectly only as meaningful variations in related observable attributes. This theory assumes we can represent these attributes through *numerical images* and capture their variation through adequate scales. Scales warrant that for any manifestation p_i of the property P in the unit u_i there is a measure q_i of the image Q such that the functional relationship between measures preserves some fundamental relationship in the variation of the attribute.

The seminal work of Stevens (1946) pinpointed four such fundamental relationships: sameness, rank, distance, and proportion, preserved by nominal, ordinal, interval, and ratio scales, respectively. Conventional textbooks have long taught that a hierarchy of scope exists among measurements with the ratio scale at the top as the most "robust" one—i.e., abstracted from actual entities and their contexts. Intended as a prudential rule for naive statisticians (e.g., Luce, 1959), the hierarchy has turned into a canon and, as such, has been disputed since its introduction. Indeed, any measurement entails a *loss function*, and the loss is admissible that allows retaining crucial information (e.g., Guttman, 1977). Thus, prominent comparatists contend that ratio scales prove robust for detecting fine-grained changes, but sacrifice the information on "critical points." The qualitative change that occurs in the state of a unit when the measure of a crucial attribute reaches a special value is better conveyed by nominal scales (e.g., Sartori, 1984, 1991; Collier & Mahon, 1993; Ragin, 2000; Goertz, 2020).

In short, scales entail a trade-off between *precision* and *meaning*. However, the trade-off can weaken when metric variables are remapped as *fuzzy sets*.

7.4.1 Gauging for QCA: The Theoretical Side

7.4.1.1 The Starting Point

Zadeh (1968, 1978) introduced fuzzy sets to widen the scope of algorithmic problem-solving. He noted how machines could deliver precise solutions, but limited to trivial problems, while the human brain tackles complex issues through linguistic structures with hazy *hedges* such as ⌜*very*⌝, ⌜*somewhat*⌝, or ⌜*almost*⌝.

Fuzzy scores translate hedges into weights (μ) ranging from 0.00 to 1.00 to convey the degrees of membership of u_i to the set of A instances. They, too, understand the membership in a set and its opposite as complements, calculated as in (7.20):

$$\mu_{i \in \bar{A}} = 1.00 - \mu_{i \in A} \tag{7.20}$$

where $\in$ reads ⌜*in*⌝.

The meaning of the relationship between complements is established by a third relevant value, the *crossover*. Conventionally weighing 0.50, the crossover is the point of neutrality and signals a membership neither in the set nor in its complement.

Logically, fuzzy scores capture the *possibility* that the statement ⌜is A⌝ is true for the actual unit u_i: 1.00 indicates the statement is *certainly* true; 0.00 indicates the statement is *certainly not* true; 0.50 indicates that the positioning of u_i is *highly* *ambiguous* given the observation. Therefore, original fuzzy scores defy a strictly bivalent logic. The advantage is that the three points allow alignment of linguistic hedges, sets, and metric variables through a triangular, trapezoidal, or bell-shaped function. This *filter function* maps the raw values ν_A—e.g., age in years—into fuzzy scores μ_A—e.g., membership in the set ‹YOUNG›—so that it conveys the certainty that a 16-year-old is in the set and a 36-year-old is almost so.

To map meanings onto fuzzy scores, then, the researcher needs to establish

- The raw value of the *inclusion* threshold, α. The threshold truncates any variation above α as irrelevant: for any value higher than α, the unit u_i does qualify as an instance of the set and takes 1.00 as its fuzzy score.
- The raw value of the *exclusion* threshold, β. The threshold truncates any variation below β as irrelevant: for any lower values, the unit u_i does not qualify as an instance of the set and takes the fuzzy score of 0.00.
- The raw value of the *crossover* γ, which makes the classification of u_i uncertain and corresponds to the fuzzy score of 0.50. In Zadeh's original system, the raw value of the crossover is the arithmetic mean of α and β.

7.4.1.2 Ragin's Reinvention

For QCA, Zadeh's original proposal is affected by a twofold ambiguity. First, linguistic hedges are seldom clearly ordered, and a straightforward correspondence with particular fuzzy scores can prove idiosyncratic. Second, triangular, trapezoidal, or bell-shaped relations can make each fuzzy score μ_A correspond to more than one raw scores on ν_A, which makes it hard to retrieve the raw value from the fuzzy score.

Ragin's fuzzy sets avoid these issues with a gauge that, before rendering natural language, includes both pieces of information of interest to comparatists—those of "differences in degree," and of "differences in kind" (Ragin, 2000). His filter functions are monotonic non-decreasing, which re-establishes the isomorphism of raw values, fuzzy membership scores, and selected hedges—as in Table 7.3.

The remapping of raw variables into fuzzy scores is especially illuminating of Ragin's rationale of conversion. He portrays it as an operation of *calibration*—defined as the fine-tuning of an instrument to improve the validity of its measurements. Although the concept best applies to continuous variables, the calibration rationale also informs the transformation of qualitative data into fuzzy scores (e.g., De Block & Vis, 2019). Indeed, the instrument to be fine-tuned is the filter function, whose shape can be decided using different methods (Ragin, 2000, 2007, 2008:96; Dușa, 2019).

The *indirect method of calibration* assigns the same "qualitative score" from a scale such as (*c*) or (*f*) in Table 7.3 to groups of cases with similar raw values. Then, the cases' raw scores may or may not be filtered into predicted fuzzy scores through the qualitative scores by fractional polynomial regression.

