NCA-guided QCA:
A Methodology for Developing Stronger Causal Statements of Necessity and Sufficiency

Claude Rubinson
University of Houston—Downtown

Jan Dul
Erasmus University Rotterdam

Erasmus Research Institute of Management
Rotterdam School of Management
Erasmus University Rotterdam
October 4, 2022

Two Approaches to the Empirical Study of Causation

1) **Effects-of-causes**

   - Examines the effects of independent variables on a dependent variable.

   - Tends to align with quantitative, variable-oriented research.

   - Key question of causation: How do changes in the independent variables affect the dependent variable?

2) **Causes-of-effects**

   - Identifies the causes the produce a specific outcome.

   - Tends to align with qualitative, case-oriented research.

   - Key question of causation: What are the ***conditions*** under which an outcome does or does not occur?

Causes-of-Effects Analysis

"What are the conditions under which an outcome does or does not occur?"

1) **Scope conditions**
   - The conditions under which a theory is applicable.
   - Instantiated via case selection process by excluding irrelevant observations.

2) **Necessary conditions**
   - Causal condition must (almost always) be present for outcome to occur.
   - Empirical research approaches: comparative case study, qualitative comparative analysis (QCA), necessary condition analysis (NCA).
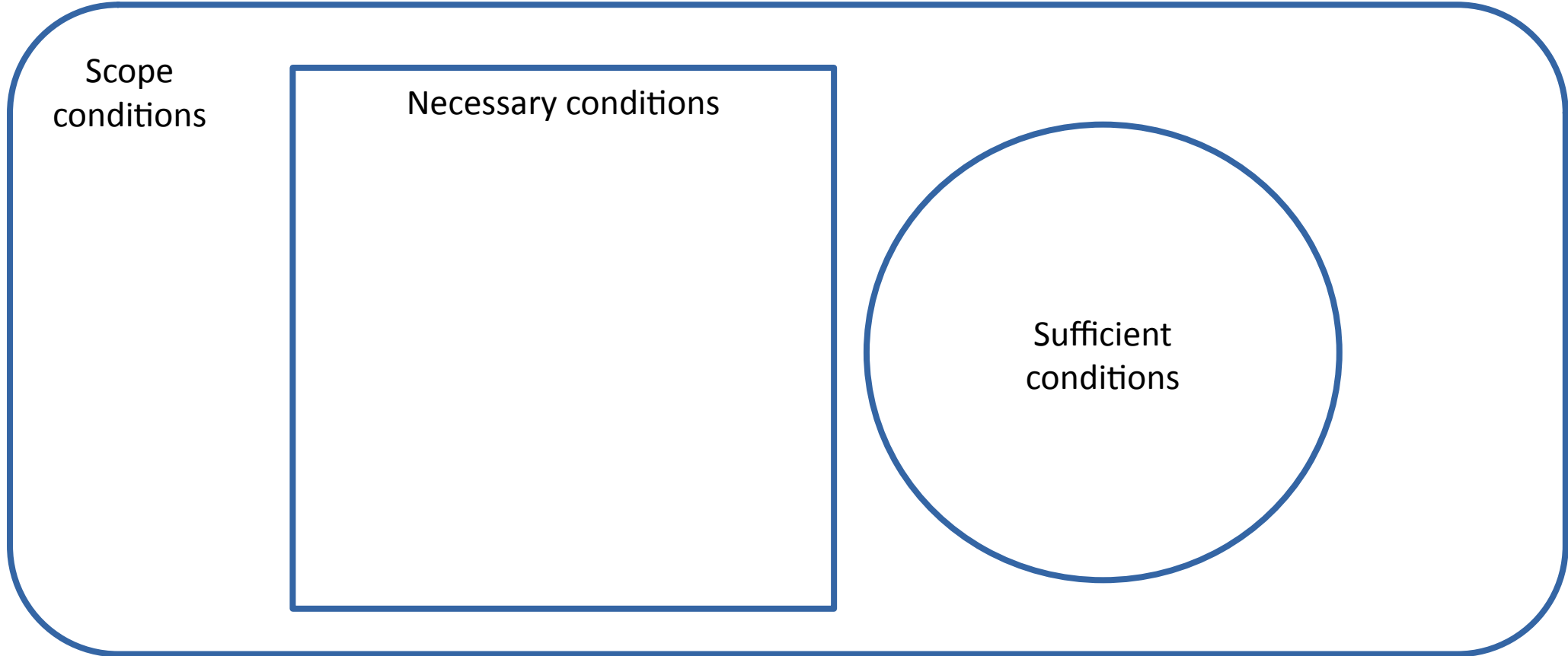
3) **Sufficient conditions**
   - When causal condition is present, outcome will (almost always) occur.
   - Empirical research approaches: small-N comparison, process tracing, QCA.

