Exercises for:

Social science concepts and measurement: new and completely revised edition (Princeton 2020)

and

Multimethod research, causal mechanisms, and case studies: an integrated approach (Princeton 2017)

and

A tale of two cultures: qualitative and quantitative research in the social sciences (Princeton 2012)

and

Politics, gender, and concepts: theory and methodology (Cambridge 2008)

and

Explaining war and peace: case studies and necessary condition counterfactuals (Routledge 2007)

Version 4 (September 2020)

Gary Goertz
University of Notre Dame
ggoertz@nd.edu
This marks the fourth version of my collected exercises. There are a lot of new exercises, particularly with the new edition of the concept book. The exercises are divided into sections with the Table of Contents given below. Exercises appear roughly in the chronological order in which they were added, more recent first. Some older exercises have been edited in the light of newer books and the like. Also, a few old ones have been removed if they have been replaced by more relevant newer ones.

This edition represents a pretty massive update to the exercises. I have completely rewritten and expanded my concept book (September 2020). This was in part motivated by exercises that I had developed over the years. As a result, there are dozens of new exercises that cover old as well as a lot of new material in the concept book. The book has much more on measurement and quantification of concepts and that is reflected in sections on scaling and aggregation.

The book also focuses on the massive industry of global indices, indexes, and barometers. A lot of this occurs in the international policy space regarding economic development, governance. International institutions like the World Bank, the European Union, United Nations, as well as many NGOs produce complex multi-level concepts and measures that are the focus of the concept book. New exercises deal with this large and expanding industry.

My current research interests focus on expanding the previous books’ discussion of causal mechanisms in a variety of new directions. Much of this has to do with process tracing and how that connects with causal mechanism analysis. For example, with Haggard and Copestake we are pursuing a new process tracing methodology called the Veil of Ignorance Process Tracing (VoiPT) methodology (Copestake, Goertz, and Haggard 2020). In addition, I continue work on what Steph Haggard and I call Large-N Qualitative Analysis (for a first discussion see chapter 7 in the multimethod book, see Goertz and Haggard 2020). In fact LNQA has its own new section. In addition, I am working a lot on external validity, generalization, and extrapolation in a variety of directions. This is absolutely central to applied policy analysis and a core issue in deciding how generalizable experiments are. One should expect quite a few new exercises on causal mechanisms, external validity, generalization, and process tracing in the next edition of the exercises.

I have long been interested in the connection between concept analysis and gender. Gender critiques of existing literature often start with an analysis of gender biases of core concepts. Early work on this resulted in the anthology edited with Amy Mazer published by Cambridge in 2008. In the new concept and measurement book gender is an important theme. Therefore, I’ve created a new section dealing with gender issues in concepts in particular, but also methodology more generally.

---

1 https://press.princeton.edu/books/paperback/9780691205489/social-science-concepts-and-measurement

2 "Generalization, case studies, and within-case causal inference: Large-N Qualitative Analysis (LNQA)" in Kincaid, H. and J. Van Bouwel (eds.), forthcoming, The Oxford handbook on the philosophy of political science.

---
The last big update to these exercises occurred when *Multimethod research, causal mechanisms, and case studies: an integrated approach* was published by Princeton University Press (August 2017). In addition to multimethod research it covers, multiple (or comparative) case study research, using case studies to systematically test hypotheses. I have included new sections below. “Case studies, case selection” deals with selecting cases for multiple case study or multimethod research. “Multimethod research” includes not only the most common statistical analyses plus case studies, but also game theory multimethod defined as a formal model plus case studies (a whole chapter of the book is devoted to this). As the book title indicates causal mechanisms are absolutely central and hence there is a new section treating this topic.

I have created categories which span all of these projects. In general there is no specific indication which book is being referred to, and in many cases an exercise could easily apply to multiple books. However, if the exercise deals with concepts it is likely to be the *Social science concepts and measurement*. If is on case selection it could be any of them except the *Case studies and necessary condition counterfactuals* book. If in contrast, it is about within-case causal explanations of World War I or the end of the Cold War then it is likely to be the *Explaining war and peace* book.

Answers to these exercises (except those that are open-ended and those that I think would be interesting but have not yet had the time to do) are included as well.

If you would like to be informed when these exercises are updated please contact me and I will put you on the email list (ggoertz@nd.edu).

Most of the articles and book chapters referred to in the following exercises are available in pdf format and complete references are found at the end. If possible I have chosen an electronically available article rather than a book. Some of the exercises work well in more than one section and hence an exercise may appear more than once.

As always, I welcome comments on these exercises and suggestions for new ones.

---

## Table of contents

**Concepts** – General issues of concept construction, guidelines, and specific concepts. – p. 4

**Semantic transformations, calibration** – Issues of semantic meaning, transformations, and creating continuous measures – p. 28

**Terminology, definitions, and semantics** – Issues of semantics, meaning, definitions, wording, etc. – p. 32

**Scaling** – Scaling issues, scale types, – p. 36

**Concept structure and aggregation** – Weakest link, best shot, mean, hybrid, family resemblance, etc. – p. 47

**Typologies** – mutually exclusive, exhaustive rules, (partially) ordered, etc. – p. 62

**Latent variable statistical models** – methodological features of and specific latent variables – p. 65

**Normative issues** – normative issues in concept construction – p. 67

**Global indices, barometers, etc.** – global indicators and barometers, popular with NGOs, UN, EU, OECD, etc. – p. 69

**Gender** – gender and concepts, research design, etc. – p. 70

**Case studies, case selection** – case selection comparative case studies, multimethod case selection, case study methodology issues – p. 77

**The Possibility Principle** – case selection comparative case studies and statistical analyses, focusing on population and negative case selection – p. 83

**Multimethod research** – multimethod research issues of all sorts – p. 99

**Theories and concepts, e.g., two-level theories** – complex concepts in complex theories – p. 103

**Causal issues, causal hypotheses, causal asymmetry, etc.** – causation, causal inference, etc. – p. 113

**Causal mechanisms** – causal complexity, causal mechanism figures, methodology and philosophy of – p. 123

**Constraint mechanisms** – necessary conditions, norms, permissive causes, etc. – p. 130

**Process tracing, within case causal inference, counterfactuals** – causal inference in individual cases, process tracing methodology, counterfactual methodology, etc. – p. 130

**Generalization, external validity, Large-N Qualitative Analysis (LNQA), scope, etc.** – descriptive generalization, extrapolation, etc. – p. 140

**QCA, set theory, Boolean algebra, fuzzy logic** – necessary and sufficient conditions, INUS, set theoretic methodology – p. 145
Concepts

1. A very difficult design decision in constructing concepts is whether to include potential causes or effects of the concept in the concept itself. Diamond and Morlino (2004, 2005) discuss various dimensions of the “quality” of democracy. Some of these might be better thought of as potential causes or effects of democracy. For example, they discuss how equality or education is important in the concept of a high-quality democracy. Discuss the various advantages and disadvantages of including dimensions such as equality or education in the concept of (quality of) democracy. Similarly, one can think about the effects of democracy: do Diamond and Morlino include effects in the concept of quality of democracy that would be independent variables in important analyses where democracy is a key dependent variable?

Answer:

One problem with the Diamond and Morlino book is that it is purely about the quality of democracy. There is no sense in which this concept is embedded in some theoretical or empirical program looking at the causes or effects of democracy. If one wants to discuss the causes of democracy (or quality of democracy) one should think hard before including potential independent variables in the concept itself. Also, they define a quality democracy as one with effective rule of law and effective administrative outputs. To what extent does one want to define democracy quality in terms of quality of bureaucracy? Would it be better to keep these as separate concepts? This way we can explore the causal relationships between the two.

2. Discuss how it is possible that Moyn (2012) can survey the historiography of human rights without talking about democracy at all, i.e., the word does not appear in the review article. What does the absence of democracy in the discussion mean about the concept of human rights guiding Moyn’s discussion?

3. There is often overlap between closely connected concepts. Should one worry about this or attempt to make them separate? For example, overlap occurs regularly between various conflict concepts such as coups and civil wars:

While we have gone to great lengths to assure that coups are not conflated with other forms of anti-regime activity in our dataset, there is no guarantee that scholars coding other events have excluded coups in theirs. For example, we have identified 38 events within the Uppsala/PRIO dataset and 4 events in the Correlates of War (Sarkees, 2000) internal conflict dataset that are best described as coups. Scholars who use these datasets to operationalize civil wars should be wary of including these events. (Powell and Thyne 2011, 256).