Table 7.3 Possible positions of u_i to A, and corresponding membership values μ_A

<table>
<tr><th rowspan="2">Position</th><th colspan="6">μ_A</th></tr>
<tr><th>(a)</th><th>(b)</th><th>(c)</th><th>(d)</th><th>(e)</th><th>(f)</th></tr>
<tr><td>Fully in</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>Mostly in</td><td></td><td></td><td></td><td></td><td rowspan="2">4/5</td><td>5/6</td></tr>
<tr><td>More in than out</td><td></td><td></td><td rowspan="2">2/3</td><td>3/4</td><td rowspan="2">4/6</td></tr>
<tr><td>More or less in</td><td></td><td></td><td></td><td>3/5</td></tr>
<tr><td>Neither in nor out</td><td></td><td>1/2</td><td></td><td>2/4</td><td></td><td>3/6</td></tr>
<tr><td>More or less out</td><td></td><td></td><td rowspan="2">1/3</td><td></td><td>2/5</td><td rowspan="2">2/6</td></tr>
<tr><td>More out than in</td><td></td><td></td><td>1/4</td><td rowspan="2">1/5</td></tr>
<tr><td>Mostly out</td><td></td><td></td><td></td><td></td><td>1/6</td></tr>
<tr><td>Fully out</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
</table>

Source: Ragin (2000:156, 2009)

The *direct method of calibration*, on the other hand, stipulates that the filter function is a growth curve of odds. The smoothness of the slopes is decided every time by suitable raw values for α_A, γ_A, β_A. These chosen raw scores are pegged to conventional fuzzy values, fixed at 0.953, 0.500, 0.047, respectively. The log-odds of $\mu\alpha$ are $\ln\left(\frac{0.953}{1-0.953}\right)=3$, while those of $\mu\alpha$ are $\ln\left(\frac{0.047}{1-0.047}\right)=-3$; thus, the fuzzy membership of the i-th unit with raw value ν_i is calculated as in (21) below:

$$\mu_i = \begin{cases} \dfrac{e^{3\frac{\nu_i-\gamma}{\alpha}}}{1+e^{3\frac{\nu_i-\gamma}{\alpha}}}, & \nu_i > \gamma \\ 0.5, & \nu_i = \gamma \\ \dfrac{e^{-3\frac{\nu_i-\gamma}{\beta}}}{1+e^{-3\frac{\nu_i-\gamma}{\beta}}}, & \nu_i < \gamma \end{cases} \tag{7.21}$$

Ragin's fuzzy sets can be conceived of as crisp sets weighted by a *classification error*. As such, they convey both qualitative and quantitative information, circumventing the trade-off between scales. Indeed, the crisp classification still holds with fuzzy scores, following the rule of conversion in (7.22):

$$A_i = \begin{cases} 1, & \mu_{i\in A} > 0.50 \\ 0, & \mu_{i\in A} < 0.50 \end{cases} \tag{7.22}$$

where A_i is the crisp membership of the i-th unit in the set A, while $\mu_{i \in A}$ is the fuzzy membership of the same i-th unit in the same set.

The preservation of crisp sets' qualitative information by QCA's fuzzy scores is further ensured by the convention that the crossover shall not be assigned to any actual unit of analysis—or of dropping the 0.5-instances under the argument that they cannot bring helpful information in the analysis (Ragin, 2008; Duşa, 2019).

Furthermore, the basic rules for calculating intersection and union as in (7.6) and in (7.10) also apply to fuzzy sets. However, fuzzy scores cannot meet the axiom of strong identity (7.1); instead, they follow the more common version (7.23) below, meaning that sameness is preserved for units with the same score.

$$A_i := A_i \tag{7.23}$$

The principles of non-contradiction and excluded middle again hold with fuzzy scores in a crisp understanding, as clarified by (7.24) and (7.25):

$$\mu_{i \in \left(A \cap \bar{A}\right)} < 0.5 \tag{7.24}$$

$$\mu_{i \in \left(A \cup \bar{A}\right)} > 0.5 \tag{7.25}$$

It is worth noting that the size of a fuzzy union calculated by (7.6) is usually smaller than its crisp versions, while the size of a fuzzy intersection calculated by (7.10) is usually larger than its crisp version due to the *residuals* that fuzzy scores leave in the partition.

7.4.1.3 Fuzzy Sufficiency and Necessity

With fuzzy scores, subset relationships are established as the *containment* (Ragin, 2000; cfr. Zadeh, 1978) of membership functions.

Therefore, fuzzy-set sufficiency is captured by Eq. (7.26):

$$\mu_{i \in \omega_{.}} < \mu_{i \in Y} \tag{7.26}$$

Equation (7.26) entails that, if we plot our units on a Cartesian plane defined by the membership scores in $\omega_{.}$ as the x-axis and the membership scores in Y as the y-axis, if $\omega_{.}$ is sufficient to Y, it distributes the units *above* the bisector in an *upper-triangular* shape.

Instead, fuzzy-set necessity corresponds to (7.27):

$$\mu_{i \in \omega_{.}} > \mu_{i \in Y} \tag{7.27}$$

Equation (7.27) means that the antecedent $\omega_{.}$ that is necessary to Y distributes the units *below* the bisector in a *lower-triangular* shape.

