Qualitative Comparative Analysis: A Cross-Disciplinary Methodology for Studying Similarities and Differences

Claude Rubinson
University of Houston—Downtown
Houston, TX

qca@grundrisse.org
http://grundrisse.org/qca/

UNAM
October 17–21, 2025

Overview

Day 1: The Logic of QCA

- Introductions and discussion of research projects

Day 2: Three Analytic Components of QCA

- Calibration, Necessity Analysis, Sufficiency Analysis

Day 3: Putting QCA into Practice

- Software for conducting QCA

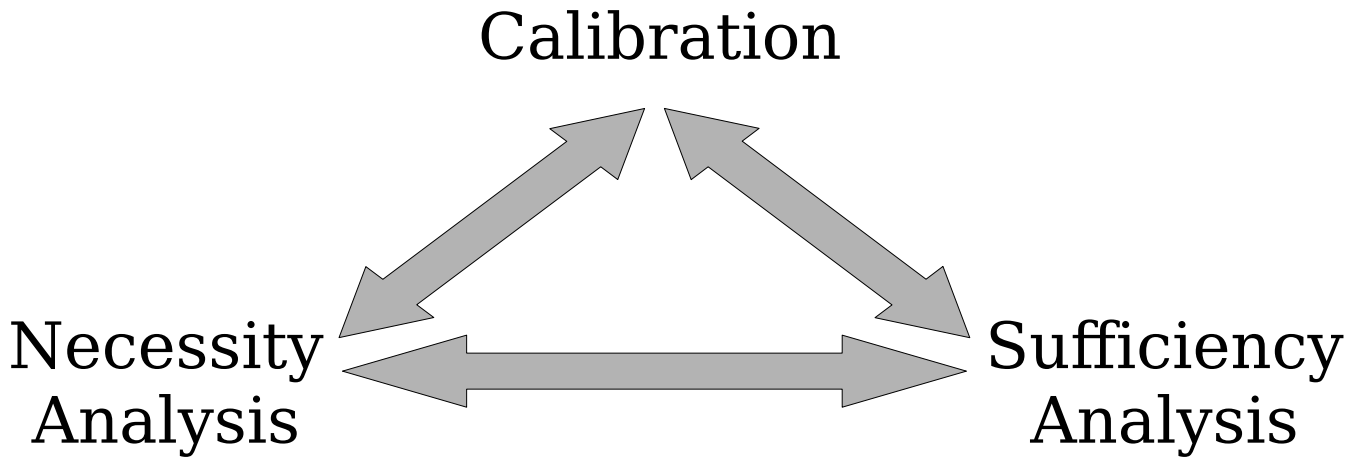
Day 4: Advances in QCA

- Time in QCA, Generalized Analytic Induction

Day 5: Pulling it all Together

- Building robust models, Visualizing and presenting QCA
- Discussion of research projects

Three Analytic Components of QCA



Time in QCA

The relevance of time in QCA

Comparative-historical research, and causal analysis more generally, is fundamentally concerned with time. *Cause precedes effect*. Yet QCA researchers tend to use cross-sectional data sets. How to incorporate time into our analyses?

The relevance of time in QCA

Comparative-historical research, and causal analysis more generally, is fundamentally concerned with time. *Cause precedes effect*. Yet QCA researchers tend to use cross-sectional data sets. How to incorporate time into our analyses?

- First *'Time-in-QCA' (TiQ) International Workshop* organized by Lasse Gerrits and Sofia Pagliarin (2022, Erasmus University Rotterdam). Report available on https://compasss.org

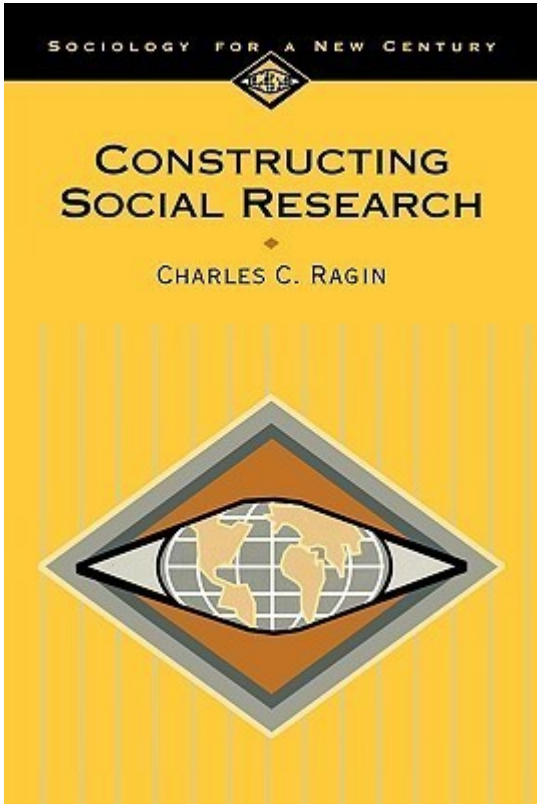
Arguments against integrating time into QCA