Relationship among scope conditions, necessary conditions and sufficient conditions

Scope
conditions

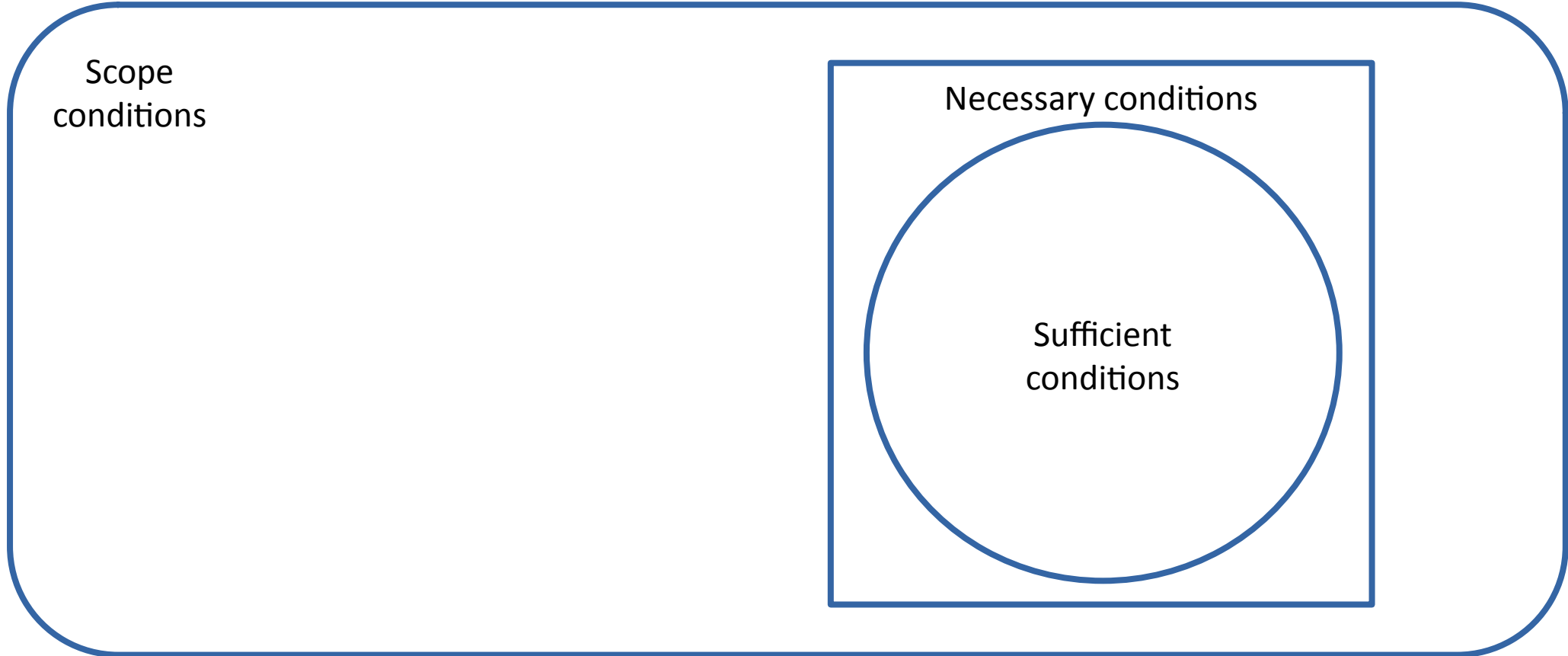Necessary conditions

Sufficient
conditions

Relationship among scope conditions, necessary conditions and sufficient conditions

Scope
conditions

Necessary conditions

Sufficient
conditions

Contributions of NCA-guided QCA

1) **Improves** data set construction by increasing case homogeneity.

2) **Reduces** the likelihood of Type I and II errors by eliminating impossible configurations from the sufficiency analysis.

3) **Resolves outstanding questions** of how to best integrate necessity and sufficiency results

Key insight: Sufficiency analysis should only be conducted on those observations where the NCs are present, because only those observations where the outcome is possible are relevant to explaining the outcome (Mahoney & Goertz 2004).

Upshot: NCA-guided QCA improves established QCA practice by simplifying the sufficiency analysis, clarifying the nature of the relationship between necessary and sufficient conditions, and improving our ability to make causal claims.

Introduction to QCA

QCA offers an analysis of invariance: Certain aspects of cases tend to co-occur:

- "All happy families are alike; each unhappy family is unhappy in its own way" (Tolstoy, *Anna Karenina*)