By extension, the relationship of necessity and sufficiency arises when the units' membership scores in a primitive (or implicant, or condition) equal those in the outcome, distributing the units *along* the bisector in a *linear* shape.

The *S.cons* parameter preserves its meaning with fuzzy scores, although they can blur the *recognition* of violations as the residuals $\mu_{i\in(Y\cap\overline{Y})}$ inflate their values. The *Proportional Reduction of Inconsistency* (*PRI*: Ragin, 2008; Schneider & Wagemann, 2012) has been introduced to deflate and complement the information from the *S.cons* calculated with fuzzy scores. The parameter builds on the rationale of the proportional reduction of error commonly employed to determine whether the information about *A* improves our prediction of *Y* (e.g., Menard, 1995). It reads as in (7.28):

$$PRI_{\omega_*\to Y} = \frac{\left|\omega_* \cap Y\right| - \left|\omega_* \cap Y \cap \overline{Y}\right|}{\left|\omega_*\right| - \left|\omega_* \cap Y \cap \overline{Y}\right|} \tag{7.28}$$

where the vertical bars again indicate the size of the fuzzy partition as the sum of the units' fuzzy membership scores in the partition—such that, for instance, $\left|\omega_*\right| := \sum_{i=1}^{N} \mu_{i\in\omega_*}$.

The set-theoretical task of the *PRI* is to establish whether the conditional relationship holds, net of fuzzy residuals. It takes the same value as the *S.cons* when the size of the residuals is null $\left|Y \cap \overline{Y}\right| = 0.00.$ It degenerates when the units systematically display higher residuals than membership in the primitive: $\mu_{i\in(Y\cap\overline{Y})} > \mu_{i\in\omega_*}$. Last, it takes lower values than the *S.cons* when the units' residuals are non-null and lower than the membership in the primitive: $0 < \mu_{i\in(Y\cap\overline{Y})} < \mu_{i\in\omega_*}$.

A *PRI* value sensibly lower than the corresponding *S.cons* points to inconsistencies that may justify the exclusion of the primitive from minimizations—or the reconsideration of gauges, conditions, or the starting hypothesis.

7.4.2 Gauging for QCA: The Empirical Side

Whether fine-grained membership scores properly render an *inus* factor only depends on how we construe our gauge—here, on how we set the thresholds. Thresholds elicit a solution to the problem of aligning the extension and the intension of an attribute (Quine, 1982; Sartori, 1984; Goertz, 2020).

A theory-driven approach to the problem clarifies the intension first to prevent the risk of stretching attributes beyond their meaning, which would introduce more hidden heterogeneity than would be desirable for the analysis (see Chap. 10). At the same time, thresholds may spoil the analysis when they enforce some ideal yardstick that none of the units can meet. In short, theoretical thresholds can become useless when decisions are not fine-tuned to actual diversity.

QCA scholars have developed several recommendations to balance these opposite risks. The recommendations assist the researcher in tackling three intertwined problems—namely, unit selection, the operationalization of causal properties, and the identification of thresholds that align meanings and empirics. In actual research, the point of attack may change; however, the resulting membership scores provide a single solution to all three issues—likely, after some iteration.

7.4.2.1 Establishing the Universe of Reference

As in any technique, units of observation provide as solid an empirical ground to the analysis as the criteria for their selection. Such criteria should prevent or minimize the later rise of threats to credible results (e.g., Geddes, 1990; Goertz, 2020).

In explanatory QCA, case selection has to ensure enough diversity to capture the causal facts of interest. Thus, the criterion cannot exclusively focus on the dependent or the independent. Units selected on the outcome of interest would artificially prevent inconsistencies—thus making the validity of results undecidable. On the other hand, units selected on the factor of interest would turn it into a constant background feature and make its causal contribution undecidable. Hence, the first criterion that unit selection shall meet is the *variability* in realized states and combinations of factors.

The broadest variability follows from open universes, but open universes may endanger the preservation of meaning (i.e., Ragin, 2008). Geographical, historical, and cultural boundaries provide the closure of the units' heterogeneity required for making interpretable decisions about thresholds. Indeed, different α, β, γ may be needed to establish whether a country qualifies as <RICH>, <DEMOCRATIC>, or <EQUAL> in different world regions and time frames. Therefore, the second and related criterion for unit selection consists of finding the meaningful *scope condition* that encloses the universe of reference and ensures interpretable membership scores. In short, the correspondence of meaning and numbers comes at the cost of a restriction in the scope of the analysis—and in the generalizability of results (e.g., Goertz, 2017; Walker & Cohen, 1985; Verweij & Vis, 2021; Findley et al., 2021). The limitation, however, might not apply to the starting explanatory hypothesis, which may travel farther than its operational specifications.

7.4.2.2 Operationalizing Intension

The operation of connecting gauges and attributes meaningfully is seldom straightforward. Again, it opens to two opposite risks of providing too a specific or generic definition of an attribute (e.g., Sartori, 1984; Ragin, 2008).

Hyper-Specificity

The fallacy of composition occurs when we recognize each "token" empirical manifestation as a different property and build a plethora of conditions with too narrow an extension (e.g., Menzies, 2004; Craver & Kaplan, 2020; cfr. Chap. 10). The problem can be solved by recognizing functional equivalences, climbing the ladder of abstraction, and gathering functionally equivalent manifestations under a single label.