1. QCA is fundamentally time-agnostic and assumes *synchronic causality*. There is no logical distinction between conditions and outcome, which in fact occur simultaneously; we only distinguish between them *semantically*.
2. The truth table has no sense of time; equivalently: the truth table assumes that time is constant across observations.
3. Time is already embedded in statements of necessity/sufficency.
4. Time is accounted when “going back to the cases,” conducing case studies, process tracing, etc.

Strategies for integrating time into QCA

- Temporal QCA: Conditions as sequences of events (Caren and Panofsky 2005; Ragin and Strand 2008)
- Embed measure of change into conditions (Ragin 2014; Walsh, et. al. 2018).
- Lagged conditions; define separate conditions for different time periods (Walsh, et. al. 2018).
- Use set coincidence to measure how truth tables change over time (Rubinson and Mueller 2016).
- Trajectory-based QCA: Examine how observations move across truth table rows over time (Pagliarin and Gerrits, 2020).
- Critical junction analysis: Within-case analysis of complex events (Rutten 2024).

Generalized Analytic Induction



Three broad categories of empirical research

- Study of covariation

  "*quantitative analysis*" tends to be **deductive & large-N**

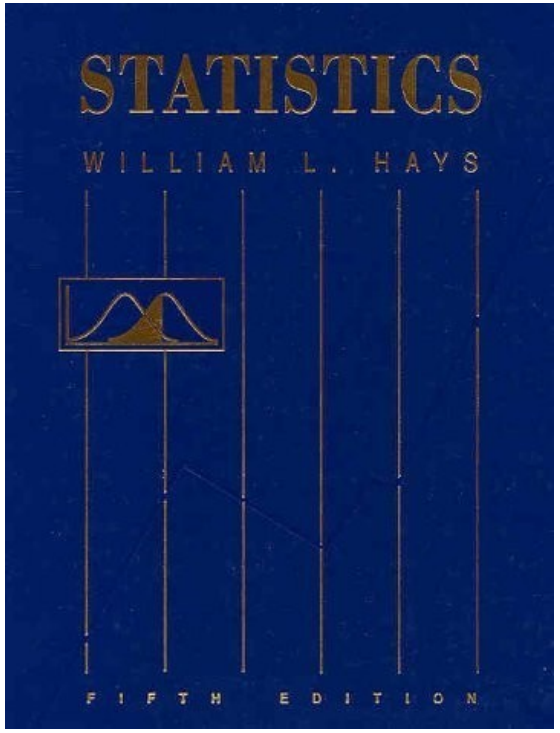
- Study of similarities and differences

  "*comparative analysis*" tends to be **retroductive & medium-N**

- Study of commonalities

  "*qualitative analysis*" tends to be **inductive & small-N**

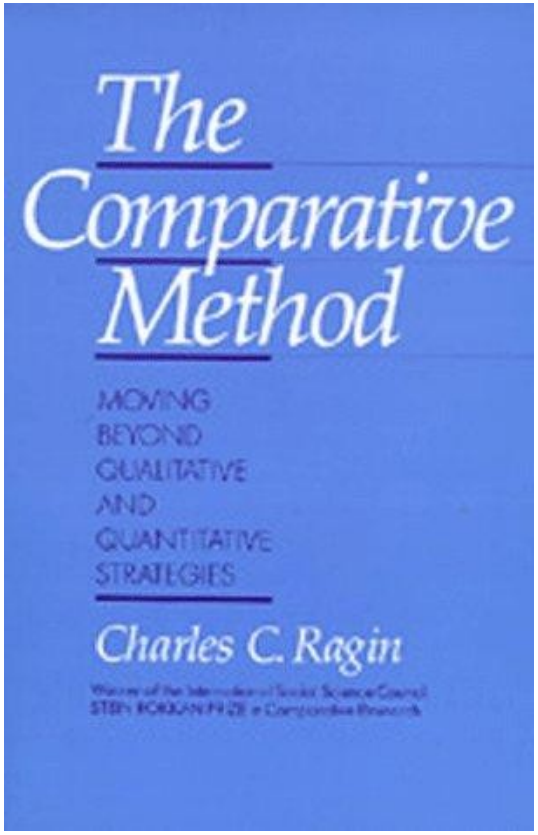
Ragin (1994) *Constructing Social Research: The Unity and Diversity of Method*