- Tenured faculty tend to have many publications

- Religious fundamentalists tend to be politically conservative

- HIV causes AIDS;
  Smoking causes lung cancer;
  SARS-CoV-2 causes COVID-19

Invariant Relationships
Certain Aspects of Cases Tend to Co-occur

Set of tenured faculty
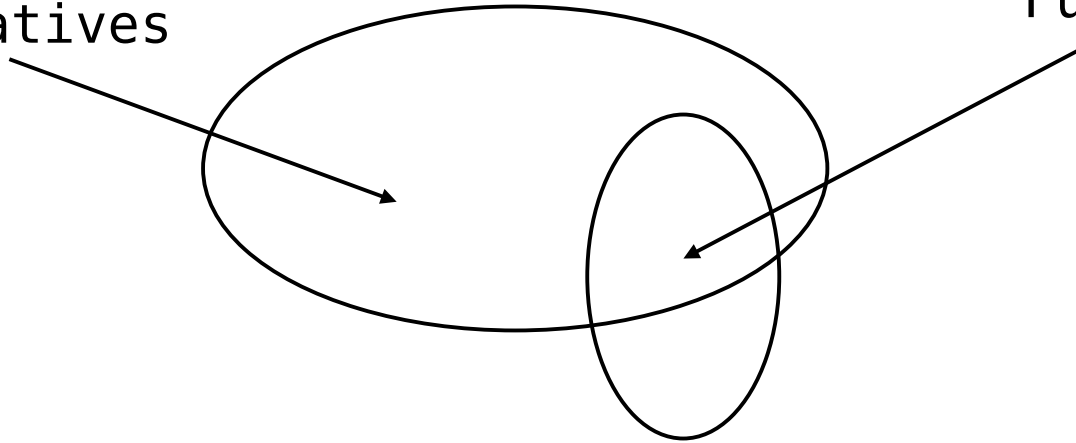
Set of people with many publications
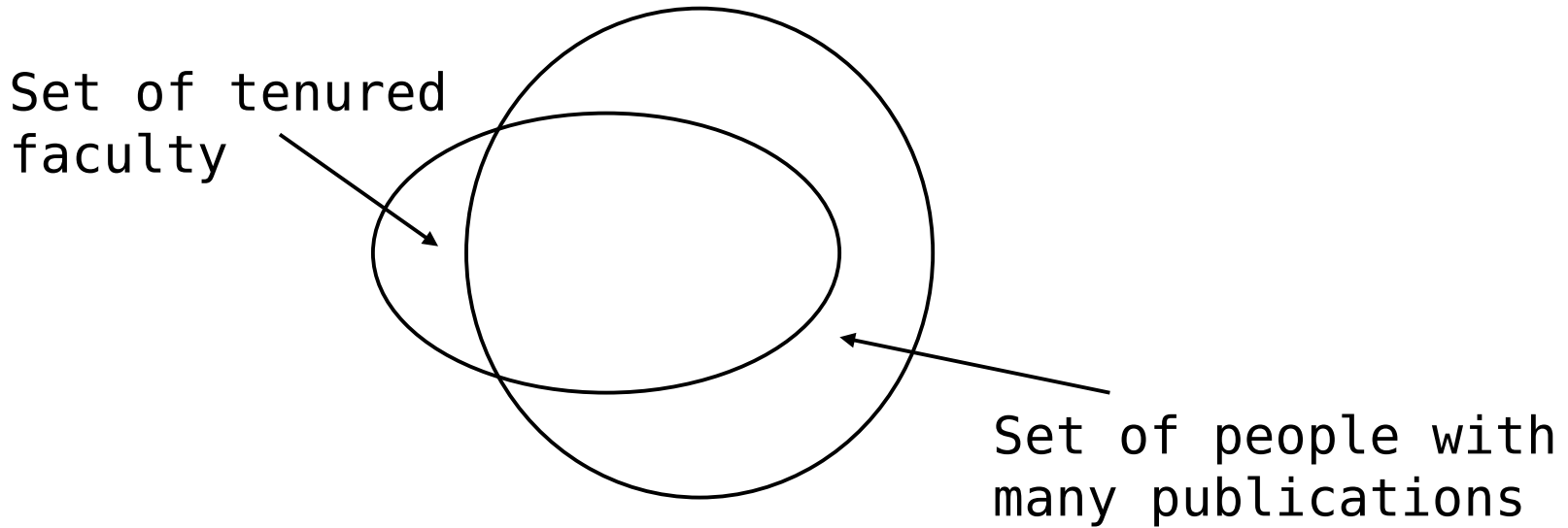
Invariant Relationships
Certain Aspects of Cases Tend to Co-occur
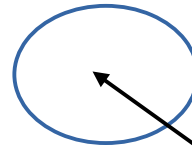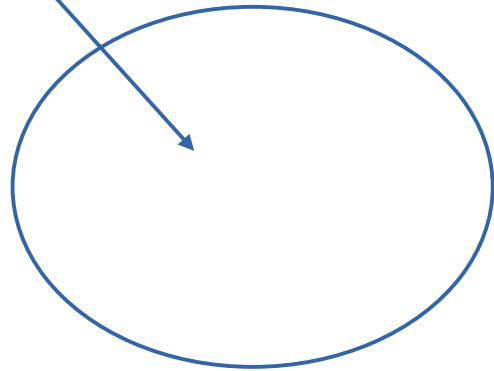
Set of
Political
Conservatives

Set of Religious
Fundamentalists

Invariant Relationships
Certain Aspects of Cases Tend to Co-occur
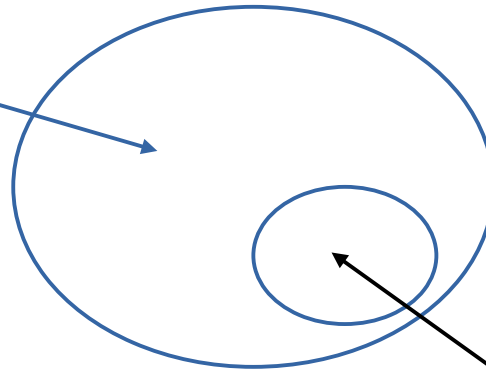
Set of people who are HIV-negative