4. Negations or negative poles often are tricky, particularly when there is concept asymmetry. Discuss the claim at the end of the following quote about two “direct opposites,” particularly in the light of the claim that populism has three core dimensions:

Beyond the lack of scholarly agreement on the defining attributes of populism, agreement is general that all forms of populism include some kind
of appeal to “the people” and a denunciation of “the elite.” Accordingly, it is not overly contentious to state that populism always involves a critique of the establishment and an adulation of the common people. More concretely, we define populism as a thin-centered ideology that considers society to be ultimately separated into two homogeneous and antagonistic camps, “the pure people” versus “the corrupt elite,” and which argues that politics should be an expression of the volonté générale (general will) of the people. . . . [T]here are at least two direct opposites of populism: elitism and pluralism. . . . Populism has three core concepts: the people, the elite, and the general will. (Mudde and Kaltwasser 2017, epub 38, 44)

5. Concepts are frequently asymmetric, e.g., democracy–autocracy. These is true for the concept pair healthy–ill. Discuss this in the context of defining health as the absence of disease:

   In a series of essays published in the 1970s, Christopher Boorse developed a powerful analysis of the theoretical notion of health—that is, the concept of health that is of concern to pathologists and physiologists. In his original essays, Boorse took health to be the absence of “disease,” where disease is understood in a wide sense that includes injuries and disabilities. To avoid confusion, he has changed his terminology (but not his view) and now takes health to be the absence of pathology. In Boorse’s view, whether a physical or mental state constitutes a pathology does not depend on judgments about how good that state is for people. Boorse defends a naturalistic, nonnormative view of pathology and health. . . . There is a large literature concerning the concept of health, most of which is critical of Boorse’s view. (Hausman 2015, 8)

6. Use the basic concept framework to simplify this complex weighting system.

   Most systems of health measurement define health states in terms of a small number of easily observable levels, along some small number of dimensions. The dimensions have been functional capacities, such as vision or cognition; consequences of multiple physical and mental capacities, such as “self-care”; and subjective states, such as pain or anxiety. So, for example, the Health Utilities Index, Mark 3 (HUI(3)) has eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. Even with only eight dimensions, if one allows several gradations along each dimension, the number of distinguishable health states becomes very large. The HUI(3) distinguishes five or six levels along each dimension for a total of 972,000 health states. . . . Once health economists have a classification of health states, such as the EQ-5D or the HUI(3), they need to assign numbers to health states that measure the value of the health state and (once one takes into account time) of health. Health economists call these numbers “quality weights.” To quantify health states requires assigning quality weights to the health state vectors defined by a health-state classification system or to some set of sequelae of diseases and injuries. The health classification system determines the magnitude of the measurement task. . . . the HUI(3) requires assigning quality weights to 972,000 health states. (Hausman 2015, 36–37)

   Answer:

   Instead of considering five or six levels for each secondary level dimension create a semantic transformation for each dimension and then decide the weights for each of the five or six secondary level dimensions.
7. The Redundancy Guideline stresses that redundancy can be problematic between concepts particularly closely related concepts. Discuss this solution to the problem of redundancy or conceptual overlaps between transparency and democracy. This arises because by definition good democracies are transparent. Discuss also the relationship between tautology and redundancy.

The study of transparency and its relationship with democracy is fraught with difficulties, one of which is the possibility of tautology. While we are concerned in our study with the transparency of policymaking and, more specifically, with the dissemination of information, the broad concept of transparency also applies to a full range factors that affect information flow within a society. Transparency thus pertains to questions of who rules, how governments might be replaced, and, indeed, how elections are contested in countries that fill key offices through an electoral process. In other words, transparency may pertain to the very question of whether one can call a political regime “democratic.” … The analytical tension is obvious: if democracy is, by definition, transparent, then the degree to which governments enjoy and promote obfuscation is simply a measure of their nondemocraticness. … We thus employ a minimalist definition of democracy. Following Schumpeter (1942) and more recently Przeworski et al. (2000), we define democracy as a regime in which the executive and the legislature are both filled by “contested elections.” … We therefore employ a similarly narrow definition of “transparency.” … We thus explicitly distinguish the transparency of the electoral system from policy transparency (Hollyer et al. 2011, 1192–93; redundancy solution and conceptual tautology)

8. Discuss the claim about bounded versus unbounded scales in Morris’s The measure of civilization how social development decides the fate of nations. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How serious critique of that approach is this?

It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country’s score in each annual report, but because the maximum possible score is always 1.0, the HDI does better at charting a nation’s relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

9. Discuss this (dubious) claim in a standard textbook on measurement in medicine:

Only unobservable constructs require a measurement theory. For observable characteristics, it is usually obvious how the items contribute to the construct being measured and no measurement theory is required. We illustrate this with a few examples. Physical activity can be characterized by frequency, type of activity and intensity. To obtain the total energy expenditure we know how to combine these items. Moreover, for some research questions we are only interested in certain types of physical activity or only in the frequency of physical activity. To assess the severity of diarrhea, a clear example of an observable characteristic, fecal output can be characterized by frequency, amount and consistency. Another example
concerns comorbidity, which is characterized by the number of accompanying diseases, the type of diseases or organ systems involved, and the disease severity or the disability or burden they cause. However, if we talk about comorbidity burden, we move in the direction of unobservable constructs. (Vet et al. 2011, 17)

10. The Ideal Type Guideline is a central one for concept construction. This can mean “conceptual creep” as the conceptual scale expands. Discuss the implications, causal and conceptual of such creep as well as the intension–extension implications of such creep. Why might having a clear semantic transformation be critical in this process?

As it turns out, abstract concepts can creep, too. For example, in 1960, Webster’s dictionary defined “aggression” as “an unprovoked attack or invasion,” but today that concept can include behaviors such as making insufficient eye contact or asking people where they are from. Many other concepts, such as abuse, bullying, mental disorder, trauma, addiction, and prejudice, have expanded of late as well. Some take these expansions as signs of political correctness and others as signs of social awakening. We take no position on whether these expansions are good or bad. Rather, we seek to understand what makes them happen. Why do concepts creep? (Levari et al. 2018, 465)

11. Draw the figure that corresponds to this concept of ‘international court authority’ and discuss the dimensions and structure.

Our measure of de facto IC authority has two key components: (1) recognizing an obligation to comply with court rulings and (2) engaging in meaningful action pushing toward giving full effect to those rulings. This is a conjunctive standard that is assessed by examining the practices of the relevant actors. A simple public statement that a judgment is legally binding is, without more, inadequate. Equally insufficient is conduct that happens to conform to a judgment. . . . Our framework identifies five types of de facto authority that correspond to the practices of different IC audiences, which we illustrate with examples from the empirical chapters of this book. . . . Figure 2.1 displays the three levels of de facto authority as nested, suggesting that a court first gains narrow, then intermediate, then extensive authority. But as we explain in the discussion that follows, we intend no teleology by this diagram. The three circles of de facto authority may not be nested; an IC could have extensive authority but lack narrow and intermediate authority. (Alter et al. 2018, 29, 31, 33, 34)

12. Do classic Mafia protection rackets count as terrorism?

Terrorism is the premeditated use or threat of use of extranormal violence or brutality by subnational groups to obtain a political, religious, or ideological objective through intimidation of a huge audience, usually not directly involved with the policy making that the terrorists seek to influence. (Enders and Sandler 2002, 145–46)

Answer:
Because the Mafia has an economic goal it is not terrorism, because it is not “political, religious, or ideological objective.” If we added “economic” to
the list (sometimes it does appear in definitions of terrorism) then the Mafia would be a terrorist group. Would it be better to define terrorism in terms of “non-economic” goals?

