THE THREE PHASES OF COMPARATIVE RESEARCH USING QCA

1. Learn as much as you can about the cases. If possible, construct a narrative for each case. Use case comparisons to refine and systematize your understandings. This process culminates in the creation of a truth table, which both represents what you've learned and disciplines your representations of the cases.

2. The Analytic Moment: Analyze the evidence using QCA. Actually, preliminary results usually send you back to phase 1.

3. Take the results back to the cases. The real "test" of the results is how useful they are. Do they help you understand the cases better? Do the different paths (causal combinations) make sense at the case level? Do the results place the cases in a new light, perhaps revealing something that would not have been evident before the analysis (phase 2).

SIMPLIFIED REPRESENTATION OF SKOCPOL'S *STATES AND SOCIAL REVOLUTIONS*

	Conditions for Political Crisis				**Conditions for Peasant Insurrection**		**Social Revolution**
	ba	*ll*	*ae*	*ip*	*ac*	*lp*	*sr*
Positive Cases							
France	1	1	0	1	1	1	1
Russia	1	0	0	1	1	1	1
China	1	1	0	1	1	0	1
Negative Cases							
Prussia/Gm. 1807	1	0	1	1	0	0	0
Prussia/Gm. 1848	1	0	1	0	0	0	0
Japan	1	0	0	1	0	0	0
England	0	1	1	0	0	0	0

ba = Bureaucratic, absolutist state
ll = Landed commercial class with political leverage
ae = Agrarian economy: transition to capitalist agriculture
ip = International pressure (at least moderate)
ac = Agrarian class structure: agrarian sector composed mainly of peasant smallholders
lp = Local politics: peasant communities are relatively autonomous
sr = Social revolution

SIMPLE EXAMPLE OF QCA USING HYPOTHETICAL DATA

A. Truth Table:

C	*L*	*H*	*G*	*U*	*N of Cases*
0	0	0	0	0	4
0	0	0	1	0	3
0	0	1	0	0	6
0	0	1	1	1	2
0	1	0	0	1	3
0	1	0	1	1	4
0	1	1	0	0	3
0	1	1	1	1	5
1	0	0	0	0	7
1	0	0	1	0	8
1	0	1	0	0	1
1	0	1	1	1	7
1	1	0	0	1	3
1	1	0	1	1	2
1	1	1	0	0	7
1	1	1	1	1	6

C = Corporatist wage negotiations
L = At least five years of rule by Left or Center-Left parties
H = Ethnic-cultural homogeneity
G = At least ten years of sustained economic growth
U = Adoption of universal pension system

B. Table simplified through row-wise comparisons (positive outcomes only)

-10- (or L*h: Left rule combined with ethnic diversity)[a]
-1-1 (or L*G: Left rule combined with economic growth)
--11 (or H*G: ethnic homogeneity combined with economic growth)

Dashes indicate that a condition has been eliminated (found to be irrelevant)

C. Finding Redundant Terms:

		Terms to be Covered (Rows with Outcome = 1)							
		0100	1100	0101	1101	0011	1011	0111	1111
Simplified Terms (from B)	-10-	x	x	x	x				
	-1-1			x	x			x	x
	--11					x	x	x	x

D. Final Results (logically minimal):

U = L*h + H*G

Lower-case letters indicate condition must be absent.
Upper-case letters indicate that condition must be present.
Multiplication (*) indicates combined conditions (logical *and*).
Addition (+) indicates alternate combinations (logical *or*).

SIMPLE EXAMPLE OF LIMITED DIVERSITY

Strong Unions (U)	Strong Left Parties (L)	Generous Welfare State (G)	N of Cases
Yes	Yes	Yes	6
Yes	No	No	8
No	No	No	5
No	Yes	????	0

Is it strong left parties (L) that causes generous welfare states (G) or is it the combination of strong unions and strong left parties (U*L) that causes generous welfare states (G)?

Limited Diversity in a Truth Table with Four Causal Conditions

A	B	C	D	Y
no	no	no	no	no
no	no	no	yes	?
no	no	yes	no	?
no	no	yes	yes	?
no	yes	no	no	no
no	yes	no	yes	no
no	yes	yes	no	?
no	yes	yes	yes	no
yes	no	no	no	?
yes	no	no	yes	?
yes	no	yes	no	?
yes	no	yes	yes	?
yes	yes	no	no	yes
yes	yes	no	yes	yes
yes	yes	yes	no	?
yes	yes	yes	yes	?