Verba (1967) elaborates on the point by discussing how case-based evidence can be turned into a causal factor. From the historical report on how the eruption of Mount Vesuvius had a significant impact on the stability of the Pompeiian political system, we may identify either ‹ERUPTION› or ‹CALAMITY› as a relevant *inus* factor; however, the latter includes the former and accommodates a broader number of functionally alternative sources of disruptions, thus widening the scope of comparisons.

According to Verba, an even better operationalization shifts the attention from contextual conditions to the properties of the unit of analysis. Instead of gauging the sources of disruption, the operationalization can narrow on those resources and arrangements that make the system respond to disruption effectively. From this viewpoint, ‹RESILIENT› better contributes to an explanatory theory of political systems' stability than ‹CALAMITY›. The system attribute can apply to the Pompeiian case, but travel farther across contexts.

Hyper-Generality

The second and opposite problem arises when the properties are encompassing to the point of losing their analytic capacity.

The problem often arises when the available measure of a concept is a composite of predictors, enabling factors, proxies, outputs, and outcomes. Such assorted content can make these composites apply "*everywhere*, as any universal should" but also "to *everything*." As a result, we incur "theoretically, a 'nullification of the problem' and, empirically, what may be called 'empirical vaporization'" (Sartori, 1991; Chap. 9; cfr. Collier & Mahon, 1993).

QCA detects these composites as trivial conditions and suggests they can be dismissed. However, composites may contain relevant explanatory information. The *inus* standing of selected components can be decided by their consistency to the outcome and by minimizations. In addition or as an alternative, suitable rules of composition by disjunction and conjunction may be devised to compress sub-properties into "superconditions" (Elman, 2005; Berg Schlosser & De Meur, 2009; Goertz, 2017; Damonte & Negri, 2019).

The Problem of Missing Values

Often, available raw measures are plagued with missing values. QCA's algorithm technique cannot handle them clearly, as the units for which the value is missing would belong to two primitives. This ambiguity can be tackled by running parallel analyses to verify whether the different classifications result in different solutions. If not, the unit and its partial information would prove irrelevant. When different classifications affect solutions—for instance, because they decide whether a primitive is realized or not—the information proves relevant, but the problem arises of how to decide between the two solutions.

Missing raw values require some credible criterion of adjudication. Alternatively, the measure can be substituted with a complete gauge of the same intension, if any. Last, the unit can be dropped from the analysis (Ragin, 2008; Basurto & Speer, 2012; Duşa, 2019). The move may increase the number of logical remainders, but remainders can be more adequately addressed with counterfactual rules in minimization.

7.4.2.3 Identifying Membership Thresholds

Thresholds explicate the rule that establishes a unit to be an instance of the set given its raw value. The default recommendation is to anchor these decisions on external theories and conventions (Ragin, 2000, 2007, 2008).

Special values of national and international policy indicators—for instance, household income to establish the risk of poverty; the share of people in an age cohort in education or training to expect a certain quality of society; the share of debt to revenue to establish the credibility of a borrower—may offer accepted anchorages to calibration decisions. However, conventional knowledge may evolve at a slower pace than actual phenomena. Under particular contingencies or within special areas, its usage for calibration may return skewed membership scores that would not survive the *RoN* test. Besides, a conventional tipping point may coincide with some units in the population, making them uninformative.

To avoid these issues, conventional knowledge can be adjusted in light of distributional considerations (Ragin, 2008). Although descriptive statistics lack qualitative meaning, considerations about quintiles seem unavoidable in large-N studies or whenever previous knowledge is wanting (e.g., Ragin & Fiss, 2017). A supplementary strategy—and consistent with the concern for non-contradictory partitions—prescribes cluster analysis to identify the raw values to be used as thresholds. The underlying rationale maintains that units close to each other belong to the same partition—and hence, that thresholds lie in the "natural gaps" between clusters.

Although long offered as a standard function for threshold setting by many software packages (e.g., Duşa, 2019), cluster analysis has driven concerns that its application might convey a deceiving sense of certitude about calibration and solutions. The risk of overconfidence can also increase when the membership scores are assigned directly following one of the scales in Table 7.3. Indeed, the researcher's classification error can always affect scoring operations in unknown directions.

To keep the risk at bay, zooming into the units around a threshold can help to support decisions with empirical knowledge when the number of cases allows it (Ragin, 2000; De Block & Vis, 2019). Frontier literature has also developed on false negatives and false positives in solutions (Braumoeller, 2015; Rohlfing, 2018) and on alternative filtering functions (Thiem, 2010). A further strategy suggests ascertaining the "robustness" of the solutions by running parallel analyses under different perturbations of units and thresholds (Marx & Duşa, 2011; Maggetti & Levi-Faur, 2013; Duşa, 2019; Oana & Schneider, 2018).

Many of these considerations are more justified in exploratory than in explanatory applications of QCA. When the driving concern is the preservation of particular meanings, seldom different gauges can render it equally well. To witness, Ostrom's theory of corruption maintains that people's perception of ineffective monitors and sanctions drives the belief of diffused wrongdoing that invites resorting to corruption along the lines of a self-fulfilling prophecy. In testing the tenability of this theory, the indexes of inefficiency in administration often used as a proxy of corruption are less suitable gauges of the phenomenon to be explained than the measures of perceived corruption.