13. The Redundancy Guideline raises the issue of semantically redundant versus empirically correlated. Discuss this issue in the context of HDI and similar indicators, for example:

On the other hand, redundancy in the three dimensions of the HDI was also pointed out right after its introduction. In this respect, McGillivray (1991) disclosed a significant and positive correlation between each dimension index. While supporting this finding, Ivanova et al. (1999) further demonstrated that the overall HDI rankings did not change significantly even if measured either by life expectancy only or by the combination of the other indicators. This point was echoed later by Cahill (2005) who thus reached the conclusion that the HDI is not sensitive to the weights employed. By contrast, McGillivray and White (1993) found redundancy in all countries but not within the same group of countries. This result was also proved by Noorbakhsh (1998c). While interpreting the high correlation in all countries as a sign of internal consistency rather than redundancy, Kovacevic (2011) further revealed that the lower two groups in the HDI category differed between themselves and from the upper two in terms of the correlation structure. (Hirai 2017, 85)

14. A key issue in the three-level framework is whether the same data-indicators can be used for different secondary-level dimensions. See how this occurs in “Measuring the rule of law”: “A careful examination of the nine factors reveals two facts. The first is that there is a partial overlap among sub-factors; that is, a sub-factor can simultaneously belong to different factors at once. This is simply to reflect the fact that various rule-of-law dimensions partially overlap in practice” (Botero and Ponce 2011, 16).

15. See the same issue of the same data-indicators in multiple dimensions in Dahl’s (1971) famous conceptualization of democracy.

I. Formulate preferences:
A. Freedom to form and join organizations
B. Freedom of expression
C. Right to vote
D. Right of political leaders to compete for support
E. Alternative sources of information
II. Signify preferences
A. Freedom to form and join organizations
B. Freedom of expression
C. Right to vote
D. Eligibility for public office
E. Right of political leaders to compete for support
F. Alternative sources of information
G. Free and fair elections
III. Have preferences weighted equally in conduct of government
A. Freedom to form and join organizations
B. Freedom of expression
C. Right to vote
D. Eligibility for public office
E. Right of political leaders to compete for support
F. Alternative sources of information
G. Free and fair elections
H. Institutions for making government policies depend on votes and other expressions of preference

16. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the “one-drop rule”?

The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the census’s history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

17. The well-known UCDP data projects have the goal of basically mapping all forms of armed conflict, both domestic and international. Discuss the extent to which their typology is mutually exclusive and exhaustive. How is this an application of the Redundancy Guideline? Why is this critical for core UCDP goals.

Organized violence 1989–2017 For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as ‘organized violence’. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting
between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the inter-ethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 539; see https://www.pcr.uu.se/research/ucdp/definitions/)

Answer:

A core UCDP goal is to give an survey the amount of armed conflict in the world. This is essentially adding up all the various kinds of armed conflict into a total. Therefore the categories need to be mutually exclusive so there is no doublecounting. In fact, the addition of new datasets to the UCDP is probably working toward the goal of completeness. This might be signaled by the change in the name of the yearly data articles to “Organized violence” in 2016 “What we now call ‘state-based armed conflict’ was called just ‘armed conflict’ in our earlier publications.” (Melander et al. 2016)

18. Compare Rawls and his typology of war with UCDP.

Rawls distinguished between different types of wars in order to define the principles that would best apply in each case. The initial typology found in his lecture notes proposes nine kinds: 1. Wars between existing via states (WW I + II) 2. Civil wars (of social justice) within via states or society (French Rev); 3. Wars of secession of minorities within region: American Civil War. 4. Colonial Wars of secession (from Empire): Algerian War; American Rev War? 5. Wars of intervention (humane intervention) 6. Wars of national unification (War of Roses; Tudors) 7. Wars of conquest, of Empire (Wars of Rome). 8. Wars of Crusade, religious or secular 9. Wars of national liberation (in present sense); guerilla wars (Armitage 2017, 335)

19. One of the obvious problems with definitions and conceptualization is when the concept being defined also appears in the definition. This is not uncommon with various “peace” concepts, such as “peacebuilding.” Barnett (2006) illustrate this:

Although peacebuilding is generically defined as external interventions that are designed to prevent the eruption or return of armed conflict, there are critical differences among actors regarding its conceptualization and operationalization. This article surveys and analyzes twenty-four governmental and intergovernmental bodies that are currently active in peace building in order to, first, identify critical differences in how they conceptualize and operationalize their mandate and, second, map areas of potential concern. (Barnett et al. 2006, 36)

See Table 1 with definitions, almost all include peace in the definition.

20. Vabulas and Snidal (2013) define an “Informal IGO”: “We thus define an IIGO as: 1. An explicitly shared expectation—rather than a formalized agreement—about purpose 2. With explicitly associated state “members” who 3. Participate in regular meetings but have no independent secretariat or
other significant institutionalization such as a headquarters and/or permanent staff.”

A core guideline in chapter 3 is to get a list of the defining features. This is closely related to coding rules for generating data sets. Discuss this list and how you might modify it to make it clear more complete. For example, a key feature is no permanent or independent secretariat. This should probably be an independent dimension. Later in the discussion it is clear that in fact there is usually some organizational basis for these informal organizations so that the key feature is in fact that it is not an independent secretariat. Also, one might want to explicitly require that there is no legally binding treaty that constitutes the organization, rather than a mere parenthetical remark. What is the implicit concept structure?

Answer:

Probably as with most data sets the concept structure is the necessary and sufficient condition one. Here is a reformulation:

(a) No legally-binding treaty or agreement that constitutes the organization.
(b) No independent secretariat or headquarters.
(c) Regular meetings.
(d) Shared expectation of purpose
(e) States form the core membership
(f) Nonstate actors can be involved (this is important since the COW IGO list does not allow nonstate actors to be members of IGOs)
(g) At least two members.

21. Legalism is an important concept in international international relations. What is the opposite of legalism? For McCall Smith (2000) according to table 1 it is diplomacy? Discuss this negative pole choice. What about “not-legalistic”?

22. It is possible to develop a definition or conceptualization based on cases that individuals or research communities identify as being exemplars of that concept. For example, one might develop a concept of populism based on countries, leaders or parties that are called populist in the literature. “Third, it uses the concept of political style to discern inductively the features of populism as a political style. This is done by examining twenty-eight cases of leaders from across the globe who are generally accepted as populists (that is, labelled as populist by at least six authors within the literature on populism), and identifying what links them in terms of political style.” (Moffitt 2016, 27

Discuss how this might lead to “minimal conceptualizations.” Which of Mill’s methods is one basically using here?

Answer:

Basically one is using Mill’s method of agreement to find the common features of all cases labeled as populist. Then implicitly one will probably assume that those necessary features are also jointly sufficient.
23. Discuss the use of the logical OR the following conceptualization of “extreme group.” For example, consider extreme nonviolent action, e.g., Gandi.

So persons or movements may be called extreme because their views are far out of the mainstream on some issue, or because they use violence to further their goals, or because they are rigid and intolerant of other points of view. A group can be extremist if it has only one of these features. Some movements – Al Qaeda is a good example – have all of them. Indeed, this is not surprising, because, as we will show, the latter two characteristics, the predilection for violence and the tendency toward rigidity and intolerance, can be derived from the first characteristic. (Wintrobe 2006, 6, my emphasis)

What about “extreme” left nonviolent campaigns? Is violence a requirement for an extreme group, or is ideology sufficient?

24. The Polity measure of democracy has a code of –77 for cases where there is no effective government, e.g., complete state failure. Some conceive of this as a problem of missing data and hence want to estimate the value. For example, sometimes these are given a value of 0 on the polity –10 to 10 scale. Discuss whether coding regime type requires that there actually be a government.

25. Call (2010) discusses the literature on the important and large concept of “failed state.” He argues that there are three large secondary-level dimensions to this concept conceptualized by three “gaps”: (1) capacity gap, (2) security gap, and (3) legitimacy gap. He shows that these gaps are not very correlated: “Moreover, the three security gaps produce remarkably different rankings of countries of the world.” (p. 309). He sees this as a negative finding for the concept of a failed state. Discuss the underlying correlational view of concepts that he is drawing on. How could this non-correlation be a positive thing from a definitional point of view?

Answer:

He works on the correlational or indicator view that all the dimensions–indicators of a concept should be correlated. A definitional view works on the notion that each definitional part should be different. Hence it is a positive feature that these dimensions are not in fact correlated, because they are in fact different dimensions of the failed state concept.

26. Fordham (2011) logs all independent and dependent variables. The key independent variable is size of country which is a major cause of major power status, the dependent variable. Explore what logging does when the data include both very small and very large countries.

27. The new “Democracy Barometer” (Bühlmann, M. et al. 2012) is a large project whose goal is to conceptualize and measure the “quality of democracy.” It is a very complicated concept with at least 4 levels. It is an interesting example of a hybrid concept structure. Discuss where necessary condition ideas are used in the aggregation and when family resemblance is used.