Territorially Based Linguistic Minorities in Western Europe

Austria:	Slovenes
	Magyars
	Croats
Belgium:	Flemings
	Walloons
	Germans
Great Britain:	Gaels (Scotland)
	Gaels (Isle of Man)
	Gaels (N. Ireland)
	Welsh
	Channel Islanders
Denmark:	Germans
	Faroe Islanders
	Greenlanders
Finland:	Swedes (mainland)
	Swedes (Aaland)
	Lapps
France:	Occitans
	Corsicans
	Alsatians
	Flemings
	Bretons
West Germany:	Danes
	North Frisians
Ireland:	Gaels
Italy:	Friulians
	Ladins
	Valdotians
	South Tyroleans
	Slovenes
	Sards
	Greeks
	Albanians
	Occitans
Netherlands:	West Frisians
Norway:	Lapps
Spain:	Catalans
	Basques
	Galicians
Sweden:	Lapps
	Finns
Switzerland:	Jurassians

SUMMARY PRESENTATION OF PREDICTIONS OF THREE THEORIES OF ETHNIC POLITICAL MOBILIZATION

	Guiding Perspective		
Characteristic	*Developmental*	*Reactive*	*Competitive*
Size of Subnation (S)	(1)[a]	(1)[a]	1
Linguistic Base (L)	1	0	(1)[a]
Relative Wealth (W)	(0)[a]	0	1
Economic Status (G)	0	?[b]	?[b]

[a] Predictions in parentheses are only weakly indicated by the theories.
[b] Question marks indicate that no clear prediction is made.

DATA ON TERRITORIALLY BASED LINGUISTIC MINORITIES

Minority	S	L	W	G	E
Lapps, Finland	0	0	0	0	0
Finns, Sweden (Torne Valley)	0	0	0	0	0
Lapps, Sweden	0	0	0	0	0
Lapps, Norway	0	0	0	0	0
Albanians, Italy	0	0	0	0	0
Greeks, Italy	0	0	0	0	0
North Frisians, Germany	0	0	0	1	1
Danes, Germany	0	0	0	1	1
Basques, France	0	0	0	1	1
Ladins, Italy	0	0	1	0	0
Magyars, Austria	0	1	0	0	0
Croats, Austria	0	1	0	0	0
Slovenes, Austria	0	1	0	0	1
Greenlanders, Denmark	0	1	0	0	1
Aalanders, Finland	0	1	1	0	2
Slovenes, Italy	0	1	1	1	1
Valdotians, Italy	0	1	1	1	2
Sards, Italy	1	0	0	0	1
Galicians, Spain	1	0	0	0	1
West Frisians, Netherlands	1	0	0	1	1
Catalans, France	1	0	0	1	1
Occitans, France	1	0	0	1	1
Welsh, Great Britain	1	0	0	1	2
Bretons, France	1	0	0	1	2
Corsicans, France	1	0	0	1	2
Friulians, Italy	1	0	1	1	1
Occitans, Italy	1	0	1	1	1
Basques, Spain	1	0	1	1	2
Catalans, Spain	1	0	1	1	2
Flemings, France	1	1	0	0	1
Walloons, Belgium	1	1	0	1	2
Swedes, Finland	1	1	1	0	2
South Tyroleans, Italy	1	1	1	0	2
Alsatians, France	1	1	1	1	1
Germans, Belgium	1	1	1	1	2
Flemings, Belgium	1	1	1	1	2

S = Size of subnation
L = Linguistic ability
W = Relative wealth of subnation
G = Growth vs. decline of subnational region
E = Degree of ethnic political mobilization

TRUTH TABLE REPRESENTATION OF DATA ON CAUSES OF ETHNIC POLITICAL MOBILIZATION

S	L	W	G	E	N
0	0	0	0	0	6
0	0	0	1	0	3
0	0	1	0	0	1
0	0	1	1	?	0
0	1	0	0	0	4
0	1	0	1	?	0
0	1	1	0	1	1
0	1	1	1	1	2
1	0	0	0	0	2
1	0	0	1	1	6
1	0	1	0	?	0
1	0	1	1	1	4
1	1	0	0	0	1
1	1	0	1	1	1
1	1	1	0	1	2
1	1	1	1	1	3

S = Size of subnation
L = Linguistic ability
W = Relative wealth of subnation
G = Growth vs. decline of subnational region
E = Degree of ethnic political mobilization