Three broad categories of empirical research

- Study of covariation

  "*quantitative analysis*" tends to be **deductive & large-N**



Three broad categories of empirical research

- Study of similarities and differences

  "*comparative analysis*" tends to be **retroductive & medium-N**

- Seeks to explain why cases do or do not exhibit the outcome:

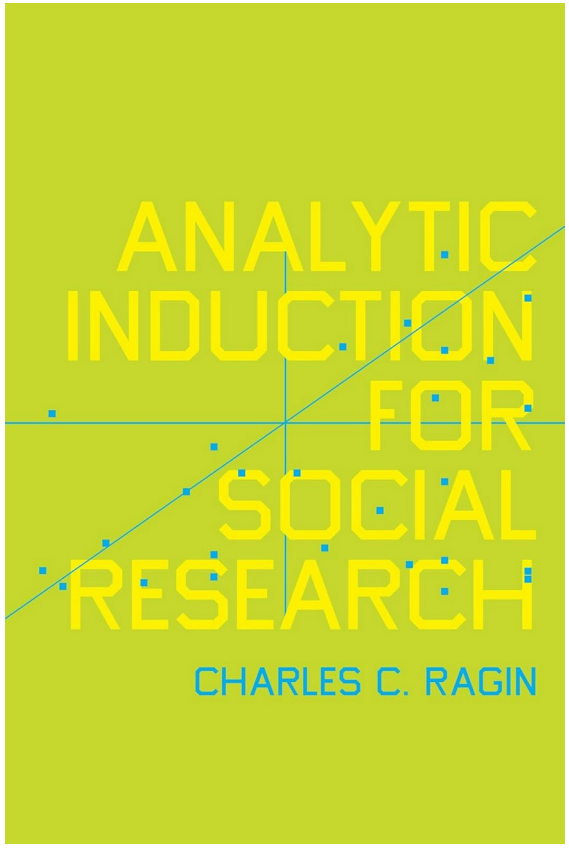
  1) Assumes a set of cases "at risk" for the outcome

  2) Outcome can be yes/no or vary by degree of presence

  3) Set of candidates for the outcome must be carefully defined

  4) There are cases that don't exhibit the outcome, in part or fully

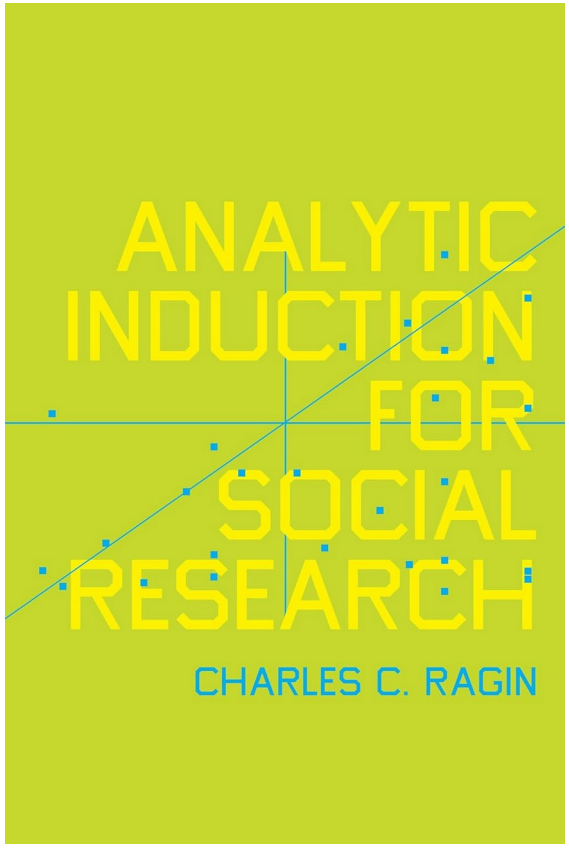
- Comparative analysis is fundamentally case-oriented and set-theoretic; QCA formalizes comparative analysis using Boolean algebra (the algebra of sets)