28. One potentially large issue is what might be called “conceptual endogeneity.” This is where a secondary-level dimension is part of the dependent variable
and also considered as an independent variable. There has been an explosion of studies of elections in hybrid democratic regimes. Should these elections be considered a cause of democratization? Does this conflict with the fact that freer and more competitive elections will change the democracy score of the country by definition? For example, “Scholars also disagree about the role of elections in promoting democratic change. (p. 12) While a familiar indicator of democratic progress, we must remember, electoral turnover does not always support democratic development. (p. 13) These issues lead to a final and more fundamental concern about attributing too much democratic influence to elections” (Bunce and Wolchik, 2011, 14). See also Gandhi and Lust-Oskar (2009) who discuss at length the causal effects of elections on authoritarianism.

29. Recently statisticians have become much more concerned with problems of “unit homogeneity.” Here is Henry Brady defining the idea:

We shall make the transformation of \( Y_B(1,0) \) into \( Y_A(0,0) \) in two steps which are depicted on Table 10.9. If \( A \) and \( B \) are identical and \( Z_A \) and \( Z_B \) [\( Z \) is the treatment] are identical as well (footnote: By saying that \( Z_A \) and \( Z_B \) have to be comparable, we mean that \( Z_A = 0 \) and \( Z_B = 0 \) are the same thing and \( Z_A = 1 \) and \( Z_B = 1 \) are the same thing.) (although we haven t indicated how this might be brought about yet) it might be reasonable to suppose that: \( Y_B(1,0) = Y_A(0,1) \) [Identicality of units and treatment or Unit Homogeneity]. That is, \( A \) and \( B \) are mirror images of one another so that the impact of \( Z_A = 1 \) and \( Z = 0 \) on \( B \) is the same as the impact of \( Z = 0 \) and \( Z = 1 \) on \( A \). (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.
Answer:
Holland (1986) says “unit homogeneity” means that units are prepared carefully “so that they ‘look’ identical in all relevant aspects” (Holland 1986, 948). To be “identical” means that the concept must produce identical units, \( A \) and \( B \) in the Brady analysis. If a study looks at all “states” then the conceptualization of state means that there are no casually important differences in “stateness.” Treatments, \( Z \) in the Brady analysis, also are assumed to be homogeneous.

30. Use the following quote to discuss the importance of predictive power in terms of evaluating a measurement model or concept. Contrast this with an approach which stresses the semantics and content of the test items.

The g factor is an artifact of linear correlation analysis. A theorem of Suppes and Zanotti (1981) informs us that for any vector of test scores from an achievement test, it is possible to construct a scalar latent factor such that, conditional on the factor, test scores are independent. The g factor exists for any vector of binary, finite-valued, or countably valued random variables. The g of conventional psychometrics is a product of mathematical conventions in factor analysis. A g also exists to account for correlations among test scores. That is a mathematical theorem of no behavioral consequence for psychometrics or for finance, another field addicted to factor models. The value of g in predicting behavior is the real test of its importance. There is much evidence that it has predictive power. (Heckman 1995, 1105)
31. Cheibub et al. (2010) is a review and updating of the Alvarez et al. and Przeworski et al. well-known measure of democracy. They strongly critique the polity measure and the Freedom House for their aggregation schemes and the large number of ways the scores on the secondary-level dimensions produce the same value on the basic level:

Regarding aggregation, for each of the ten categories in the political rights checklist and the 15 categories of the civil liberties checklist, coders assign ratings from zero to four and the points are added so that a country can obtain a maximum score of 40 in political rights and 60 in civil rights. With five alternatives for each of ten and 15 categories, there are \( 5^{10} = 9,765,625 \) possible ways to obtain a sum of scores between zero and 40 in political rights, and \( 5^{15} = 30,517,578,125 \) possible ways to obtain a sum of scores between zero and 60 in civil liberties. In all of these cases, the aggregation rules are arbitrary. (Cheibub et al. 2010, 75)

While their dichotomous coding scheme with three “categories” does not produce the same huge option of possibilities, the same issue arises in their coding scheme. Explain how the aggregation issue arises in their measure. Are their aggregation rules “arbitrary?”

Answer:

While Cheibub et al. do not think of their coding rules as involving an aggregation procedure, they certainly do. They use the necessary and sufficient condition one. While there is only one way to be coded as a democracy, i.e., one on all the rules, there are multiple ways to be coded as a zero. Assuming four rules there are for example, \((0,0,1,0), (1,1,0,0), (0,0,0,0)\), etc. ways to be a dictatorship. In short there are \(2^4 - 1 = 15\) ways to be a “dictatorship.”

32. Alvarez, Przeworski, Cheibub, Vreeland, and Gandhi who are all active working together on the concept and data on regime type have a clear preference for the term “dictatorship” as the opposite pole to democracy. In contrast, scholars using the polity data prefer the term “autocracy” or “authoritarian.” Is there anything of theoretical or conceptual importance at stake in this terminological difference?

33. One of the best ways to find out about concepts is via the codebook for datasets. Sometimes reading the codebook provides some surprises about the concept (and hence the data) that most users are probably not aware of. An example of this phenomenon are the GTD terrorism data (CETIS 2007), the standard dataset for the study of terrorism. If one reads the GTD codebook the problematic nature of the concept is acknowledged in the introduction, but almost all the codebook is about the data. To find out what the GTD concept of terrorism actually is one must read an appendix! As an exercise I recommend that one read the codebook and outline what you think that concept of terrorism used is. I think you will then be surprised by the content of the appendix.

34. In the international relations there is a growing literature on what is called the “capitalist” peace (e.g., see the special issue of *International Interactions* 2010(2)). This literature usually has two goals, (1) so show that capitalist
countries are less likely to have militarized conflicts, and (2) capitalism is more important than democracy in reducing conflict. Discuss what is the concept of “capitalism” used in these various studies. Then discuss what are the actual quantitative indicators used. How do they match up? Could you think of better indicators or tests? When left-wing, e.g., socialist or communist, critics discussed “capitalism” how was their meaning of capitalism different from the capitalist peace literature? Discuss McDonald (2010) and the various forms of the concept of capitalism from the beginning of the article and how they are implemented the quantitative analyses at the end.

Answer:

If you look at the Gartzke, Mousseau, and McDonald pieces (in the *International Interactions* special issue) “capitalism” is discussed in terms of (1) economic development, (2) (free) markets, (3) contract intensive economies, (4) integration into the global economy, (5) financial openness, (6) degree of property owned by government, (7) free trade and (8) government nontax revenues.

Most of these are either too narrow (e.g., government nontax revenues, financial openness) or different (e.g., integration into world economy which is not directly about the domestic economy, e.g., oil-exporting states).

One might reasonably define capitalism as a “market-based” economy. There seem to be at least two reasonable ways one might go about operationalizing this. First, one might contrast communist/socialist countries as a dummy variable against all other types. Another way to think about this is to consider the amount of economic activity in a country that goes through the government, which might be operationalized as government expenditures as a percentage of GDP (ideally this would be government at all levels).

McDonald (2010) is a good example of the conceptual and operational problems of this literature. He starts with general claims about the “capitalist” peace. Then he moves to a claim that capitalist economies have less “public property.” One might think that public property is property owned by the government. This would then include land, highways, military bases, buildings, etc., not to mention financial capital assets (both at home and abroad). In practice, i.e., quantitative analyses, he uses nontax revenue of the government. This variable seems to be quite distant from the starting point of capitalism.

35. Draw the three-level figure that corresponds to Levitsky and Way’s (2010 Appendix I) concept of “competitive authoritarianism.”

36. Concepts with adjectives, e.g., competitive authoritarianism, often confuse to two quite different issues. Sometimes they mean regions in the middle of a continuum as is often the case with democracy. They can also two separate concepts that are put together, which is the classic use, e.g., “pet fish.” He and Warren (2011) illustrate this potential confusion. The concept they propose is “deliberative authoritarianism” which they discuss in the context of hybrid regimes. However, they argue that “deliberation” is separate from “democracy”: “If deliberation and democracy are distinct in theory – the one a kind of communication, the other a distribution of powers to decide – they have
often been distinct in practice as well.... These observations can be ideal typed. If deliberation is a phenomenon different in kind from democracy, then (in theory) it might combine with non-democratic (authoritarian) distributions of power. We illustrate the ideal types in Table 1, where the terms “authoritarian” and “democratic” refer to the relative dispersion of means of empowerment (dispersion, by implication, provides more opportunities for the affected to exercise power), while communication can vary from “instrumental” to “strategic” and “deliberative.”. The combinations produce five familiar types, and one unfamiliar type, deliberative authoritarianism” (He and Warren 2011, 273). Analyze the difference between middle of the continuum and the separate concept approaches in dealing with concepts with adjectives.