EQUATION: E = SG + LW

USING BOOLEAN ALGEBRA TO EVALUATE THEORIES

1. Intersection with the reactive ethnicity perspective

R = lw
E = SG + LW
R(E) = SlwG

Conforming cases (6): West Frisians (Netherlands), Catalans (France), Occitans (France), Bretons (France), Corsicans (France), and Welsh (Great Britain)

2. Intersection with the ethnic competition perspective

C = SW
E = SG + LW
C(E) = SW(G + L)

Conforming cases (9): Germans (Belgium), Flemings (Belgium), Swedes (Finland), Alsatians (France), Friulians (Italy), Occitans (Italy), South Tyroleans (Italy), Basques (Spain), Catalans (Spain)

3. Intersection with the developmental perspective

D = Lg
E = SG + LW
D(E) = LWg

One case uniquely covered: Aalanders (Finland)

4. Cases not covered by any theoretical perspective

H = lw + SW + Lg
h = (L + W)(s + w)(l + G) (using De Morgan's Law)
= slW + sLG + sWG + LwG

h(E) = (slW + sLG + sWG + LwG)(SG + LW)
= sLWG + SLwG

Cases covered by sLWG: Slovenes (Italy) and Valdotians (Italy)
Case covered by SLwG: Walloons (Belgium)

Analyzing Qualitative Data on "Multiple Instances"

Ethnographers and other qualitative researchers often have "multiple instances," meaning they have many cases of "the same thing." A researcher who conducts in-depth, open-ended interviews with 20 Olympic athletes in order to find out how they maintain their commitment has 20 instances of commitment to study. How should qualitative researchers generalize across instances?

One simple way to generalize across cases is to see if they all display the same ingredients for the outcome. These ingredients can be viewed as **causally relevant** or as **constitutive**. If they are viewed as constitutive, the outcome, in effect, is inherent or embedded in the conditions.

For example, the researcher studying how 20 Olympic athletes maintain their commitment might list some common ingredients, based on his/her field work: (1) devotion to a daily exercise routine, (2) a feeling of separateness from (or superiority to) non-athletes, (3) development of pre- or post work-out rituals, (4) food preferences and practices that make eating with others (especially non-athletes) problematic, and (5) associating primarily with others athletes. Do the 20 Olympic-level athletes share these features?

Case #	devotion to exercise	feeling of separateness	workout rituals	assoc. w/ athletes	separate food
1	yes	yes	yes	yes	yes
2	yes	no	yes	yes	yes
3	yes	yes	yes	no	no
4	yes	yes	yes	no	yes
5	yes	yes	yes	no	no
6	yes	yes	no	yes	yes
7	yes	no	no	yes	yes
8	yes	no	yes	yes	no
9	yes	no	yes	yes	yes
10	yes	yes	no	no	yes
11	yes	yes	yes	no	yes
12	yes	yes	no	yes	yes
13	yes	no	yes	yes	yes
14	yes	no	yes	yes	yes
15	yes	yes	yes	no	yes
16	yes	yes	no	yes	yes
17	yes	yes	yes	no	yes
18	yes	yes	yes	yes	yes
19	yes	yes	no	yes	yes
20	yes	no	yes	yes	no

Analyzing the Cross-Case Evidence

There are (at least) three ways to analyze this evidence.

1**. Analytic Induction**: The conditions must be shared uniformly. without exceptions. This is the technique developed by Alfred Lindesmith and
used in his study *Opiate Addiction*. It is strict and rigorous. According to the table (which is hypothetical), the only thing they all share is "devotion to exercise." So in very strict terms, there is not a whole lot of agreement across the 20 cases and not much to talk about.

2. **Composite Portrait**. What characteristics do “most” cases share? Notice that 13/20 have a feeling of separateness; 14/20 practice workout rituals; 13/20 associate primarily with other athletes; and 16/20 have distinctive food preferences or habits. The "composite portrait" has all five features, therefore, because all five features are more commonly present than absent across all 20 cases. This "composite portrait" can be reinforced simply by counting the number of "yes's" each case displays. Two cases have five yes's; twelve cases have four yes's; and six cases have three yes's. No case has less than 3 out of the 5 traits, which means even the "worst" fitting case has 3 of the 5 traits. It is possible to apply probabilistic criteria to these assessments, but it is rarely done.