Three broad categories of empirical research

- Study of commonalities

  "*qualitative analysis*" tends to be **inductive & small-N**

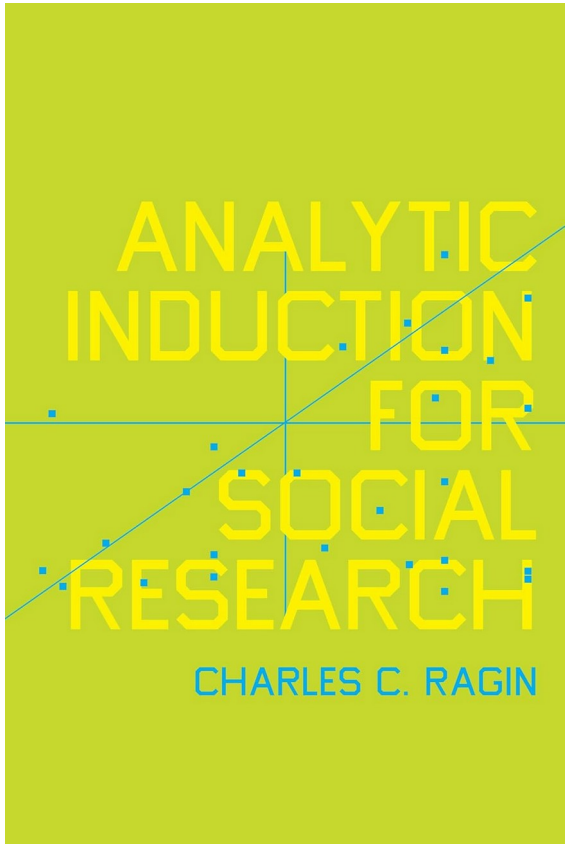


Three broad categories of empirical research

- Study of commonalities

  "*qualitative analysis*" tends to be **inductive & small-N**

- Focus on: interpreting significance, giving voice, developing new theory, understanding "how" an outcome occurs.



Three broad categories of empirical research

- Study of commonalities

  "*qualitative analysis*" tends to be **inductive & small-N**

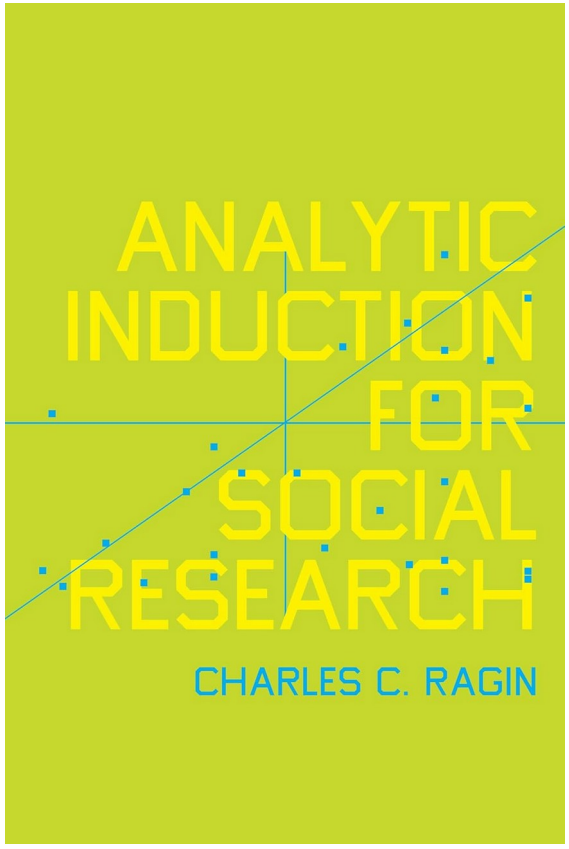
- Focus on: interpreting significance, giving voice, developing new theory, understanding "how" an outcome occurs.

- In contrast to comparative research:

  1) Outcome is (more or less) the same for every case, a constant.

  2) The set of cases with the outcome is the set of relevant cases.

  3) There are no "negative cases." Cases lacking the outcome are not relevant to understanding "how."



Three broad categories of empirical research

- Study of commonalities

  "*qualitative analysis*" tends to be **inductive & small-N**

- Example: Becker (1953) "Becoming a Marihuana User

BECOMING A MARIHUANA USER*

HOWARD S. BECKER

ABSTRACT

An individual will be able to use marihuana for pleasure only when he (1) learns to smoke it in a way that will produce real effects; (2) learns to recognize the effects and connect them with drug use; and (3) learns to enjoy the sensations he perceives. This proposition, based on an analysis of fifty interviews with marihuana users, calls into question theories which ascribe behavior to antecedent predispositions and suggests the utility of explaining behavior in terms of the emergence of motives and dispositions in the course of experience.