Set of people with AIDS
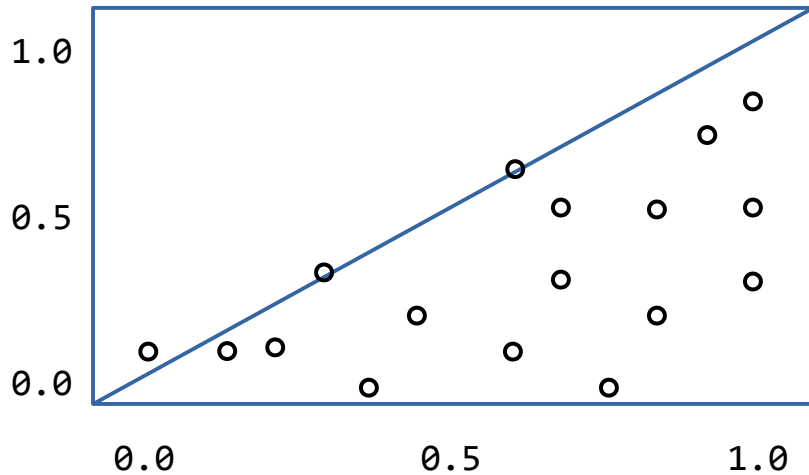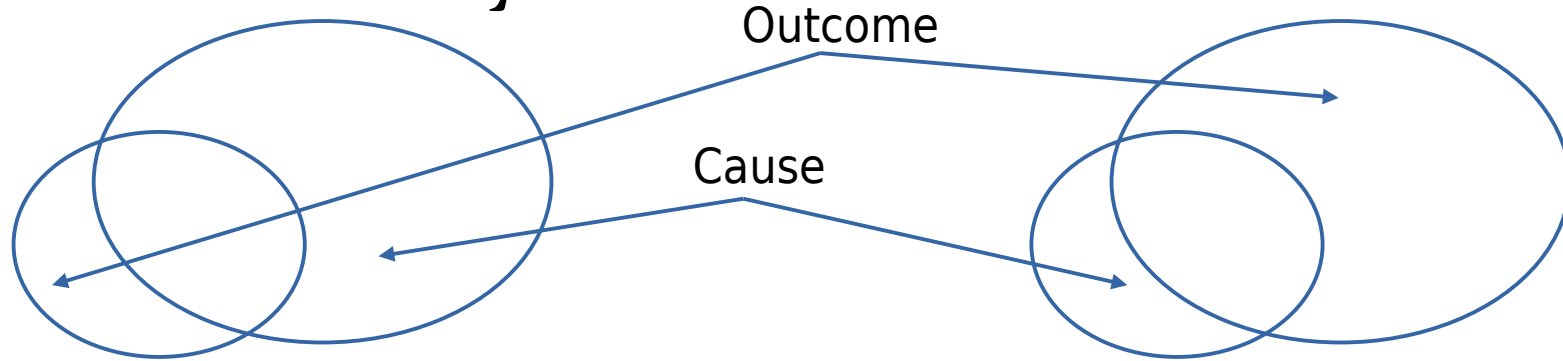
Invariant Relationships
Certain Aspects of Cases Tend to Co-occur

Set of people who are HIV-positive

Set of people with AIDS

QCA uses Boolean Algebra to Identify Necessary and Sufficient Conditions

Outcome

Cause

Subset relationship consistent with *necessity* (X≥Y)
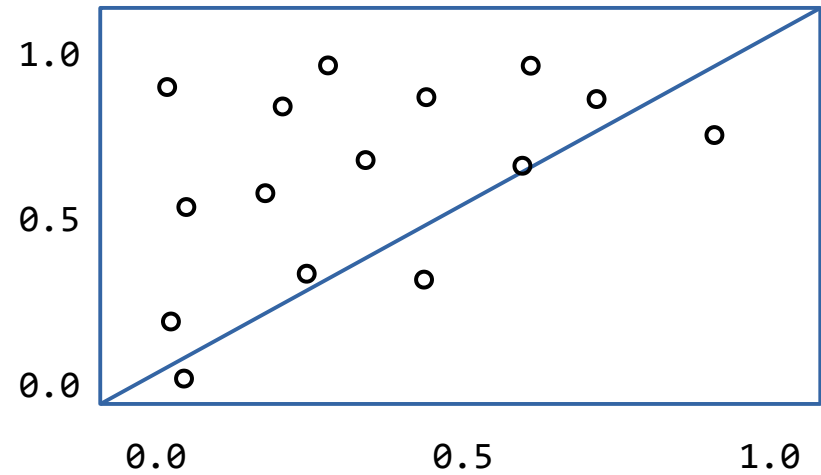
Subset relationship consistent with *sufficiency* (Y≥X)

Distinguishing Features of QCA

- Cross-discipline technique for studying similarities and differences
- Assumption of causal complexity:
  - Conditions may be multifaceted/compound
  - There can be multiple paths to the same outcome (equifinality)
- No degrees-of-freedom restrictions; appropriate for small-, medium-, and large-N analysis
- Encourages retroductive analysis (moving back and forth between theory and data)
  - Uses a malleable analytic frame
  - Must identify, measure, and scale (calibrate) your explanatory conditions and outcome
  - Data set must include both positive and negative outcomes

Why NCA if QCA already has necessity testing?

- When QCA researchers conduct necessity tests (they often don't!), they frequently don't find evidence of necessity.

- This is surprising b/c necessity is the focus of much comparative-historical research (the foundation of QCA).

- The reason: QCA only detects necessity *in kind.* But necessity often exists *in degree:* A minimum amount of X is needed in order to achieve a minimum amount of Y.

- For dichotomous data, NCA and QCA produce the same results. NCA is more sensitive to necessity in continuous data.

- Combining both techniques leverages the strengths of both NCA's necessity analysis and QCA's sufficiency analysis.

NCA and QCA complement one another

- Both conceive of causation in terms of necessity/sufficiency and focus on causes of effects rather than the effects of causes.

- Both identify "invariant" relationships (i.e., that when X is present, Y is or is not present) but neither is deterministic. Both encourage exceptions to the necessity/sufficiency relationship to be explained empirically and theoretically.