37. While scholars often use the “ideal type” to think about concepts there is virtually no methodological literature on building ideal types. He and Warren (2011) illustrate some of the issues of trying to do ideal type concepts with two dimensions: “we then develop the ideal type of deliberative authoritarianism—a regime style that makes frequent use of authoritarian deliberation. In developing this ideal type, we depart from much of the literature on hybrid regimes (He and Warren 2011, 270). If one thinks about about “deliberation” as a separate concept (see the question above) consider their view of the extreme positive pole of “deliberative authoritarianism” (e.g., Table 1, 273). Is it is really the extreme for deliberation? (Hint: think about the positive extreme for deliberation for democracy).

Answer:

If one thinks of deliberation as a concept in its own right the ideal type, i.e., positive pole, should continue for at least one more column. Certainly good deliberative democracy includes more deliberation than deliberation in authoritarian regimes like China. Since the authors focus on ideal type which is a combination of authoritarianism and deliberation, they naturally want to stop the scale at the end of the authoritarian row, without considering deliberation as a concept in its own right. Hence it would be less natural to have an empty cell at the end of the second row which is high quality deliberation which one would naturally have to include if one wanted to include the positive pole of deliberation for democracy.

38. Adjectives are often used to distinguish among subtypes. Discuss Rathbun’s (2011) use of “qualitative” to distinguish between different types of multilateralism; for example, “RH1: State preferences for qualitative multilateralism will be accompanied by preferences for organizations with a smaller number of members.” Discuss the other type which is “quantitative” multilateralism.

39. Gandhi and Lust-Oskar (2009) review the literature on “authoritarian elections.” Discuss this interesting example of a concepts and adjectives. Fair or competitive elections are of course part of the definition of democracy. Presumably their absence is a requirement for an authoritarian regime. Do such cases fall into the gray zone between democracy and authoritarianism?
40. There is an extensive cognitive psychology literature on the use of prototypes in categorization (see Murphy (2002) for a nice review). The basic idea is that people (not social scientists) often think about concepts via prototypes. For example, the concept of a “bird” is based on prototypical birds, e.g., sparrows. It is less clear whether social science concepts should or are based on prototype models of concepts. Shadish, Cook, and Campbell (2002; essentially the 3rd edition of the classic Cook and Campbell) provides an example where concepts (aka “constructs”) are discussed using the prototype idea. Discuss the use of prototypes in social science. Are they different from ideal types? Contrast prototypes with concepts based on defining, essential features (a common contrast in the cognitive psychology literature).

41. Colaresi et al. (2007) in their prominent work on the concept of “strategic rivalries” argue (chapter 6) that there are two types of rivalry “positional” (i.e., roughly competition in the regional or global system) or “spatial” (i.e., over territory). Table 6.1 argues that there is a continuum from pure positional to pure spatial, i.e., the “spatial–positional continuum.” Discuss whether one continuum with the two combined would be better than two continuums one for each type of rivalry going from presence to absence on each continuum.

42. Abbott, Keohane, Moravcsik, Slaughter, and Snidal (2000) published a very influential article on “The concept of legalization.” This article could serve as a good example for the discussion of concepts. (1) Draw a diagram that represents the three-level structure of their concept. The three secondary-level dimensions are quite clear and the indicators are given in various tables. (2) What is the structure that they implicitly use to connect secondary-level dimensions and indicators? (3) Thinking about the negation of a concept is important; what about the negations given in figure 1 as well as table 1? (4) “Delegation” (table 4) might be separated into two, or even three, separate dimensions? Are “dispute resolution,” “rule-making” and “implementation” part of the higher level concept of “delegation?” (5) In the delegation dimension the concept of “binding” appears frequently, but is also central to the “obligation” dimension. Is this a problem?

Answer:

It seems like the implicit structuring used at both levels is the family resemblance one. There is not much sense that any of these are necessary. There are potentially other negations that could be used in figure 1, such as “nonbinding” for obligation. For precision it is interesting that the positive pole uses the term “rule” while the negative pole uses the term “norm.” Delegation is problematic because it includes both implementation and arbitration. In terms of the delegation dimension the concept of “binding” appears frequently, but it is also central to the “obligation” dimension. Is this a problem?

Also missing is a potential dimension about who makes international rules, e.g., states, United Nations, IGOs, etc.
Gates et al. (2006) reformulate a new three-dimensional concept of democracy using polity and Vanhanen. They use a cube to think about their concept of democracy, anocracy, and autocracy. In their statistical analyses they use a dichotomous coding of democracy, autocracy, and anocracy. How could they have used continuous [0,1] variables for each of these?

The ideal types also include polities that are close to the corners. As the cube in Figure 1 defines a space, it is possible to examine the distances within this space. In order to classify a regime as either Ideal or Inconsistent, the distance from the point given by the polity’s coordinates to the eight corners and the midpoint of the cube is calculated. A regime is defined as Democratic or Autocratic if it is closer to either of the ideal type corners of the cube than to the other corners or the midpoint. The Democratic ideal type will hence include observations that are closer to the corner [1,1,1]. Since it is the distance from the democratic corner that defines the ideal type, scores close to 1 on one of the dimensions to some extent offset low scores on the other dimensions. The autocratic ideal type includes all polities that are closer to the autocratic corner than any of the other reference points. All polities that are not coded as Autocratic or Democratic are coded as Inconsistent. For the 1800–2000 and 1900–2000 periods the respective distributions of the three types of polities were as follows: Autocratic (43%, 39%), Democratic (14%, 17%), and Inconsistent (43%, 44%). (Gates et al. 2006, 898, note that they conceptualize the three dimensions this with the cube).

Answer:

Using the cube they can calculate three distances for each data point in the cube: (1) the distance to (1,1,1) which would be the democracy score, (2) the distance to (0,0,0) which would be the autocracy score, and (3) distance to (.5,.5,.5), this distance will have to be normalized to be in the [0,1] range.

Volgy et al. (2008) is a nice example of an implicit three-level concept structure. Draw a figure of the concept. What are the dimensions and the structural principles used at each level?

In literature on the concept of ethnicity, perhaps the biggest problem is defining the secondary-level dimensions that are used to conceptualize ethnicity. Some commonly used dimensions are religion, language, and race. Discuss what the complete list should be. Should it include factors like class or ideology?

Answer:

Here is one answer: “What is ethnic identity? Since the publication of Horowitz’s (1985) Ethnic Groups in Conflict, there has been a convergence among comparative political scientists on which identities we classify as ethnic. For Horowitz, ethnicity is an umbrella concept that “easily embraces groups differentiated by color, language, and religion; it covers ‘tribes,’ ‘races,’ ‘nationalities,’ and castes” (Horowitz 1985, 53). Much of the recent theoretical literature on ethnic politics explicitly follows this umbrella classification (e.g., Varshney 2002, Chandra 2004, Htun 2004, Wilkinson 2004, Posner 2005). Even more importantly, the four principal datasets on ethnic groups that constitute
the foundation for cross-national empirical studies of the effect of ethnic identity in comparative politics – the Atlas Narodov Mira (Bruk and Apenchenko 1964), Alesina et al.’s (2003) dataset on ethnic groups in 190 countries, a comparable count of ethnic groups in 160 countries (Fearon 2003), and the Minorities at Risk project (http://www.cidcm.umd.edu/inscr/mar/) – also generally employ this umbrella classification. Only some quibbles remain on the margin about whether castes should be excluded (e.g., Fearon 2003) or retained (e.g., Bruk & Apenchenko 1964, Varshney 2002, Chandra 2004, Htun A 2004, Wilkinson 2004, Posner 2005) and whether region and clan should be included.” (Chandra 2008, 397–98)

46. For German speakers. Opp (2005) is a very rare methods textbook that devoted significant attention to concepts – most textbook only look at measurement; it is also a rare textbook that devotes a whole chapter to teaching logic – most research design and methodology textbooks do not discuss this at all. In Opp’s discussion of concepts (chapter 4) he distinguishes between (1) “Analytische operationalisierung” and (2) “empirische operationalisierung.” Discuss the extent to which (2) corresponds to the standard latent variable model of concepts and (1) corresponds to what the concept book calls ontological concepts.