Analytic Induction

Negative cases:

- From AI's perspective, "negative cases" *per se* don't exist; they are positive cases of something else, deserving of separate treatment.

Explaining the outcome:

- AI identifies shared *antecedent conditions* across a range of *cases* with the same outcome, focusing especially on any reconceptualization that is prompted by disconfirming cases.
- The researcher explains a *constant* (the shared outcome) with *another constant* or, more often, *a set of constants* (shared antecedent conditions).


WHOA. NO VARIABLES?


EXCELLENT!
Metro Goldwyn Mayer

Analytic Induction's Checkered History

- In various incarnations, AI was a popular technique in U.S. sociology in the early decades of empirical research:
  - Znaniecki (1934) *The Method of Sociology*
  - Angell (1936) *The Family Encounters the Depression*
  - Lindesmith (1947) *Opiate Addiction*
  - Cressy (1950) *Criminal Violation of Financial Trust*
  - Becker (1953) *Becoming a Marijuana User*
- Challenged by W.S. Robinson's (1951) "The Logical Structure of Analytic Induction" as being useless for prediction. In combination with other critiques, AI was eclipsed by "enumerative induction" (i.e., analysis of covariation).

QCA and AI

	Cause absent	Cause present
Outcome present		
Outcome absent		

QCA and AI

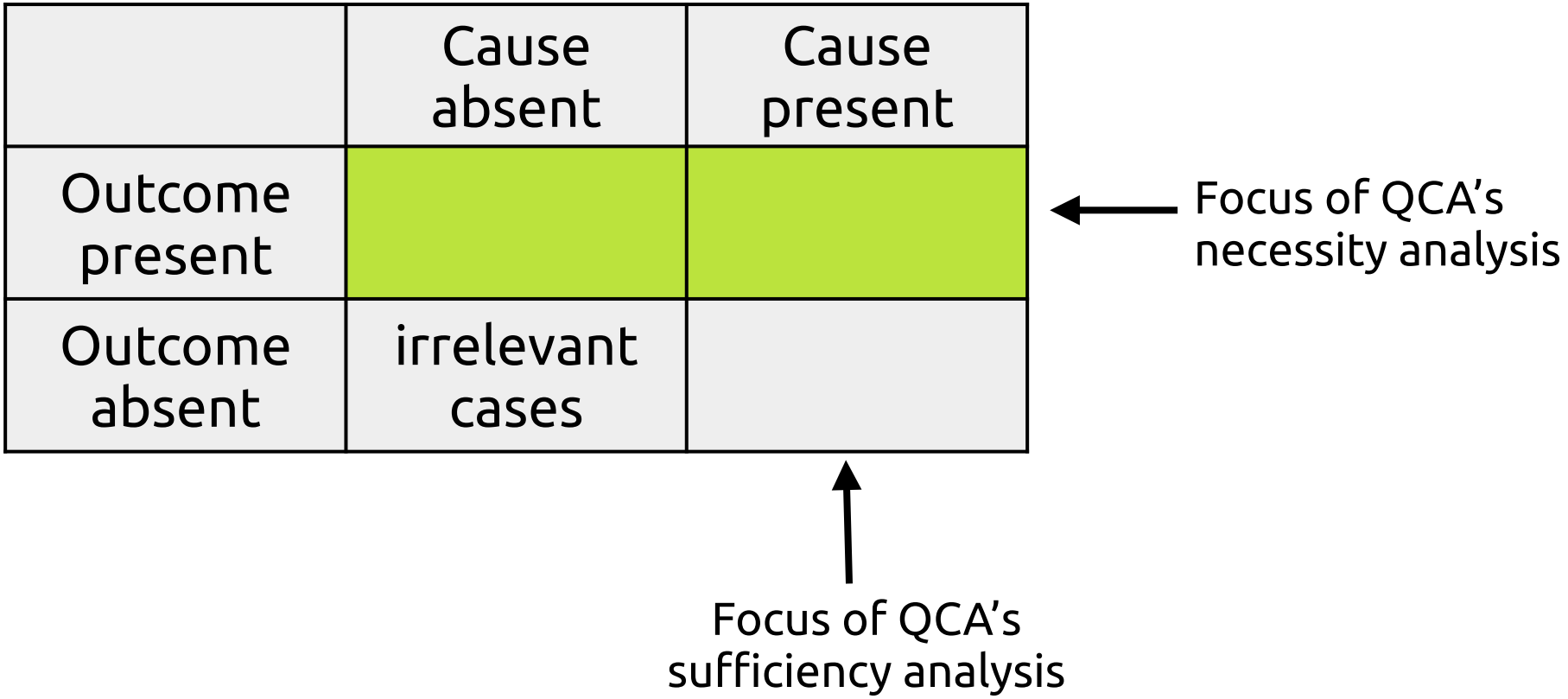
	Cause absent	Cause present
Outcome present		
Outcome absent	irrelevant cases	