In explanatory usages, however, coder's biases are possible, and this possibility can be explored by simulating some systematic tendencies toward strictness, generosity, confidence, or coyness in assigning membership scores. These tendencies can be rendered by calculating the *concentration* (7.29), *dilation* (7.30), *intensification* (7.31), or *moderation* (7.32) of the original fuzzy scores (Smithson & Verkuilen, 2006):

$$\mu_{i\in A}^{Conc} = \mu_{i\in A}^{2.0} \tag{7.29}$$

$$\mu_{i\in A}^{Dil} = \mu_{i\in A}^{0.5} \tag{7.30}$$

$$\mu_{i\in A}^{Int} = \begin{cases} \mu_{i\in A}^{0.5}, & \mu_{i\in A} > 0.5 \\ \mu_{i\in A}^{2.0}, & \mu_{i\in A} < 0.5 \end{cases} \tag{7.31}$$

$$\mu_{i\in A}^{\text{mod}} = \begin{cases} \mu_{i\in A}^{2.0}, & \mu_{i\in A} > 0.5 \\ \mu_{i\in A}^{0.5}, & \mu_{i\in A} < 0.5 \end{cases} \tag{7.32}$$

These transformations expose the worsening or the improvement that coders' biases can impart to solutions. They prove that truth tables and solutions inevitably change with scoring strategies—and the intensification, by bringing the fuzzy truth table closer to its crisp version, inevitably enhances the consistency and symmetry of observed primitives. In the end, the relative fragility of findings mirrors the specificity of our operationalization —but also its local value. It counts less as a problem of the technique or the algorithm than an issue in our knowledge, models, and gauging strategies.

7.5 Summing Up

To run a credible explanatory QCA, a researcher may want to

1. *Define the outcome of interest, the causal stories about its generative process, and the conditions that make it "certain."* This step implies reviewing the theoretical and empirical literature to find testable definitions of the outcome, and identifying a convincing (type of) data-generation mechanism beneath it. Based on the mechanism, triggering, enabling, and shielding *inus* conditions can be hypothesized that, jointly given in an ideal unit, would compel the generation process and ensure it unfolds unimpeded. This bundle provides the starting *inus* hypothesis.
2. *Identify the universe of reference and the raw variables that render the hypothesis, then declare the directional expectations about each factor*. Define a scope condition for a population ensuring meaningful units' diversity. Choose the raw measures at the proper level of abstraction to render each factor as faithfully as possible. Estimate the missing values, or discard the corresponding unit. Then, declare the directional expectations about the contribution of each factor to the occurrence and failure of the outcome.
3. *Turn raw data into membership scores*. Explore the variation in the raw measures; identify thresholds; assign membership scores to instances with proper operations. Different scaling may affect the assessment of set-relationships; consider applying the same scaling. Consider whether the specification of the hypothesis may benefit from the compression of some factors; in that case, add the new superconditions to the dataset. Calculate different datasets with diluted, concentrated, moderated, and intensified scores to run parallel analyses for robustness.
4. *Assess the claim of individual consistency.* Calculate the necessity parameters for single conditions against the outcome and its negation. Identify those conditions from the starting hypothesis with *N.cons* above 0.95 and low *RoN*, and fork the analysis by running the next steps with and without them. If compressed conditions obtain better *N.cons* and *N.cov* values than the original ones, consider dropping the latter. *N.cons* and *N.cov* values can also be used to establish whether the directional expectations stand in the population.
5. *Assess the claims of sufficiency.* Build the truth tables of the positive and negative outcome, assign instances to primitives, and calculate the *S.cons* and the *PRI* of the realized primitives. Check for inconsistent instances in configurations; if found, re-run the calibration. Be it of no help, add a further condition in line with the starting hypothesis to improve the consistency of each primitive to one outcome.
6. *Minimize*. Establish the cut-off in the values of *S.cons* below which the observed primitives will not be deemed consistent with the claim of sufficiency—in case, with the help of *PRI* values—to both the positive and the negative outcome. Find the conservative, parsimonious, and plausible solutions. Consider the difference in the composition of each prime implicant from the parsimonious and the plausible solution. If new conditions appear in the latter, check whether the *S.cons* values of the plausible solution are higher than the parsimonious. Higher consistency values indicate the addition is detectably meaningful, and the plausible solution is more credible than the parsimonious. If the additional conditions in

the plausible solution do not improve the *S.cons* values on the parsimonious, consider re-running the analysis from step 5 without these additional conditions to verify the robustness of minimizations.

7. *Plot the solutions to the outcome and its negation*. Check the fitting of the instances to the upper triangular shape, assuming the shape is met when instances fall above the $y = x + 0.1$ line (Ragin, 2000). Discuss which implicants explain which instances of the outcome. Consider the unexplained instances.
8. *Return to theory*. Consider the logical relationship between the solutions to the outcome and its negation. Identify the strategies that a negative instance can adopt to reach the closer positive group.
9. *Run re-analyses and extensions for robustness*. Run the analysis with different calibrations and scope conditions, and compare the raising of contradictory configurations, the change in necessity, the differences in solutions.

You can find the example here https://doi.org/10.5281/zenodo.7117973.
Enjoy your explanatory QCA!