3. **Substitutable Conditions.** It is also possible to look at the evidence in terms of "substitutability." The basic idea here is that there is more than one way to satisfy a general condition and that some of these feature may be alternate ways of satisfying the same condition. Notice, for example, that whenever there is a "no" in the "feeling of separateness" column there is a "yes" in the "associates mostly with other athletes" column. The reverse is also true: whenever there is a "no" in the "associates with athletes" column, there is a "yes" in the "feeling of separateness" column. If you think of these two things as "equivalent" or substitutable, then every Olympic athlete displays EITHER a feeling of separateness OR a preference for associating with other athletes. Notice that this same general relation exists between "workout rituals" and "separate foods." Whenever there is a "no" in one column, there is a "yes" in the other. Thus, these two features can also be seen as equivalent or substitutable. The general picture that emerges from this analysis is that the shared features are not one (devotion to exercise) but three:

1. Devotion to a daily exercise routine
2. EITHER feeling of separateness OR associates primarily with other athletes
3. EITHER performs workout rituals OR eats separate food

Substitutability and Concept Elaboration

Generally, when substitutability is identified, it goes hand-in-hand with concept development or elaboration. What do the pairs of substitutable "ingredients" represent at a more abstract level?

The second commonalities (feeling of separateness or associates primarily with other athletes) can be described at a more general level as the construction of a *social boundary* between athletes and non-athletes, either through patterns of association or through feelings related to personal identity.

The third commonality (workout rituals or separate food) could be seen as *daily activities and practices that reinforce that identity*.

The three ingredients for maintaining commitment as an Olympic athlete, therefore, could be described as:

1. Devotion to a daily exercise routine
2. Construction of a boundary separating athletes from non-athletes
3. Everyday activities that reinforce identity as an athlete

Substitutability and QCA

It is possible to analyze substitutability using truth table analysis. The steps are:

(1) create an outcome variable that is coded "1" for every case (maintains commitment),

(2) change all the 0's to dashes (according to our theoretical and substantive knowledge, if an individual has a "no" on one trait, changing it to "yes" would only bring about a stronger case for commitment), and

(3) analyze the resulting truth table, setting the "remainders" to false.

Here's the data set:

case	exr	feel	ritls	assoc	food	outcome
1	1	1	1	1	1	1
2	1	–	1	1	1	1
3	1	1	1	–	–	1
4	1	1	1	–	1	1
5	1	1	1	–	–	1
6	1	1	–	1	1	1
7	1	–	–	1	1	1
8	1	–	1	1	–	1
9	1	–	1	1	1	1
10	1	1	–	–	1	1
11	1	1	1	–	1	1
12	1	1	–	1	1	1
13	1	–	1	1	1	1
14	1	–	1	1	1	1
15	1	1	1	–	1	1
16	1	1	–	1	1	1
17	1	1	1	–	1	1
18	1	1	1	1	1	1
19	1	1	–	1	1	1
20	1	–	1	1	–	1

Here are the results:

EXERCISE FEEL RITUALS +
EXERCISE RITUALS ASSOC +
EXERCISE FEEL FOOD +
EXERCISE ASSOC FOOD

which can be factored to show the three conditions:

EXERCISE * (FEEL + ASSOC) * (FOOD + RITUALS)

FREQUENCIES AND CODES FOR VARIABLES USED IN BOOLEAN ANALYSIS OF CHALLENGING GROUPS

	Value	Freq.	Percent
Bureaucracy			
	0	29	54.7
	1	24	45.3
Lower Strata Constituency			
	0	28	52.8
	1	25	47.2
Displacement as Primary Goal			
	0	37	69.8
	1	16	30.2
Help From Outsiders			
	0	35	66.0
	1	18	34.0
Acceptance Achieved			
	0	28	52.8
	1	25	47.2
New Advantages Won			
	0	27	50.9
	1	26	49.1

Values show coding in qualitative comparative analysis: 1 indicates presence; 0 indicates absence.