QCA and AI

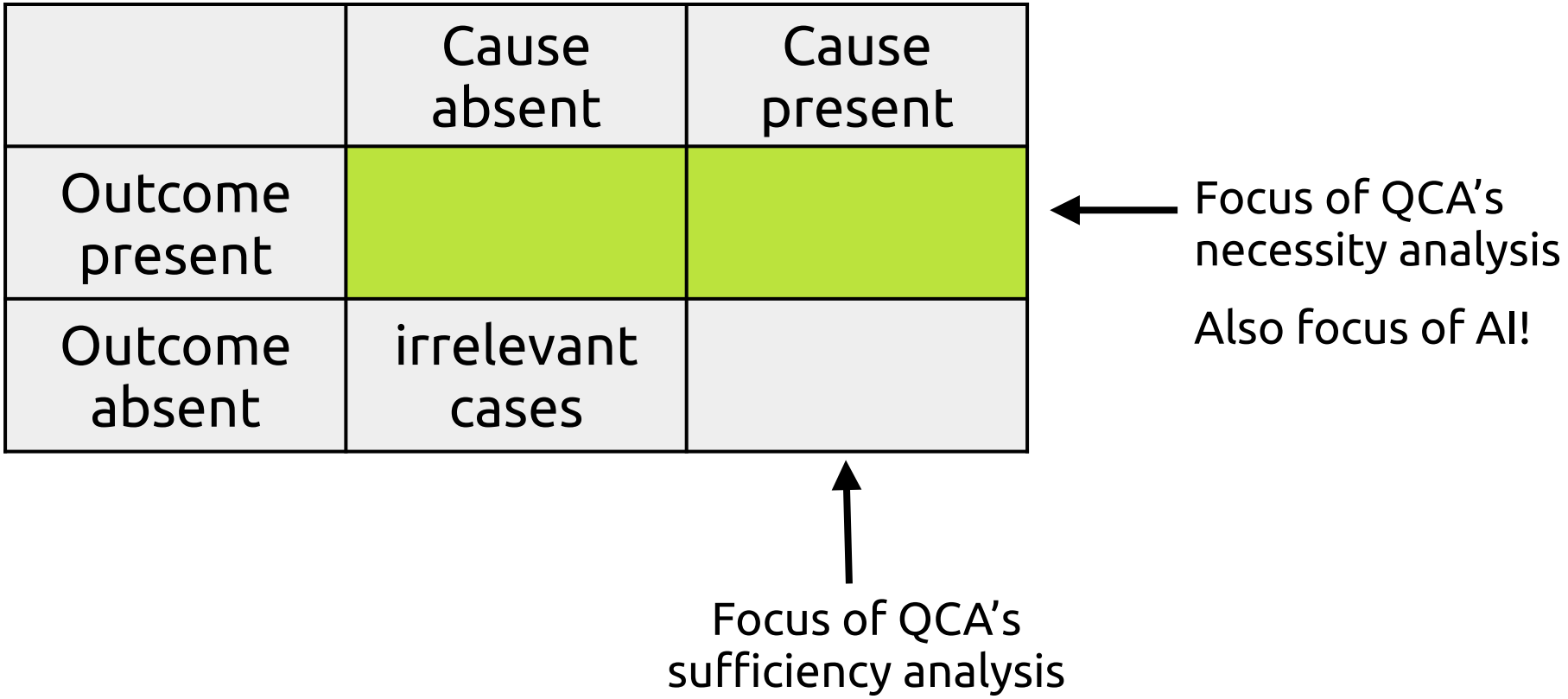
	Cause absent	Cause present
Outcome present		
Outcome absent	irrelevant cases	

Focus of QCA's sufficiency analysis

QCA and AI


Cause absent
Cause present
Outcome present
Outcome absent
irrelevant cases
Focus of QCA's necessity analysis
Focus of QCA's sufficiency analysis

QCA and AI


Cause absent
Cause present
Outcome present
Outcome absent
irrelevant cases
Focus of QCA's necessity analysis
Also focus of AI!
Focus of QCA's sufficiency analysis

QCA and AI

	Cause absent	Cause present
Outcome present		
Outcome absent	irrelevant cases	

← Focus of AI

- AI invites us to take the concept of necessity more seriously by understanding necessary conditions as enabling forces and not merely barriers (cf. Goertz, Dul).
- But Ragin uses the term *antecedent condition*, which does not imply a causal relationship.