Suggested Readings

The full-fledged version of the original proposal remains Charles C. Ragin, 2008. *Redesigning social inquiry: Fuzzy sets and beyond.* University of Chicago Press. An updated version and close to the original proposal is Patrick A. Mello's *Qualitative Comparative Analysis: An Introduction to Research Design and Application* (Georgetown University Press, 2021). A more case-oriented version is Ioana-Elena Oana, Carsten Q. Schneider, and Eva Thomann's *Qualitative Comparative Analysis Using R: A Beginner's Guide* (Cambridge University Press, 2021).

The detailed documentation of the R functions for QCA is in Adrian Dușa's *QCA with R: A comprehensive resource* (Springer, 2019). Additional functions are in Ioana-Elena Oana and Carsten Q. Schneider's *SetMethods: an Add-on R Package for Advanced QCA* (The R Journal https://doi.org/10.32614/RJ-2018-031).

The standards of transparency in reporting QCA are detailed in Schneider, Carsten Q., Vis, Barbara and Koivu, Kendra, 2019. *Set-Analytic Approaches, Especially Qualitative Comparative Analysis (QCA)*, https://doi.org/10.2139/ssrn.3333474

Review Questions

Section 7.2

(a) What is *inus* causation?
(b) What is an *inus* machine?
(c) How are the two concepts related to directional expectations?

Section 7.3

(a) What is a literal?
(b) What is a set?

(c) What is the relationship between the membership in a set and the truth value of a proposition?
(d) What is a truth table?
(e) How many primitives has a truth table of seven literals?
(f) Construe the truth table of literal A and the ⌜not⌝ connective.
(g) What does the principle of non-contradiction say?
(h) What does the weakest link rule say?
(i) How do you calculate the membership of a unit in an intersection?
(j) Construe the truth table of literals A, B, C, D and compute the truth function of the ⌜*and*⌝ connective.
(k) What does the principle of the excluded middle say?
(l) What does the strongest link rule say?
(m) Construe the truth table of literals A, B, C, D and compute the truth function of the ⌜*or*⌝ operator.
(n) What is the consistency of sufficiency?
(o) How can the consistency of sufficiency support the assessment of underspecification?
(p) What is the consistency of necessity?
(q) How can the consistency of necessity support the assessment of overspecification?
(r) What is in a parsimonious solution?
(s) What is a hard counterfactual, and what is an easy one? In which round of minimizations are they employed?

Section 7.4

(a) How do fuzzy scores accommodate qualitative and quantitative information?
(b) What are the shapes of the filter function in Zadeh's fuzzy sets, and how do they differ from Ragin's?
(c) What is the meaning of the inclusion and exclusion points in terms of relevant and irrelevant variation?
(d) What is the rule for turning fuzzy into crisp scores? Can we reverse the transformation?
(e) The membership score of u_1 in set A is 0.3. Calculate the value of its membership in the intersection $A \cap \overline{A}$.
(f) Do fuzzy scores violate the principle of non-contradiction?
(g) The membership score of u_1 in set A is 0.3. Calculate the value of its membership in the union $A \cup \overline{A}$.
(h) Do fuzzy scores stretch the principle of the excluded middle?
(i) What is the *PRI* for?
(j) How can you ascertain the robustness of configurational solutions?
(k) Calculate the concentrated, dilated, intensified, and moderated scores of unit u_i with original membership in Y of 0.9 and in A of 0.8.
(l) Calculate the *S.cons* of each transformation from exercise 11, and order them from the strongest to the weaker. Which fares better, and which worse?

References

Álamos-Concha, P., Pattyn, V., Rihoux, B., Schalembier, B., Beach, D., & Cambré, B. (2021). Conservative solutions for progress: On solution types when combining QCA with in-depth process-tracing. *Quality and Quantity*. https://doi.org/10.1007/s11135-021-01191-x

Amenta, E., & Poulsen, J. D. (1994). Where to begin. A survey of five approaches to selecting independent variables for qualitative comparative analysis. *Sociological Methods and Research, 23*(1), 22–53. https://doi.org/10.1177/0049124194023001002

Ansell, C., Doberstein, C., Henderson, H., Siddiki, S., 't Hart, P.: Understanding inclusion in collaborative governance: A mixed methods approach. Policy and Society 39(4), 570–591 (2020). https://doi.org/10.1080/14494035.2020.1785726.

Basurto, X., & Speer, J. (2012). Structuring the calibration of qualitative data as sets for qualitative comparative analysis (QCA). *Field Methods, 24*(2), 155–174. https://doi.org/10.1177/1525822X11433998

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

Baumgartner, M., & Thiem, A. (2020). Often trusted but never (properly) tested: Evaluating qualitative comparative analysis. *Sociological Methods & Research, 49*(2), 279–311. https://doi.org/10.1177/0049124117701487

Befani, B. (2013). Between complexity and generalization: Addressing evaluation challenges with QCA. *Evaluation, 19*(3), 269–283. https://doi.org/10.1177/1474022213493839

Berg Schlosser, D., & De Meur, G. (2009). Comparative research design: Case and variable selection. In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 19–32). London. https://doi.org/10.4135/9781452226569.n2

Boole, G. (1853). *An investigation of the Laws of thought on which are founded the mathematical theories of logic and probabilities*. Walton and Maberly.

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

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge University Press.

Cartwright, N. (2017). Causal powers. Why Humeans can't even be instrumentalist. In J. D. Jacobs (Ed.), *Causal powers* (pp. 9–23). Oxford University Press.

Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*(2), 367–405. https://doi.org/10.1037/0033-295X.104.2.367

Clarke, K. A. (2020). Logical constraints: The limitations of QCA in social science research. *Political Analysis, 28*(4), 552–568. https://doi.org/10.1017/pan.2020.7

Colby, M. E. (1991). Environmental management in development: The evolution of paradigms. *Ecological Economics, 3*(3), 193–213. https://doi.org/10.1016/0921-8009(91)90032-A

Collier, D., & Mahon, J. E. (1993). Conceptual stretching revisited. *American Political Science Review, 87*(4), 845–855. https://doi.org/10.2307/2938818

Craver, C. F., & Kaplan, D. M. (2020). Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science., 71*(1), 287–319. https://doi.org/10.1093/bjps/axy015

Damonte, A. (2013). Policy tools for green growth in the EU15: A qualitative comparative analysis. *Environmental Politics, 23*(1), 18–40. https://doi.org/10.1080/09644016.2013.817759

Damonte, A. (2021a). Gauging the import and essentiality of single conditions in standard configurational solutions. *Sociological Methods & Research, 50*(2), 683–707. https://doi.org/10.1177/0049124118794678

Damonte, A. (2021b). Modeling configurational explanations. *Italian Political Science Review, 51*(2), 18U 2–18U97. https://doi.org/10.1017/ipo.2021.2

Damonte, A., & Negri, F. (2019). Gauging fiscal worlds: How the EU countries balanced equality and wealth between 2007 and 2016. *Quality and Quantity, 53*(4), 1675–1692. https://doi.org/10.1007/s11135-018-00833-x

De Block, D., & Vis, B. (2019). Addressing the challenges related to transforming qualitative into quantitative data in qualitative comparative analysis. *Journal of Mixed Methods Research, 13*(4), 503–535. https://doi.org/10.1177/1558689818770061

De Meur, G., & Berg-Schlosser, B. (1994). Comparing political systems: Establishing similarities and dissimilarities. *European Journal of Political Research, 26*(2), 193–219. https://doi.org/10.1111/j.1475-6765.1994.tb00440.x

De Morgan, A. (1847). *Formal logic: Or, the calculus of inference, necessary and probable*. Taylor and Walton.

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

Elman, C. (2005). Explanatory typologies in qualitative studies of international politics. *International Organization, 59*(2), 293–326. https://doi.org/10.1017/S0020818305050101

Findley, M. G., Kikuta, K., & Denly, M. (2021). External Validity. *Annual Review of Political Science, 24*(1), 365–393. https://doi.org/10.1146/annurev-polisci-041719-102556

Fiss, P. C., Sharapov, D., & Cronqvist, L. (2013). Opposites attract? Opportunities and challenges for integrating large-N QCA and econometric analysis. *Political Research Quarterly, 66*(1), 191–198. jstor.org/stable/23563602

Geddes, B. (1990). How the cases you choose affect the answers you get: Selection bias in comparative politics. *Political Analysis, 2*, 131–150. https://doi.org/10.1093/pan/2.1.131

Goertz, G. (2017). *Multimethod research, causal mechanisms, and case studies: An integrated approach*. Princeton University Press.

Goertz, G. (2020). *Social science concepts and measurement: New and completely. Revised Edition*. Princeton University Press.

Guttman, L. (1977). What is not what in statistics. *The Statistician, 26*(2), 81–107. jstor.org/stable/2987957

Hájek, A. (2011). Conditional probability. In P. S. Bandyopadhyay & M. R. Forster (Eds.), *Handbook of philosophy of science* (*Philosophy of statistics*) (Vol. 7, pp. 99–135). North-Holland.

Hinterleintner, M., Sager, F., & Thomann, E. (2016). The politics of external approval: Explaining the IMFs evaluation of austerity programmes. *European Journal of Political Research, 55*(3), 549–567.

Huntjens, P., Pahl-Wostl, C., Rihoux, B., Schlüter, M., Flachner, Z., Neto, S., Koskova, R., Dickens, C., & Kiti, I. N. (2011). Adaptive water management and policy learning in a changing climate: A formal comparative analysis of eight water management regimes in Europe, Africa and Asia. *Environmental Policy and Governance, 21*(3), 145–163. https://doi.org/10.1002/eet.571

Kogut, B. K., & Ragin, C. C. (2006). Exploring complexity when diversity is limited: Institutional complementarity in theories of rule of law and national systems revisited. *European Management Review, 3*(1), 44–59. https://doi.org/10.1057/palgrave.emr.1500048

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

Lauri, T., Põder, K., & Ciccia, R. (2020). Pathways to gender equality: A configurational analysis of childcare instruments and outcomes in 21 European countries. *Social Policy & Administration, 54*(5), 615–863. https://doi.org/10.1111/spol.12562

Luce, R. D. (1959). On the possible psychophysical laws. *Psychological Review, 66*(2), 81–95. https://doi.org/10.1037/h0043178

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

Mackie, J. L. (1966). The direction of causation. *The Philosophical Review, 75*(4), 441–466. jstor.org/stable/2183223

Mackie, J. L. (1974). *The cement of the universe. A study of causation*. Clarendon Press.

Maggetti, M., & Levi-Faur, D. (2013). Dealing with errors in QCA. *Political Research Quarterly, 66*(1), 198–204. jstor.org/stable/23563603

Mahoney, J. (2021). *The logic of social science*. Princeton University Press.