- Both establish empirical relationships by examining the distribution of observations across a property space. Observations are treated as holistic entities, rather than mere carriers of variables.

- Both operate independently of sample size and make no assumptions regarding how data was collected. Both may be used to analyze small, medium, and large samples; randomly- and purposively-collected data are equally valid.

- Both fundamentally rely upon meaningful measurement and the application of theoretical and empirical knowledge to establish causal claims.

NCA and QCA complement one another

But: QCA is based on Boolean algebra; NCA, on linear algebra. One can't simply use NCA as a "drop-in" replacement for QCA's necessity testing.

- NCA-guided QCA is a *multi-method* technique: NCA first, then QCA.

- The NCA-guided QCA protocol ensures alignment of the two techniques and that they properly target the same outcome.

Overview of NCA

NCA Plot : Sociability - Sales.performance
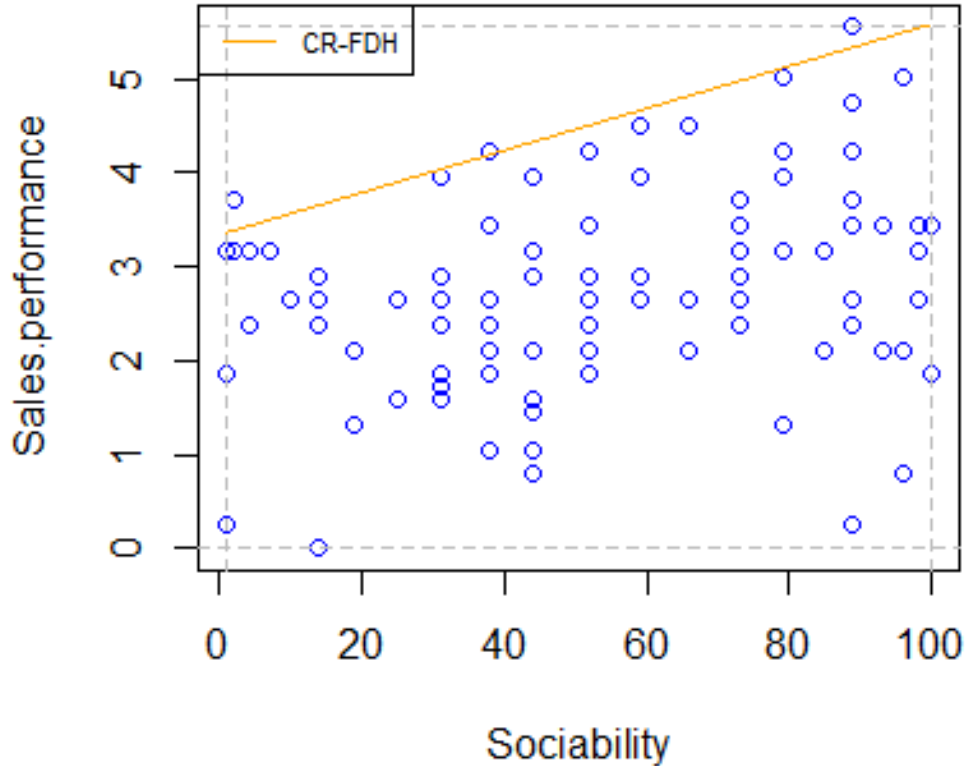
Necessity in kind established via:

- Ceiling line that distinguishes empty space

- Metrics of effect size & statistical significance

- Applicable theory

Overview of NCA

NCA Plot : Sociability - Sales.performance

Necessity in kind established via:

- Ceiling line that distinguishes empty space

- Metrics of effect size & statistical significance
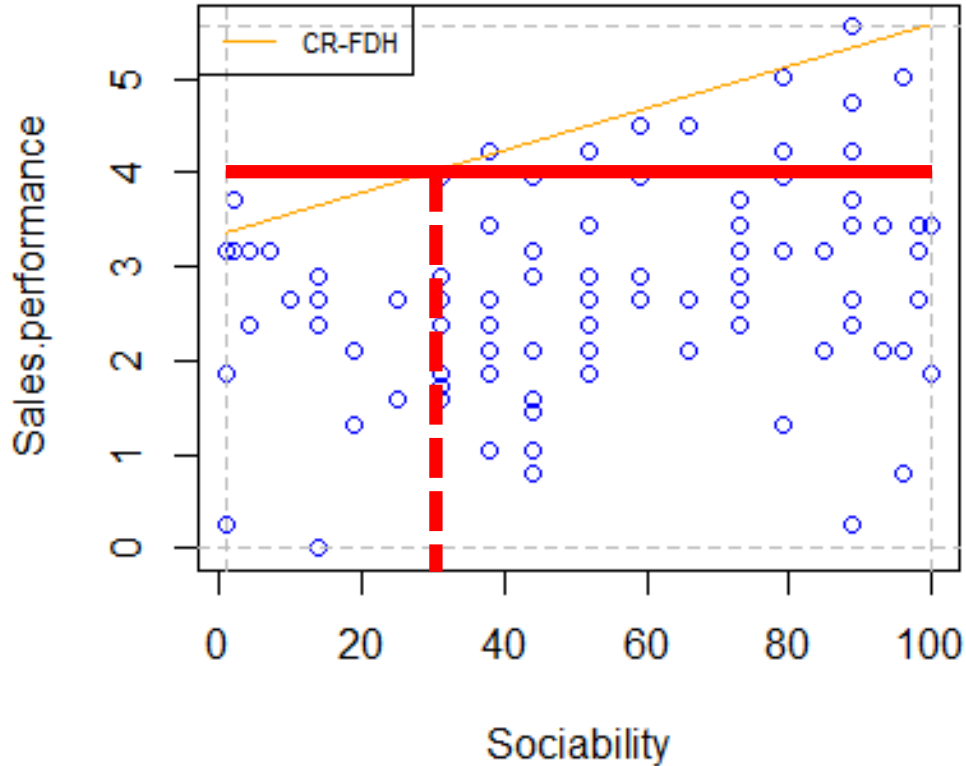
- Applicable theory