QCA and AI

	Cause absent	Cause present
Outcome present		
Outcome absent	irrelevant cases	

← Focus of AI

Like QCA, AI is:

- a descriptive technique, not inferential. Causation is established by identifying mechanism(s), using substantive and theoretical knowledge.
- sensitive to causal complexity, conceiving of conditions as configurations and recognizing equifinality (but see discussion of PoS formulas).

QCA and AI

	Cause absent	Cause present
Outcome present		
Outcome absent	What happened instead?	

← Focus of AI

Unlike QCA, AI:

- considers negative cases to be positive cases of something else, also worthy of investigation.
- does not homogenize negative cases. Usually, these so-called negative cases represent several different “happenings.”

Conducting AI

1. Specify the outcome to be explained. Often, the outcome is a happening.
2. Collect evidence on observations that clearly exhibit the outcome.
3. Identify causally-relevant antecedent conditions for the outcome shared by these initial observations. Formulate a working hypothesis about these observed commonalities.
4. Collect evidence on additional instances of the outcome. Try to identify instances of the outcome that challenge your working hypothesis.
5. When you find observations that challenge your working hypothesis, you will need to reformulate your scope, antecedent conditions and/or outcome. When the outcome is reformulated, its scope is typically narrowed so that disconfirming cases are excluded. When antecedent conditions are reformulated, the causal argument is usually altered to embrace the disconfirming cases in some way.
6. Repeat steps 4 and 5 until the evidence derived from additional instances no longer prompts reformulations of the hypothesis or its empirical scope. The goal is to achieve operational closure between your analytic frame and images.

Conducting AI

AI is fiercely attentive to disconfirming observations that exhibit the outcome and not the cause. Goal is to empty the upper-left cell.

Four strategies, all of which involve embracing case-oriented, retroductive analysis:

1. Narrow the domain
2. Identify substitutable conditions
3. Restrict the outcome
4. Expand the cause

(From the variable-oriented, deductive perspective, these strategies are either neutral or regressive.)

Reconciling Disconfirming Observations
Strategy 1: Narrow the Domain

Panel A. Initial definition of relevant observations (all democracies)

	Elites cohesive	Elites fractionalized
Unstable democracy	5 (25%)	25 (62.5%)
Stable democracy	15 (75%)	15 (37.5%)

N = 60

Panel B. Revised definition of relevant observations (only parliamentary democracies)

	Elites cohesive	Elites fractionalized
Unstable parliamentary democracy	0 (0%)	22 (63%)
Stable parliamentary democracy	10 (100%)	13 (37%)

N = 45

Reconciling Disconfirming Observations
Strategy 2: Identify Substitutable Conditions

Panel A. Initial results

	Elites cohesive	Elites fractionalized
Unstable democracy	5 (25%)	25 (62.5%)
Stable democracy	15 (75%)	15 (37.5%)

N = 60

Panel B. Resolution via a more inclusive argument

	Elites cohesive and no dominant ethnic minority	Elites fractionalized or dominant ethnic minority
Unstable democracy	0 (0%)	30 (62.5%)
Stable democracy	12 (100%)	18 (37.5%)

N = 60

Reconciling Disconfirming Observations
Strategy 3: Restrict the Outcome

Panel A. Initial conceptualization of outcome

	Elites cohesive	Elites fractionalized
Unstable democracy	5 (25%)	25 (62.5%)
Stable democracy	15 (75%)	15 (37.5%)

N = 60

Panel B. Reconceptualized outcome

	Elites cohesive	Elites fractionalized
Constitutional crisis	0 (0%)	20 (50%)
No constitutional crisis	20 (100%)	20 (50%)

N = 60

Reconciling Disconfirming Observations
Strategy 4: Expand the Cause

Panel A. Initial conceptualization of antecedent condition

	Elites cohesive	Elites fractionalized
Unstable democracy	5 (25%)	25 (62.5%)
Stable democracy	15 (75%)	15 (37.5%)

N = 60

Panel B. Reconceptualized antecedent condition

	Elites not conflictual	Elites discord
Unstable democracy	0 (0%)	30 (60%)
Stable democracy	10 (100%)	20 (40%)