Truth Table For Causes of New Advantages*

					Number of Cases	New Adv.	No New Adv.	Outcome Variables		
BUR	**LOW**	**DIS**	**HLP**	**ACP**				**U**	**L**	**P**
0	0	0	0	0	4	2	2	0	0	1
0	0	0	0	1	2	2	0	1	1	1
0	0	0	1	0	2	2	0	1	1	1
0	0	0	1	1	2	2	0	1	1	1
0	0	1	0	0	4	0	4	0	0	0
0	0	1	0	1	1	1	0	–	–	1
0	0	1	1	0	2	0	2	0	0	0
0	0	1	1	1	1	0	1	0	0	0
0	1	0	0	0	2	0	2	0	0	0
0	1	0	0	1	0	remainder		?	?	?
0	1	0	1	0	0	remainder		?	?	?
0	1	0	1	1	2	2	0	1	1	1
0	1	1	0	0	5	0	5	0	0	0
0	1	1	0	1	0	remainder		?	?	?
0	1	1	1	0	2	0	2	0	0	0
0	1	1	1	1	0	remainder		?	?	?
1	0	0	0	0	3	0	3	0	0	0
1	0	0	0	1	4	1	3	0	0	1
1	0	0	1	0	1	1	0	–	–	1
1	0	0	1	1	1	1	0	–	–	1
1	0	1	0	0	1	0	1	0	0	0

1	0	1	0	1	0	remainder		?	?	?
1	0	1	1	0	0	remainder		?	?	?
1	0	1	1	1	0	remainder		?	?	?
1	1	0	0	0	2	1	1	0	0	1
1	1	0	0	1	7	6	1	0	1	1
1	1	0	1	0	0	remainder		?	?	?
1	1	0	1	1	5	5	0	1	1	1
1	1	1	0	0	0	remainder		?	?	?
1	1	1	0	1	0	remainder		?	?	?
1	1	1	1	0	0	remainder		?	?	?
1	1	1	1	1	0	remainder		?	?	?

* Column headings: BUR = bureaucratic organization; LOW = lower strata constituency; DIS = displacement as primary goal; HLP = help from outsiders; ACP = acceptance of the organization. 1 indicates presence; 0 indicates absence. The output is coded as follows: U = uniform new advantages; L = new advantages likely; P = new advantages possible. The don't care output coding is indicated with a dash.

Minimization of Truth Tables for New Advantages*

Remainder Terms = False (0)	Remainder Terms = Don't Care (-)
UNIFORM = dis*HLP*ACP + low*dis*HLP + bur*low*dis*ACP	UNIFORM = dis*HLP + bur*dis*ACP
LIKELY = dis*HLP*ACP + low*dis*HLP + bur*low*dis*ACP + BUR*LOW*dis*ACP	LIKELY = dis*HLP + bur*dis*ACP + LOW*ACP
POSSIBLE = dis*HLP*ACP + low*dis*HLP + bur*low*dis + BUR*dis*ACP + BUR*LOW*dis*hlp + bur*low*hlp*ACP	POSSIBLE = dis*HLP + bur*low*dis + BUR*LOW + hlp*ACP

*Variable names in upper case letters indicate presence; variables names in lower case letters indicate absence; addition indicates logical OR; multiplication (*) indicates logical AND.

CONSTRUCTING TRUTH TABLES

- 1. Define the outcome, based on the best positive cases: What is it? Sharpen definition by comparing the outcome to related-but-different outcomes (e.g., comparing social revolutions to "mere" political revolutions). Identify the positive cases of the outcome (preliminary list). Some "positive cases" may not have full membership in the set defined by the outcome. Think of the outcome as an ideal type.

- 2. Identify causal conditions shared by positive cases. Which causal conditions seem important in the best examples of the outcome? Shared causal conditions do not have to perfectly universal, but must be important across a range of cases. If this search seems to fail (i.e., you cannot identify general causes or shared features), you may need to return to step #1 and define types of the outcome or focus on a more narrowly defined outcome. Alternatively, you may conclude at this early point that there are no shared causal conditions for the outcome. This makes it more difficult to identify negative cases (i.e., good candidates for the outcome which nevertheless did not experience it). Generally, it is possible to find very general commonalities, even if they are causally distant from the outcome (e.g., "debtor countries" in a comparative study of protest against the IMF). The important thing to remember here is that you are looking for raw material for defining the population when you examine commonalities across positive cases.

- 3. Identify, if possible, **candidates** for the outcome, based on your analysis of features shared by the positive cases. The goal is to find negative cases that have as much as possible in common with the positive cases. Compare and contrast positive cases with these "candidate" (i.e., negative) cases. Perhaps sharpen definition of the outcome (e.g., identify different-but-equivalent outcomes among the candidate cases). You may need to return to step #1 or #2 as you try to work through the problem of identifying negative cases. Such cases may be impossible to identify. For example, the common causes identified in set #2 may be necessary and sufficient for the outcome, in which case you won't be able to find a case that shares these causal conditions without sharing the outcome. If this is happens, you should try to find negative cases with as many of the shared causal conditions as possible.