Marx, A., & Duşa, A. (2011). Crisp-set qualitative comparative analysis (csQCA), contradictions and consistency benchmarks for model specification. *Methodological Innovations, 6*(2), 103–148. https://doi.org/10.4256/mio.2010.0037

Mello, P. A. (2021). *Qualitative comparative analysis: An introduction to research design and application*. Georgetown University Press.

Menard, S. (1995). *Applied logistic regression analysis*. Sage.

Menzies, P. (2004). Causal models, token causation, and processes. *Philosophy of Science, 71*(5), 820–832. https://doi.org/10.1086/425057

Most, B. A., & Starr, H. (2015). *Inquiry, logic, and international politics: With a new preface by Harvey Starr*. University of South Carolina Press.

Oana, I.-E., & Schneider, C. Q. (2018). SetMethods: An add-on R package for advanced QCA. *The R Journal, 10*(1), 507–533. https://doi.org/10.32614/RJ-2018-031

Oana, I.-E., Schneider, C. Q., & Thomann, E. (2021). *Qualitative comparative analysis using R: A Beginner's guide*. Cambridge University Press.

Pahl-Wostl, C. (2008). Requirements for adaptive water management. In C. Pahl-Wostl, P. Kabat, & J. Möltgen (Eds.), *Adaptive and integrated water management* (pp. 1–22). Springer. https://doi.org/10.1007/978-3-540-75941-6_1

Quine, W. V. O. (1982). *Methods of logic*. Harvard University Press.

Ragin, C. C. (1987/2014). *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. (2007). Measurement versus calibration: A set-theoretic approach. In J. M. Box-Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 174–198). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199286546.003.0008

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

Ragin, C. C., & Fiss, P. (2017). *Intersectional inequality: Race, class, test scores, and poverty*. University of Chicago Press.

Rihoux, B., & De Meur, G. (2009). Crisp-set qualitative comparative analysis (csQCA). In B. Rihoux & C. C. Ragin (Eds.), *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 33–68). Sage.

Rohlfing, I. (2018). Power and false negatives in qualitative comparative analysis: Foundations, simulation and estimation for empirical studies. *Political Analysis, 26*(1), 72–89. https://doi.org/10.1017/pan.2017.30

Rohlfing, I. (2020). The choice between crisp and fuzzy sets in qualitative comparative analysis and the ambiguous consequences for finding consistent set relations. *Field Methods, 32*(1), 75–88. https://doi.org/10.1177/1525822X19896258

Rosenberg, A. S., Knuppe, A. J., & Braumoeller, B. F. (2017). Unifying the study of asymmetric hypotheses. *Political Analysis, 25*(3), 381–401. https://doi.org/10.1017/pan.2017.16

Sabatier, P., & Mazmanian, D. (1980). The implementation of public policy: A framework of analysis. *Policy Studies Journal, 8*(4), 538–560. https://doi.org/10.1111/j.1541-0072.1980.tb01266.x

Salmon, W. C. (2020). *Scientific explanation and the causal structure of the world*. Princeton University Press. https://doi.org/10.1515/9780691221489

Sartori, G. (1984). *Social science concepts: A systematic analysis*. Sage.

Sartori, G. (1991). Comparing and Miscomparing. *Journal of Theoretical Politics, 3*(3), 243–257. https://doi.org/10.1177/0951692891003003001

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

Smithson, M., & Verkuilen, J. (2006). *Fuzzy set theory: Applications in the social sciences*. Sage.

Sprenger, J. (2011). Hempel and the paradoxes of confirmation. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Handbook of the history of logic* (Vol. 10, pp. 235–263). North-Holland. https://doi.org/10.1016/B978-0-444-52936-7.50007-0

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677–680. jstor.org/stable/1671815

Stiller, S. (2017). The interplay of actor-related strategies and political context: A fuzzy-set QCA analysis of structural reforms in continental welfare states. *Journal of European Public Policy, 24*, 81–99. https://doi.org/10.1080/13501763.2015.1118146

Stone, M. H. (1936). The theory of representations of Boolean algebras. *Transactions of the American Mathematical Society, 40*(1), 37–111. https://doi.org/10.1090/S0002-9947-1936-1501865-8

Thiem, A. (2010). Set-relational fit and the formulation of transformational rules. fsQCA. *COMPASSS WP Series, 2010*(61) http://www.compasss.org/wpseries/Thiem2010.pdf

Verba, S. (1967). Some dilemmas in comparative research. *World Politics, 20*(1), 111–127. jstor.org/stable/2009730

Verweij, S., & Vis, B. (2021). Three strategies to track configurations over time with Qualitative Comparative Analysis. *European Political Science Review, 13*(1), 95–111. Cambridge University Press. https://doi.org/10.1017/S1755773920000375

Walker, H. A., & Cohen, B. P. (1985). Scope statement: imperatives for evaluating theory. *American Sociological Review, 50*(3), 288–301. jstor.org/stable/2095540

Wittgenstein, L. (1922). *Tractatus Logicus Philosophicus*. London, UK.

Zadeh, L. A. (1968). Fuzzy algorithms. *Information and Control, 12*(3), 94–102. https://doi.org/10.1016/S0019-9958(68)90211-8

Zadeh, L. A. (1978). PRUF a meaning representation language for natural languages. *International Journal of Man-Machine Studies, 10*(4), 395–460. https://doi.org/10.1016/S0020-7373(78)80003-0