47. Concept asymmetry appears in various contexts. In survey research there are often major differences depending on the how the question is worded. For example, a classic article by Rugg in 1941 found differences of up to 20 percent of two questions “that seem logically equivalent.” (Holleman 1999, 210), “Should the USA forbid public speeches against democracy” versus “Should the USA permit public speeches of democracy.” One of the main hypotheses in the literature is that there are different scales for the two responses. Holleman expresses this clearly: “The main issue is whether questions worded differently, although conceptually equivalent, measure similar attitudes (a test of the different attitudes hypothesis) and, if so, whether the similar attitudes are expressed on similar scales (the different scales hypothesis)” (Holleman 1999, 213). She finds empirical support for the two scales hypothesis.

Using semantic transformations discuss whether these the response of yes to one question is the equivalent to the answer of no on the other. Does your answer suggest that perhaps fuzzy logic scaling or concepts would be appropriate here.

Answer:

Concept asymmetry means that “poor” does not necessarily equal “not wealthy.” Hence it might be the case that “not forbidden” is not necessarily equivalent to “permitted.” The fact that empirically two scales are found might suggest that a fuzzy logic approach might be better.

48. The GTD (global terrorism database) database (CETIS 2007) uses a hybrid concept of terrorism involving 3 necessary conditions and then 2 of 3 other conditions for an incident to be considered terrorist. (1) Discuss whether various kinds of Mafia violence could be considered terrorism under this concept. (2) Discuss if nonviolent demonstrations would count. (3) One important aspect of concept construction is considering the negative pole or phenomenon
that do not fall under the concept. The coding manual implies that insurgency is different from terrorism, e.g., “Appendix A: guidelines for differentiating between insurgency and terrorism,” but it does not define or conceptualize “insurgency”; how might this be an issue in studying terrorism and, say, civil war? (4) Discuss the secondary-level dimension “violates international humanitarian law of war.”

Answer:

Mafia activities could count because while they fail on one of the 2 out of 3 conditions (economic motives only) they can fulfill the other two conditions. Also there are coding categories for “criminal groups” (p. 32).

“Pacifist” / Anti-War is one possible category of terrorist group.

In addition to appendix A there is a variable “alternative designation” for incidents which are not considered terrorist, which include categories “Insurgency or guerilla action,” and “Internecine conflict action.” It is not at all clear what the relationship is between terrorism, insurgency, state breakdown, and civil war.

49. Democracy is a key variable in the theories about civil war. Vreeland (2008) discusses how the polity coding of anocracy is problematic for this use. Explain why these problems make sense given the original goals of the polity concept (Gurr 1974).

Answer:

Gurr (1974) was interested in the stability of regimes, and by consequence regime change or viability. A country in civil war can be considered and coded as nonstable.

50. The concept of “civilian power” engages the nature of foreign policy in countries like the EU and Japan (e.g., see Mauell et al. 2006; Smith 2006). Discuss how the adjective “civilian” works in this concept. For example, how does it modify the concept of power? Also discuss the role of the negative pole in this concept.

Answer:

Civilian is clearly discussed as the opposite of “military.” Note that many discussions of power would include nonmilitary forms of power.

51. Some key concepts are basically defined in a negative way (e.g., see exercise 50 on “civilian” power). A good example of the issues involved is the debate over the “democratic deficit” in the EU (see Moravcsik 2002; Follesdal and Hixs 2005). One seems to be taking away dimensions from the concept of democracy. In practice “democratic deficit” is often defined without a clear concept of democracy. Discuss the (implicit) concept of democracy used by various authors. Is it defendable or common as a stand-alone concept of democracy?

52. There is a very large literature in American politics on the concept and measurement of racism, and it is a topic that arouses much controversy: “Is white opposition to policies driven by racial prejudice or is it grounded in race-blind
ideological principles?” (Feldman and Huddy 2005, 168; this is first sentence of the article). There are many aspects of this debate of interest regarding concepts. (1) One interesting aspect is the variety of terminology, such as “overt prejudice,” “new racism,” “racial resentment,” “old-fashioned racism,” and “symbolic racism.” In addition there are related concepts such as “prejudice” and “discrimination.” In particular the use of adjectives is of interest: discuss the issues and implications of the use of the concept “symbolic racism” (e.g., Tarman and Sears 2005) as a particular use of adjectives. Notice as well the relevance of the negation guideline given in the use of adjectives. (2) The negation guideline suggests that much can be gained by focusing on the opposite concept. Much of the debate revolves around the concept of a “principled conservatism” which argues against policies such as affirmative action but which is not racist (Sniderman et al. 1996). One could focus on that concept instead of directly on racism. For example, if one is principled then those principles should apply across groups by gender and sexual orientation and not just race. (3) It is important to think about causal relationships in concepts. One might suggest that, at least in the case of the south, principled conservatism is a consequence of racism, given that race was a central component of Southern culture and politics for centuries. See Quillian (2006) for a nice survey of the political science, sociology, and psychology literatures.

53. One serious issue with dichotomous data is the potential for serious heterogeneity in the zero or one category. This is particularly likely to be a problem in zero category. The issue of heterogeneity arises almost by definition with nonordered, categorical data. For example, quantitative studies of international conflict that look at the outcome dispute variables that have anywhere from 3 to 9 categories. Often these can be reduced to settlement/compromise, win, draw. If the focus is, say, settlement or compromise then there is a tendency to code that as one and lump win and draw together as zero. Discuss the merits of this procedure. What would be alternatives?

Answer:
Lumping all these together in the zero seems particularly dubious when the causal mechanisms producing these zeros are likely to be quite different.

Within the context of the conflict literature and the militarized disputes dataset, it might make sense to just include draw/stalemate as the contrasting case since this constitutes often 70+ % of the zero cases (e.g., Goertz et al. 2005). This choice could well be brought into the research design and theory if the research were interested in making specific causal contrasts.

54. Typically, one attacks theoretically and empirically concept issues from the positive pole first. However, sometimes it might make sense to really focus on the negative. For example, the concept of “judicial independence” is key in the literature on comparative judicial systems and democratization. Evaluate Larkins (1996) proposal to focus on the dimensions of judicial dependence as the best way to attack the problem. Also, discuss whether the necessary and sufficient condition or family resemblance approach makes more sense.

Answer:
The answer to the second question about concept structure informs the response to the first. Typically, there are many ways in which governments can make judiciaries dependent. Usually you only need one to remove judicial independence; i.e., using the contrapositive of the necessary and sufficient condition structure means that the absence of all the means of judicial dependence is sufficient for judicial independence.

55. Sometimes the “positive” pole of interest is itself defined in negative terms. The concept of an “antisystem party” has a long history in comparative politics, notably Sartori (1976): “a party that would change, if it could, not the government, but the system of government” (cited in Capoccia 2002, 18).

Discuss the problems of defining what the “system” is that forms the positive pole. Frequently, “system” means “democracy.” Capoccia (2005, 2002) discusses two types of antisystem parties one “ideological” and the other “relational.” Relational means distant and extreme parties in ideological space. Discuss the relationship between “anti” in terms of the ideological space of the parties versus the characteristics of the political system itself. How would a family resemblance perspective on democracy be more problematic than a necessary and sufficient condition one in looking at system as democracy?

Answer:

See Capoccia (2002) for an extended discussion. (1) Inherently it is going to be hard to define what the “system” is. (2) One party could be distant ideologically from the center and not be opposed to the system of government itself. (3) Using the necessary and sufficient condition view of democracy makes it easier to think about antisystem parties since if they are opposed to any secondary-level dimension of democracy then it is an antisystem party.

56. In many debates the role of the negative pole plays a key role. One important point in the debate between Duffield (2003) and Koremenos et al. (2003) is the role of “centralization” in the study of international institutions. What is the opposite pole? What would be the opposite pole in the study of domestic institutions?

Answer:

Duffield and Koremenos et al. assume that the opposite pole of centralization is “delegation;” for example, “Duffield makes a good point in stating that greater focus on delegation may provide a way to further refine the concept of centralization, because delegation and centralization are indeed closely related” (Koremenos et al. 2003, 434).