- 4. Constitute initial population of cases, embracing cases with outcome and candidate cases. These case should display many (if not most) of the necessary conditions for the outcome (the shared causal conditions identified in step #2). There are other ways to constitute populations. This method (working from commonalities shared by positive cases) is the most inductive and avoids accepting "given" populations as the starting point,

- 5. Consider possible sufficient conditions: What kinds of factors distinguish positive from candidate cases? Often it is useful at this point to compare the "best" cases of the outcome with the "best" negative cases (i.e., the best candidates for the outcome, defined as those that share as many causally relevant conditions as possible). Add any newly identified causally relevant conditions to the analytic frame.

- 6. Construct the property space, showing all the cases and the relevant causes. Think of each combination of causally relevant conditions as an initial grouping. Each combination constitutes (potentially) a different type of case. Notice that you are viewing cases as configurations. The property space is not simply a listing of cases by variables, but a definition of the different configurations of relevant causal conditions. Notice also that you are interested in looking at all possible combinations of causal conditions, not just those with cases.

- 7. Study the groupings of cases established in the truth table (or property space). Ask: do the cases grouped together (in a row) really belong together, based on wider knowledge (and given goals of study)? If so proceed. If not return to step #1 (or step #2, #3, or #5). This examination should be conducted without regard for their scores on the outcome.

- 8. Study the outcomes associated with each grouping of cases (on a given row). Do the cases with each combination of attributes have the same (or more-or-less the same) outcome? If not, return to step #1, #2, #3, or #5.

The Distinctiveness of Case-Oriented Research

In a forthcoming comment in *Studies in Comparative International Development*, Christopher Achen, a well-known quantitative researcher, notes:

> Few social scientists dispute the need to combine qualitative and quantitative methods and evidence in the profession. The question is how. As . . . [many] scholars have said, first-rate social science theorizing seems to integrate the two in ways we do not fully understand. For example, contemporary case-study methods are difficult to explicate in conventional statistical theory, and yet they are frequently quite powerful and successful in ways that no statistical methods could match. *An important clue is that they often carry out an implicit comparison against known background relationships*, most obviously so in single-case studies (Ragin 2000:206). But what is the precise inferential logic of this step and why is it so successful? No one knows. (*Italics added*)

Olav Stokke's Truth Table for Causes of Successful Shaming in International Regimes

Advice (A)	Commitment (C)	Shadow (S)	Inconvenience (I)	Reverberation (R)	Success (Y)
1	0	1	1	1	1
1	0	0	1	0	0
1	0	0	1	1	0
0	0	0	1	0	0
1	1	1	1	1	1
1	1	1	1	0	0
1	1	1	0	0	1
1	0	0	0	0	1

1. Advice (A): Whether the shamers can substantiate their criticism with reference to explicit recommendations of the regime's scientific advisory body.
2. Commitment (C): Whether the target behavior explicitly violates a conservation measure adopted by the regime's decision-making body.
3. Shadow of the future (S): Perceived need of the target of shaming to strike new deals under the regime--such beneficial deals are likely to be jeopardized if criticism is ignored.
4. Inconvenience (I): The inconvenience (to the target of shaming) of the behavioral change that the shamers are trying to prompt.
5. Reverberation (R): The domestic political costs to the target of shaming for not complying (i.e., for being scandalized as a culprit).

HOW STOKKE'S EVIDENCE IS TYPICAL

- The number of cases (8) is more than a handful, but still small enough to permit familiarity with each case.

- From the viewpoint of conventional quantitative social science, however, the number of cases is very small relative to the number of causal conditions (5). This ratio essentially eliminates the possibility of any form of multivariate statistical analysis.

- If the cases are viewed configurationally, then the prospects seem even more discouraging, for there are 2^5 logically possible combinations of five causal conditions. We have empirical evidence on only eight of the 32 combinations.

- This pattern of **limited diversity** is characteristic of comparative research and, more generally, of research on naturally occurring social and political phenomena.

- Causal combinations without cases are potential counterfactual cases.

SIMPLE EXAMPLE OF LIMITED DIVERSITY

Strong Unions (U)	Strong Left Parties (L)	Generous Welfare State (G)	N of Cases
Yes	Yes	Yes	6
Yes	No	No	8
No	No	No	5
No	Yes	????	0 (they don't exist)

Is it strong left parties (L) that cause generous welfare states (G) or is it the combination of strong unions and strong left parties (L*U) that causes generous welfare states (G)?