Necessity in degree established via:

- Presence of necessity in kind

- XY plot/Bottleneck table

NCA-guided QCA Protocol
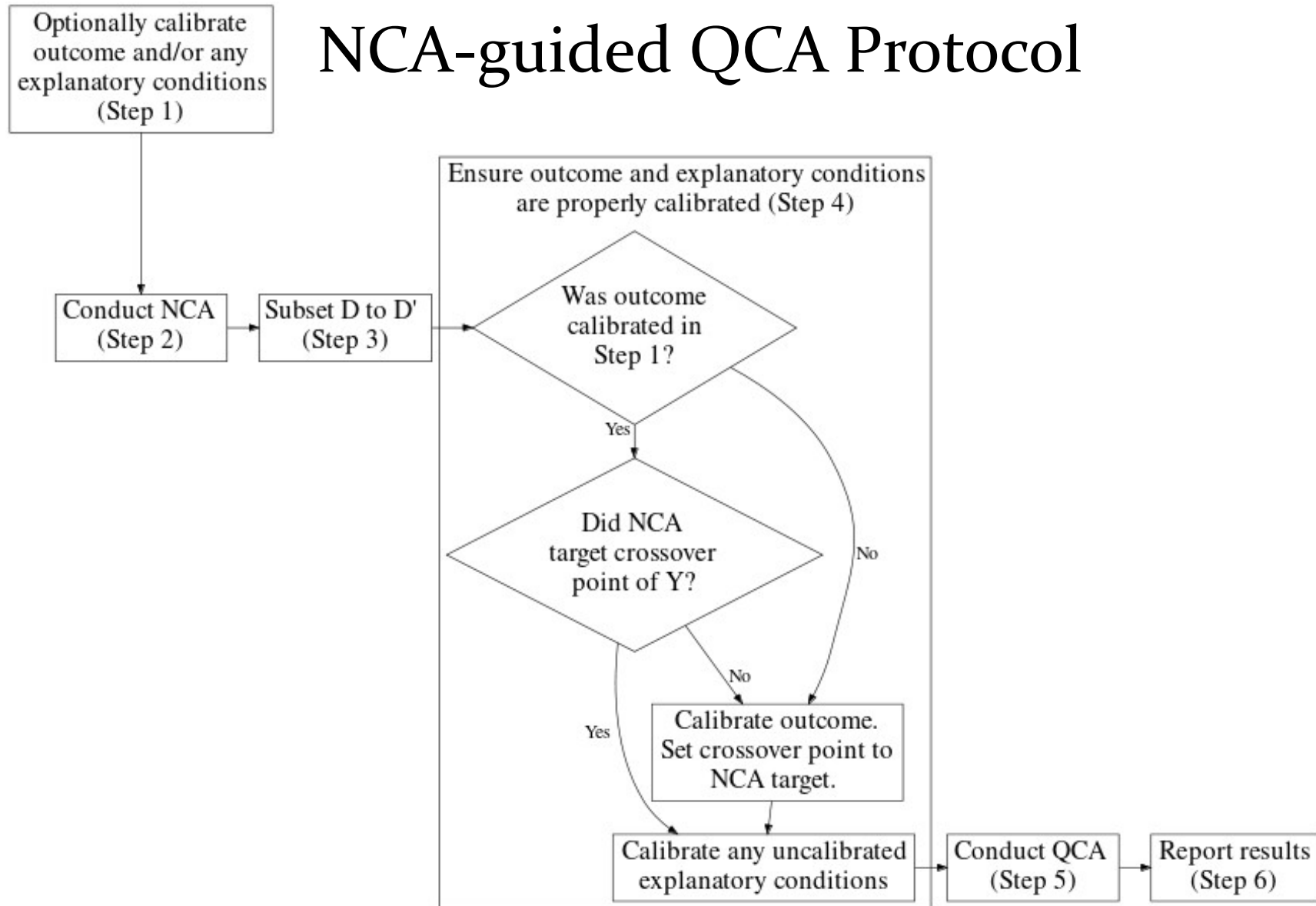
NCA-guided QCA Protocol

Optionally calibrate outcome and/or any explanatory conditions (Step 1)

Pure NCA

Conduct NCA (Step 2)

Subset D to D' (Step 3)

Ensure outcome and explanatory conditions are properly calibrated (Step 4)

Was outcome calibrated in Step 1?

Yes

Did NCA target crossover point of Y?

No

Yes

No

Calibrate outcome. Set crossover point to NCA target.

Calibrate any uncalibrated explanatory conditions

Pure QCA

Conduct QCA (Step 5)

Report results (Step 6)

Value-added is Step 3:

Subset data set to observations where all NCs are present.