It is likely that within the literature on domestic institutions design (e.g., constitutions), the opposite pole would be “federalism.”

57. A key issue when creating concepts is their eventual theoretical and policy usage. Of great policy importance today is the concept of “indigenous people” (see Corntassel 2003 for an excellent review of different concepts). Clearly, one policy goal is protect indigenous peoples. Discuss how this influences one’s definition. Notice that many important international organizations such as the UN, ILO, and the World Bank have developed definitions that have practical and legal import.
58. Describe the three-level structure of Dasgupta’s concept of human well-being or destitution (Dasgupta 1990; Dasgupta and Weale 1992). What is its structure? What are the relative weights attached at the indicator and secondary levels?

Answer:

The structure is very similar to the one used by Schmitter for corporatism, since it uses the sum of the ranks. The secondary-level level variables such as “health,” “education,” “wealth” will not be equally weighted in general unless they have the same number of indicators. More generally, is it reasonable to ranking literacy equally with per capita income?

59. Milner and Kubota (2005) discuss a number of important issues regarding concepts and measurement. (1) Discuss their measure of “open/closed trade regimes.” What is the strategy used for building the measure? (2) Contrast this with their concept(s) of “economic crisis.” What is the strategy used for building these measures? Would it be better to have one or two economic crisis variables? What impact or role does the information that the two crisis variables are not correlated (i.e., \( r = .01 \)) have on your argument?

Answer:

See Milner and Kubota (2005, 122–24). For the open/closed trade regime variable they use a family resemblance strategy by applying the m-of-n rule: if one of the four secondary-level factors is present then the variable is coded one. Milner and Kubota probably separated out the two economic crisis variables because of their low correlation.

60. Examine Doner et al.’s concept of the “developmental state” (Doner et al. 2005, Table 1) and draw a figure of their three-level concept.

61. Przeworski and his colleagues have proposed an influential concept and measure of democracy:

Operationally, a regime was classified [dichotomously] as a democracy if none of the four rules listed below applied . . . .

Rule 1: Executive selection: the chief executive is not elected.
Rule 2: Legislative selection: the legislature is not elected.
Rule 3: Party: there is no more than one party. Specifically, this rule applies if (1) there were no parties, or (2) there was only one party, or (3) the current tenure in office ended up in the establishment of a nonparty or one-party rule, or (4) the incumbents unconstitutionally closed the legislature and rewrote the rules in their favor.

Rule 4: Type II error: a regime passes the previous three rules, the incumbents held office in the immediate past by virtue of elections for more than two terms or without being elected, and until today or the time when they were overthrown they have not lost an election. (Przeworski and Limongi, 1997, 178; see also Alvarez et al. 1996, Przeworski et al. 2000, chapter 1)

Rule 4 (called “Alternation” in Przeworski et al. 2000) only applies if “a regime passes the previous three rules.” Would anything change in the final dichotomous democracy codings if we just considered this a fourth rule? Explain.
Answer:
Nothing would change in the final dichotomous democracy coding. Because failure on any rule, including rule 4, eliminates the country as a democracy nothing is gained (except perhaps time) in the overall democracy coding by applying rule 4 only when a country passes rules 1–3.

62. An important concept in the study of international conflict is that of the “similarity” of regime types. One central hypothesis is that international conflict is more likely between dissimilar regimes. One can compare two countries on their level of democracy in at least two ways with a multilevel democracy concept: (1) via the similarity of the secondary-level dimensions or (2) via the similarity of the basic-level measure. Discuss if one procedure is more valid than the other.

Answer:
Similarity comparisons at the level of secondary-level dimensions with necessary and sufficient condition structures are not valid because the claim is that the absence of one dimension makes the a country a nondemocracy, independent of what is happening with other secondary-level dimensions. In general, similarity should be addressed at the basic level not using secondary-level dimensions because using secondary-level dimensions ignores how the concept is structured (see Bueno de Mesquita et al. 2005 for an extensive justification for examining the secondary-level dimensions individually).

63. One important guideline for concept-building is to think about the negative pole. Esping-Anderson (1990; 1999) focused attention on three “worlds” of welfare capitalism, which has informed 15 years of research. This view of the welfare state has three dimensions (1) socialist, (2) liberal, and (3) conservative. Evaluate Hicks and Kenworthy’s argument that instead of three dimensions there are two with (2) being the negative pole of (1):

we can characterize and differentiate welfare states in terms of the “progressive liberalism” and “traditional conservatism” of their policies and programs. The first of these two dimensions is fairly novel. It rearranges Esping-Andersen’s separate social democratic and liberal dimensions into two poles of a single dimension. (Hicks and Kenworthy 2003, 51).

64. There exists numerous measures of corporatism (see Siaroff 1999 or Kenworthy 2003 for surveys). Examine Cameron (1984) who uses a hybrid measure. What is the concept model that leads to hybrid measures?

65. Explicitly contrast Schmitter’s concept of corporatism with his concept of pluralism. To what extent is pluralism strictly the negation of corporatism?

66. One important issue is redundancy across indicators or dimensions. The CIRI measure and state on human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a given year; their final score is the sum of the scores of
the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How could you reformulate the structure to deal with this critique?

Answer:
Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of either. With the new scoring system it would also have a score of eight.

The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

67. Often the issue of dichotomous versus continuous lies hidden in conceptual discussions. Vu’s discussion of the analysis of the state makes a distinction between “variables” and “attributes.” How does the issue of dichotomous dimensions seem to lie in the distinction between the two?

Weber defines it [the state] as “a human community that (successfully) claims the monopoly of the legitimate use of physical force within a given territory.” Later students of the state have modified this definition somewhat but the essential elements remain the same. Tilly defines states as “coercion-wielding organizations that are distinct from households and kinship groups and exercise clear priority in some respects over all other organizations within substantial territories.” For this definition he means to include city-states, empires, and theocracies and to exclude tribes, lineages, firms, and churches. “National states” are a subset of states that are relatively powerful, centralized, and differentiated sovereign organizations. Levi offers a more restrictive definition than Tilly: “a state is a complex apparatus of centralized and institutionalized power that concentrates violence, establishes property rights, and regulates society within a given territory while being formally recognized as a state by international forums.” These later definitions avoid such terms as “monopoly” and “legitimate,” perhaps because these are seen as variables, not attributes. (Vu 2010, 164–65; Levi 1982, 40 makes the same distinction between “variables” and “elements of a definition.”)

Answer:
“Variables” are continuous dimensions, while Vu seems to assume that “attributes” and dichotomous. One could certainly think that all the attributes of a state discussed in this quote are continuous.

68. The topic of concepts with adjectives arises with great force in thinking about the EU as a state. Scholars talk about Westphalian, regulatory, post-modern, regional, international, e.g., (Caporaso 1996). Analyze these various concepts in terms of their coverage of cases the degree to which they are subsets of a Weberian concept of a state. Should one think of the EU in terms of a gray zone? Discuss how the conceptualization of the EU as a state influences your view of the “stateness” of the members.
One potential solution is to consider the EU and its members as “quasi-states.” Basically, in some areas such as trade the EU acts like a sovereign state. In other areas the members retain sovereign control. Hence both the EU and its members are in the gray zone.

69. For some important concepts like “democracy” and “peace” scholars have attached the adjective “stable.” In early work (e.g., 1960s) this was common for democracy, but the “stable” adjective was dropped until people began to work on democratic transitions. In contrast, it is still very commonly used to denote a high level of peace (e.g., Kacowicz and Bar-Siman-Tov 2000). What is a potential problem with using stable as an adjective?

Answer:
In the early literature on democracy and the literature on peace, often stable often means “high level.” For many “stable peace” and “positive peace” mean the same thing. It useful to distinguish between the level of a variable and how stable the phenomenon is at any given level. This is not a problem in the literature on democratic transitions since the emphasis is really on stability. See Goertz et al. (2016) for a survey of the literature on the concept of peace.

70. Terrorism is a big policy as well as scientific issue. There are a large variety of different adjectives attached to the concept “terrorism” including international, transnational, state-sponsored, state, domestic, political, revolutionary. In making sense of this one must consider how the adjective is modifying the terrorism concept. Related to this is that many of the adjectives involve the negative pole of a dichotomous concept, e.g., state versus nonstate. Analyze the conceptual issues involved in defining terrorism, with or without adjectives. Discuss how the decisions one makes can have a significant impact on case selection. For a survey of different concepts of terrorism see for example Crenshaw (1972), Gibbs (1989), Held (2003).