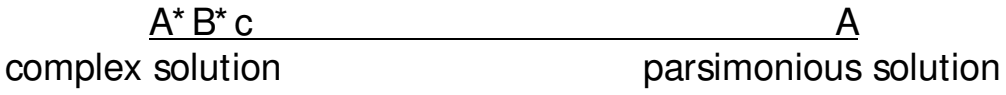
From a correlational viewpoint, having a strong left party (L) is perfectly correlated with having a generous welfare state (G). A parsimonious explanation has been achieved.

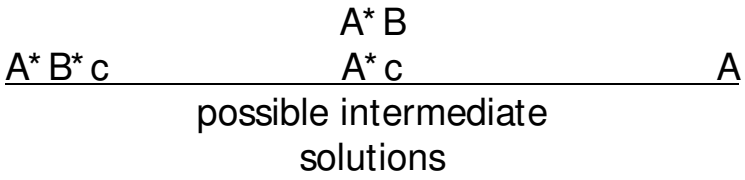
From a case-oriented perspective, however, all instances of generous welfare state share two causally relevant conditions (strong left parties and strong unions) and none of the negative cases display this combination. This pattern suggests a more complex explanation.

Limited Diversity in a Truth Table with Four Causal Conditions

A	B	C	D	Y
no	no	no	no	no
no	no	no	yes	?
no	no	yes	no	?
no	no	yes	yes	?
no	yes	no	no	no
no	yes	no	yes	no
no	yes	yes	no	?
no	yes	yes	yes	no
yes	no	no	no	?
yes	no	no	yes	?
yes	no	yes	no	?
yes	no	yes	yes	?
yes	yes	no	no	yes
yes	yes	no	yes	yes
yes	yes	yes	no	?
yes	yes	yes	yes	?

PARSIMONY VERSUS COMPLEXITY (HYPOTHETICAL DATA)





At the left end of the continuum is the complex solution; the right end shows the parsimonious solution. The complex solution is a subset of the parsimonious solution.

Assume theoretical and substantive knowledge indicates that it is the presence of these four conditions (A, B, C, D) and not their absence (a, b, c, d) that should be linked to the outcome (Y). This knowledge defines A*B*C as an **easy** counterfactual, yielding solution A*B; it defines A*b*c as a **difficult** counterfactual. (This second counterfactual is what is required to produce A*c as a solution.)

Olav Stokke's Truth Table for Causes of Successful Shaming in International Regimes

Advice (A)	Commitment (C)	Shadow (S)	Inconvenience (I)	Reverberation (R)	Success (Y)
1	0	1	1	1	1
1	0	0	1	0	0
1	0	0	1	1	0
0	0	0	1	0	0
1	1	1	1	1	1
1	1	1	1	0	0
1	1	1	0	0	1
1	0	0	0	0	1

1. Advice (A): Whether the shamers can substantiate their criticism with reference to explicit recommendations of the regime's scientific advisory body.
2. Commitment (C): Whether the target behavior explicitly violates a conservation measure adopted by the regime's decision-making body.
3. Shadow of the future (S): Perceived need of the target of shaming to strike new deals under the regime--such beneficial deals are likely to be jeopardized if criticism is ignored.
4. Inconvenience (I): The inconvenience (to the target of shaming) of the behavioral change that the shamers are trying to prompt.
5. Reverberation (R): The domestic political costs to the target of shaming for not complying (i.e., for being scandalized as a culprit).

PARSIMONY VERSUS COMPLEXITY IN STOKKE'S EVIDENCE

A·c·s·i·r +
A·C·S·i·r +
A·S·I·R

complex

i +
S·R

parsimonious

A·c·s·i·r +
A·C·S·i·r +
A·S·I·R

A·i +
A·S·R

i +
S·R

intermediate

In the complex solution, none of the combinations without cases is used as a counterfactual case. In the parsimonious solution, any combination without cases that yields a simpler solution is incorporated into the solution (i.e., both easy and difficult counterfactuals). The assumptions are: A, C, S, i, R. These assumptions yield the intermediate solution.

Combination A·S·I·R:

1. Causal conditions S and R cannot be removed because they appear in the corresponding parsimonious term at the other end of the continuum.
2. The support of the regime's the scientific advisory body (A) is certainly linked to the success of shaming. This causal condition should be retained.
3. The fact that it is inconvenient for the targets of shaming to change their behavior (I) does *not* promote successful shaming. Thus, inconvenience (I) can be dropped from the combination A·S·I·R because inconvenience of behavioral change to the target of shaming is not central to the success of A·S·R in generating conformity.