Step 4 ensures that NCA and QCA target the same outcome

NCs are embedded into sufficiency results

NCA-guided QCA Protocol

- **Step 2** is a conventional NCA necessity analysis on D to identify the conditions that are necessary to realize the outcome.

- Key is **Step 3**: Subsetting data set D is to those observations that meet/exceed the specified bottleneck thresholds. This effectively dichotomizes the necessary conditions, *calibrating* them as crisp sets.

  - D' now *implicitly* includes all necessary conditions for all observations.

  - Key insight: Lacking the necessary conditions, it is impossible to realize the outcome.

- **Step 4** ensures that NCA and QCA target the same outcome.

- **Step 5** is a conventional QCA sufficiency analysis on D' but ignoring the necessary conditions, which are now constants.

- When reporting results (**Step 6**) include necessary conditions in each sufficiency recipe.

Empirical Examples

Depending upon the nature of one's data, NCA-guided QCA may produce results that are the same as, similar to, or very different from conventional QCA.

1) Emmenegger (2011) "Job Security Regulations in Western Democracies"

2) Martinez, Molina, and Rubinson (in progress) "Future Crossing Intentions of Recently-Repatriated Migrant Women at the Arizona-Sonora Border"

Emmenegger (2011)

- fsQCA to identify determinants of job security regulations (restrictions on hiring and firing) in 19 Western democracies: AU AT BE CA DK FI FR DE IE IT NL NZ NO PT ES SE CH GB US
- 6 explanatory conditions: statist capitalism (S); managed capitalism (C); strong labor movement (L); degree Catholic (R); strong religious parties (P); Few veto points (V)

Original findings

- Necessary condition: **S+C**: "In Western democracies, JSR does not occur under market capitalism but only in the presence of statist or managed capitalist economies."
- Countries inconsistent with necessity: AU CA NL US
- Sufficient conditions: **S*R*v + S*R*P*(C+L) + C*L*P*V**
- Solution describes 12 states (FR IE IT PT; AT BE DE ES; DK NL NO SE)

Reanalysis, using NCA-guided QCA

- For JSR >=0.5, NCA identifies **L (0.29), R (0.2), and P (0.2)** as individually necessary: "For a Western democracy to achieve *at least moderate JSR* requires, the presence of (a) a labor movement, (b) a Catholic political culture, and (c) Christian political parties are required. These may be weakly expressed but none can be entirely absent."
- Drop countries that don't meet required level of L, R, *or* P (AU CA IE NZ CH GB US)
- Conduct QCA on remaining 12 countries, using S, C, & V: **(L*R*P) * ([S*c] + [C*V])**
- Solution describes 8 states (ES FR IT PT; AT BE DE NL)

Emmenegger (2011)

- Here, the conventional QCA necessity testing and the NCA offer different perspectives on what is necessary to achieve JSR. The conventional QCA's distinguishes statist or managed capitalism as necessary. NCA, however, identifies the strength of a country's labor movement and the influence of the Catholic church as crucial.

- The NCA results are more restrictive. The QCA results exclude 4 countries from the possibility of achieving JSR is possible; the NCA results exclude 7. (3 excluded by both.)

- The two techniques are not in competition but suggest different ways of "telling the story" of how JSR is achieved. Emmenegger's discussion is framed around the differences among state capitalist countries, continental European managed capitalist countries, and Nordic managed capitalist countries. Our results tell a story that ignores the continental/Nordic distinction and simply distinguishes between state and managed capitalist countries, all of which share the presence of a Catholic political culture as well as labor and Christian political parties advocating for job security regulations.

- Importantly, the NCA-guided QCA leaves 4 countries unexplained. It would be worth returning to these 4 cases to understand what's going on with them.

Martinez, Molina, & Rubinson (in progress)

- fsQCA to identify determinants of future crossing intentions of women recently deported from U.S. to Mexico

- Outcome: Intends (1.0, 0.7, 0.0)

- 6 explanatory conditions:
  - Prior crossing experience,
  - High social capital,
  - Currently employed,
  - High school educated,
  - Short journey,
  - Encountered bandits

- QCA found no necessary conditions; closest was HiSoCap (ncon=0.85)

Martinez, Molina, & Rubinson (in progress)