N = 60

Using fs/QCA to Conduct AI

Obs	Daily exercise (exer)	Feeling of separateness (feel)	Workout rituals (ritl)	Associate w/ athletes (assoc)	Separate food (food)	Maintains commitment (commit)
1	1	1	1	1	1	1
2	1	0	1	1	1	1
3	1	1	1	0	0	1
4	1	1	1	0	1	1
5	1	1	1	0	0	1
6	1	1	0	1	1	1
7	1	0	0	1	1	1
8	1	0	1	1	0	1
9	1	0	1	1	1	1
10	1	1	0	0	1	1
11	1	1	1	0	1	1
...	— Rows 12–19 excluded —					
20	1	0	1	1	0	1

Using fs/QCA to Conduct AI

Initial truth table

Daily exercise (exer)	Feeling of separateness (feel)	Workout rituals (ritl)	Associate w/ athletes (assoc)	Separate food (food)	N	Maintains commitment (commit)
1	1	1	0	1	4	1
1	1	0	1	1	4	1
1	0	1	1	1	4	1
1	1	1	0	0	2	1
1	0	1	1	0	2	1
1	1	1	1	1	2	1
1	1	0	0	1	1	1
1	0	0	1	1	1	1

Using fs/QCA to Conduct AI

Revised truth table

Daily exercise (exer)	Feeling of separateness (feel)	Workout rituals (ritl)	Associate w/ athletes (assoc)	Separate food (food)	N	Maintains commitment (commit)
1	1	1	–	1	4	1
1	1	–	1	1	4	1
1	–	1	1	1	4	1
1	1	1	–	–	2	1
1	–	1	1	–	2	1
1	1	1	1	1	2	1
1	1	–	–	1	1	1
1	–	–	1	1	1	1

- Based on substantive and theoretical knowledge, the researcher *reinterprets* the presence (absence) of each condition as contributing to the realization of the outcome, or as irrelevant.
- A dash indicates which, if any, state is irrelevant.

Using fs/QCA to Conduct AI

Revised truth table

Daily exercise (exer)	Feeling of separateness (feel)	Workout rituals (ritl)	Associate w/ athletes (assoc)	Separate food (food)	N	Maintains commitment (commit)
1	1	1	–	1	4	1
1	1	–	1	1	4	1
1	–	1	1	1	4	1
1	1	1	–	–	2	1
1	–	1	1	–	2	1
1	1	1	1	1	2	1
1	1	–	–	1	1	1
1	–	–	1	1	1	1

- Truth table minimization (remainders are treated as ‘false’):

  exer•feel•ritl + exer•ritl•assoc + exer•feel•food + exer•assoc•food

- Four recipes for sustained commitment, with only one shared antecedent condition.

Dude, where are the constants?

Converting “sums-of-products” to “products-of-sums”

- Multiple recipes (equifinality) indicates multiple pathways to realizing the outcome, not a single set of shared antecedent conditions. Where’s the constant?
- Strategy 2 (“Identify substitutable conditions”) and Strategy 4 (“Expand the cause”) can be implemented using Boolean algebra.
- When multiple conditions serve the function, they can be joined using logical OR

Example:

- 13 athletes exhibit “associates primarily with other athletes (assoc)”
- 7 remaining athletes exhibit “feels separate from or superior to non-athletes (feel)”
- Can these conditions be considered substitutable for one another? If so, we can create a macro-condition of “identifies boundary between athletes and non-athletes,” defined algebraically as “assoc + feel.”
- Are “workout rituals (ritl)” and “separate food (food)” substitutable as the macro-condition “everyday rituals that reinforce the athlete identity”?

Converting “sums-of-products” to “products-of-sums”

The initial four-recipe solution is presented in “sums-of-products” form, in which each pathway is separated by logical OR:

exer•feel•ritl + exer•ritl•assoc + exer•feel•food + exer•assoc•food

Via logical substitutions, we recognize that committed athletes share three antecedent conditions, not just one (devotion to exercise):

1. Devotion to a daily exercise regime (exer),
2. Construction of a boundary between athletes and non-athletes (feel + assoc), and
3. Everyday practices that reinforce one’s identity as an athlete (food + ritl)

We can therefore reformulate the solution into a “product-of-sums” form, with pathways separated by logical AND. Result is a set of shared antecedent conditions (i.e., a constant):

exer • (feel + assoc) • (food + ritl)

SOP expressions can be converted to POS form using fs/QCA software.