The intermediate combination is **A·S·R**.

Combination A·C·S·i·r:

1. Condition i (the behavioral change is not inconvenient) cannot be dropped because it appears in the corresponding parsimonious term.
2. Condition A (support from the regime's scientific advisory board) should remain because this condition is clearly linked to the success of shaming.
3. Condition C (the offending behavior clearly violates a prior commitment) also should not be dropped, for this too is something that should only contribute to the success of shaming.
4. Condition S (the violator will need to strike future deals with the regime) is also a factor that should only promote successful shaming.
5. Condition r (absence of domestic reverberations for being shamed) can be removed. Clearly, the presence of domestic reverberation (R) would promote successful shaming.

The intermediate combination is **A·C·S·i**.

Combination A·c·s·i·r:

1. Condition i must be retained because it appears in the corresponding parsimonious term.
2. Condition A is retained as well, for the reasons stated previously.
3. Condition r (absence of domestic reverberations) can be removed, as it was from the previous combination, for the same reason provided.
4. Condition c (absence of violation of a commitment) can be removed, for surely these instances of successful shaming would still have been successful if there had been an explicit violation of a commitment (C).
5. Condition s (absence of a need to strike future deals with the regime) can be safely removed because only its presence (S) should contribute to the success of shaming.

The intermediate term is **A·i**.

These three intermediate terms can be joined into a single equation:

$$\mathbf{A \cdot S \cdot R + A \cdot C \cdot S \cdot i + A \cdot i \longrightarrow Y}$$

which can then be simplified to:

$$\mathbf{A \cdot S \cdot R + A \cdot i \longrightarrow Y}$$

because the term A·C·S·i is a subset of the term A·i and is thus logically redundant. (All cases of A·C·S·i are also cases of A·i.) These results indicate that there are two paths to successful shaming: (1) support from the regime's scientific advisory body (A) combined with the need to strike future deals (S) and domestic reverberations for being shamed (R), and (2) support from the regime's scientific advisory body (A) combined with the fact that the behavioral change is not inconvenient (i).

Limited Diversity in Commonly Used Data Sets

college	p_income	afqt	married	kids	number	consistency
0	0	0	0	0	327	0.452813
0	0	0	1	1	154	0.488127
0	0	0	0	1	65	0.25492
1	0	0	0	0	57	0.792723
0	0	0	1	0	41	0.815288
1	0	0	1	1	24	0.804862
0	1	0	0	0	14	0.746632
1	0	0	1	0	13	0.962559
0	1	0	1	1	10	0.85515
1	1	0	0	0	10	0.941069
1	0	0	0	1	7	0.466333
1	0	1	1	1	6	0.949928
0	1	0	1	0	5	0.972446
1	1	0	1	0	4	1
1	1	1	0	0	4	0.986691
1	0	1	0	0	3	0.915131
1	1	0	1	1	3	0.944169
1	1	1	1	0	3	1
0	0	1	1	1	2	0.782026
1	1	1	1	1	2	0.922
0	0	1	0	0	1	0.993634
0	0	1	0	1	1	0.881229
0	1	0	0	1	1	0.610116
1	0	1	1	0	1	1
0	0	1	1	0	0	1
0	1	1	0	0	0	1
0	1	1	0	1	0	1
0	1	1	1	0	0	1
0	1	1	1	1	0	0.638853
1	0	1	0	1	0	
1	1	0	0	1	0	0.908452
1	1	1	0	1	0	

CONCLUSIONS

1. Limited diversity is a characteristic feature of naturally occurring social phenomena.

2. The resolution of the problem of limited diversity involves counterfactual analysis, in some way.

3. In case-oriented comparative research, the resolution of the problem of limited diversity is knowledge and theory dependent. "How" this happens in case-oriented research (Achen's query) is through the incorporation of "easy" counterfactuals.

4. In order to define "easy" counterfactuals, researchers apply their substantive and theoretical knowledge to the "remainder" combinations. In practice, this allows them to craft an intermediate solution, situated between the "most complex" and "most parsimonious" QCA solutions. maintaining the subset relationship among possible solutions.

5. In quantitative research, the problem is also addressed through assumptions. However, these assumptions (e.g., linearity and additivity) are usually invisible to users. They are rarely examined or challenged, even when they are known to be unreasonable.