Conventional QCA
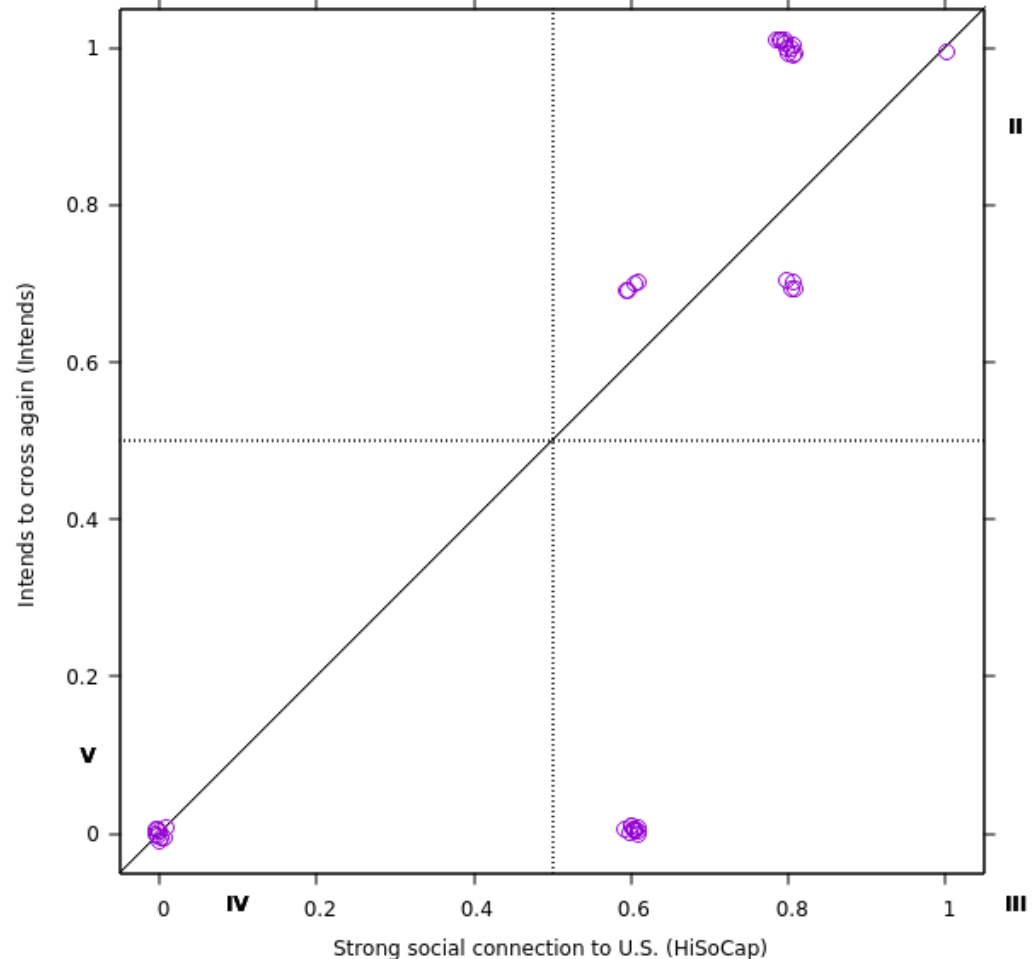
	Recipes			
	1a	1b	1c	1d
High Social Capital	●	⬤	⬤	⬤
Experienced	⊖	●	⬤	⬤
Employed		⊖	⊖	⬤
Educated		⊖		⬤
Short trip	⊖		●	⬤
Bandits	⊖	⊖	⬤	⊖
Consistency	0.87	0.99	1.00	0.87
Raw coverage	0.22	0.24	0.07	0.10
Unique coverage	0.13	0.19	0.07	0.06

Solution consistency: 0.94

Solution coverage: 0.53

● / ●   Core/contributory condition present

⊖ / ⊖   Core/contributory condition absent

Conventional QCA

NCA-guided QCA

	Recipes			
	2a	2b	2c	2d
Moderate Social Capital	◆	◆	◆	◆
Experienced	⊖	●	⬤	⬤
Employed		⊖	⊖	⬤
Educated			⊖	⬤
Short trip	⊖		●	⬤
Bandits	⊖	⊖	⬤	⊖
Consistency	0.88	0.99	0.98	0.87
Raw coverage	0.22	0.26	0.08	0.10
Unique coverage	0.13	0.21	0.08	0.06

Solution consistency: 0.94

Solution coverage: 0.57

◆ / ◇   Presence/absence of condition is necessary

● / ●   Core/contributory condition present

⊖ / ⊖   Core/contributory condition absent

NCA-guided QCA

NCA-guided QCA strengthens causal arguments

I. Increases case homogeneity and implements Mahoney and Goertz's (2004) Possibility Principle:

- "cases where the outcome is *impossible* [are] uninformative and hence irrelevant observations"

- Impossible cases increase the likelihood of committing type I & II errors

- Mahoney and Goertz develop a *theoretically-informed* possibility principle; using NCA makes it *theoretically- and empirically-informed* (the impossible can't happen)

II. Resolves the question of how to integrate necessary conditions into sufficiency results

- Current QCA practices—appending NCs to sufficiency recipes vs selective incorporation of remainders (IS, ESA)—fail in the face of "almost always necessary conditions" (e.g., ncon ≥ 0.9).

- NCA-guided QCA embeds the presence of the NCs into the analysis of sufficient conditions; the NCA results are implicitly part of the QCA results.

Limitations of NCA-guided QCA

1. Need to learn and apply two techniques, NCA and QCA.

2. Step 4 of the protocol, which guarantees the validity of the method, can be confusing; if a mistake is made here, there's nothing in the protocol to alert you.

3. Protocol may encourage mechanistic analysis.

4. If QCA tends to find "too few" necessary conditions, NCA risks finding "too many." The application of theoretical and empirical knowledge is crucial to assess whether the ceiling line is causally meaningful and not simply descriptive.

Recommendations

1. Conducting NCA offers an additional opportunity for getting to know one's cases, which is a critical element of a successful QCA.

2. When D' has few observations, they can be interpreted directly and QCA may not be needed. If all observations in D' possess the outcome, the necessary conditions are also sufficient and there is nothing left for QCA to explain.

3. When working with calibrated data, first examine the XY plots. Observations scattered close to the diagonal may suggest conditions that are consistent with both necessity (according to NCA) and sufficiency (according to QCA). Think carefully about which interpretation makes more sense.

4. Both NCA and QCA require a strong theoretical basis and deep substantive knowledge in order to claim causation. NCA-guided QCA is no different. What NCA-guided QCA offers is the ability to generate and leverage additional empirical knowledge, strengthening one's basis for making causal claims.