71. Sometimes the ideal point is not at the end of the scale but in the middle somewhere. Analyze the concept of “judicial (in)dependence” as an example of this. See Larkins (1996) for a review of the literature.

Answer:
Typically, one wants the courts to be not completely unresponsive to elected officials. At the same time the courts cannot be completely submissive to these officials. Hence the ideal point is some degree of independence but also the possibility that judges can be removed by elected officials. Also, there are typically legislative mechanisms for overriding some judicial decisions.

The five essays that follow are part of a collaborative effort, launched at a conference at Stanford University, to elaborate and refine the concept of
democratic quality and to apply it to a series of six paired country comparisons. We asked each author to discuss a particular dimension of the quality of democracy such as freedom, the rule of law, vertical accountability, responsiveness, and equality (our own list, and by no means exhaustive). We wanted each author to explain how the dimension in question relates to other dimensions in our framework, to suggest possible indicators for measuring the dimension, to identify ways in which this element of democratic quality is subverted in the real world, and to offer (where possible) policy recommendations. Our full framework features eight dimensions: the five outlined above, plus participation, competition, and horizontal accountability. Other dimensions might include transparency and the effectiveness of representation. The different aspects of democratic quality overlap, however, and we choose to treat these latter two as elements of our principal dimensions. We attempt here to identify some of the ways in which the different elements of democracy not only overlap, but also depend upon one another, forming a system in which improvement along one dimension (such as participation) can have beneficial effects along others (such as equality and accountability). At the same time, however, there can be trade-offs between the different dimensions of democratic quality, and it is impossible to maximize all of them at once. In this sense at least, every democratic country must make an inherently value-laden choice about what kind of democracy it wishes to be. Talk of a “good” or “better” democracy implies knowing what democracy is. At a minimum, democracy requires: 1) universal, adult suffrage; 2) recurring, free, competitive, and fair elections; 3) more than one serious political party; and 4) alternative sources of information. If elections are to be truly meaningful, free, and fair, there must be some degree of civil and political freedom beyond the electoral arena so that citizens can articulate and organize around their political beliefs and interests. Once a country meets these basic standards, further empirical analysis can ask how well it achieves the three main goals of an ideal democracy—political and civil freedom, popular sovereignty (control over public policies and the officials who make them), and political equality (in these rights and powers)—as well as broader standards of good governance (such as transparency, legality, and responsible rule). (Diamond and Morlino 2004, 20–21)

73. How does Sartori view the question of sufficiency in conceptualization and concept structure? See also Collier and Levitsky (1997) for the same question.

Answer:

He focuses his attention on the minimal necessary conditions and does not explicitly make claims about sufficiency. In one footnote he explicitly rejects sufficiency ideas. Collier and Levitsky make no claims about sufficiency and focus only on the minimum requirements for democracy.

74. In the example showing that extension can increase with extension with family resemblance concepts, the book used a constant “half or more” rule. In practice, it seems to be case that as the total number of possible characteristics increases that the threshold does as well. For example, one might use 2 out of 4, but when the total is 10 the rule might be 7 out of 10. Discuss the validity of this tendency. What are the implications in terms of extension?

Answer:
It seems like there is some intuition that when the total number of characteristics increases that extension is likely to increase (as demonstrated in the text). To keep the extension in line, the threshold is thus increased.

75. One part of the conventional wisdom is that measures of democracy are highly correlated. It may be the case that this high correlation masks significant disagreement, particularly in the gray zone. Bogaards finds big potential differences between concepts and measures of democracy for African cases. Contrast Bogaards's (2007) examination of African data with Treier and Jackman's (2008) discussion of errors in measurement. Where do each see the biggest problem of measurement error?

Answer:

Bogaards almost by definition looks at only gray zone cases, i.e., elections in Africa. His analysis suggests that there can be major differences between democracy measures. In contrast, Treier and Jackman argue that there is most uncertainty at the ends of the authoritarian–democracy scale; measurements are most accurate in the middle (i.e., the gray zone). Bogaards analyses imply that uncertainty and error are more likely in the middle.

76. A key assumption in multidimensional concepts is that the various means of achieving a given aggregate level are equivalent (e.g., 2+3 and 4+1 are really the same). What is one criterion for defending equivalence? Bueno de Mesquita et al. (2005) provides a very good discussion of these issues.

Answer:

One common criterion for equivalence is similar causal effect in some key hypotheses. Bueno de Mesquita et al. (2005) illustrates this.

Semantic transformations, calibration

77. Discuss and graph this semantic transformation from the EU statistics bureau:

Often, the scores are based on the percentiles of the distribution of the indicator across the countries. For example, the top 5% of the units receive a score of 100, the units between the 85th and 95th percentiles receive 80 points, the units between the 65th and the 85th percentiles receive 60 points, the units between the 35th and the 65th percentiles receive 50 points, the units between the 15th and the 35th percentiles receive 40 points, the units between the 5th and the 15th percentiles receive 20 points, and, finally, the bottom 5% of the units receive 0 points (see Table 1). This is a way to prize the most performing countries and penalize the less performing ones. (Eurostat 2005, Tools for Composite Indicators Building, 49)

78. Discuss the claim about bounded versus unbounded scales in Morris’s The measure of civilization how social development decides the fate of nations. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How a serious critique of that approach is this?
It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country’s score in each annual report, but because the maximum possible score is always 1.0, the HDI does better at charting a nation’s relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

79. A core Ideal Typology guideline is a central one. This can mean “conceptual creep” as the conceptual scale expands. Discuss the implications, causal and conceptual of such creep as well as the intension–extension implications of such creep. Why might having a clear semantic transformation be critical in this process?

As it turns out, abstract concepts can creep, too. For example, in 1960, Webster’s dictionary defined “aggression” as “an unprovoked attack or invasion,” but today that concept can include behaviors such as making insufficient eye contact or asking people where they are from (1). Many other concepts, such as abuse, bullying, mental disorder, trauma, addiction, and prejudice, have expanded of late as well (2). Some take these expansions as signs of political correctness and others as signs of social awakening. We take no position on whether these expansions are good or bad. Rather, we seek to understand what makes them happen. Why do concepts creep? (Levari et al. 2018, 465)

80. Discuss the SDG index and semantic transformations, particularly linear versus other options. Describe the semantic transformations for each of the indicators in table 7 using the transformations implicit in the red-yellow-green categorization schemes. How do they differ or not from linear transformations.

The procedure for calculating the SDG Index has four steps: (i) perform statistical tests for normality and remove extreme values from the distribution; (ii) rescale the data to ensure comparability; (iii) aggregate the indicators within and across SDGs; and (iv) conduct sensitivity and other statistical test. . . . Z-scores are the most commonly used method in constructing composite indices (OECD 2008), but we see several reasons for not applying this approach to data underlying the SDG Index and Dashboards. . . . To remove the effect of extreme values, which can skew the results of a composite index, the OECD (2008) recommends truncating the data by removing the bottom 2.5 percentiles from the distribution. We apply this approach to the lower threshold and truncate data at this level. In this way we attenuate the impact of extreme values at the bottom end of the distribution on the SDG Index scores. . . . Where possible, the thresholds are derived from the SDGs, their targets, or other official sources. All thresholds are specified in absolute terms and described in Table 7. (Sachs et al. 2016, 31)

81. The MPI (Alkire et al. 2018) defines “severely poor” as a deprivation score of 50 percent or higher. What would be a fuzzy logic way of doing this?

Answer:
“Severely” means essentially the same as “very” so one can use the strategy for adjectives discussed in the concept book.
82. Key in the massive industry of conceptualizing and measuring well-being is the use of development goals and aspirational goals to perform key scaling operations on individual dimensions. These typically indicate what the 1.0 values are on the fuzzy logic dimension. Once you achieve this threshold then you are good. Discuss this example from Women, Peace, and Security Index: “Other indicators require setting maximum values. We set aspirational maximum values of 15 years for mean years of schooling and 50 percent for parliamentary representation.” (Georgetown University’s Institute for Women, Peace and Security, PRIO 2017, 57)

83. Many datasets contain ordinal variables. The Minorities at Risk dataset illustrates this nicely. Take variable REPGENCIV and discuss a semantic transformation of it.

Repression of group civilian populations (those not engaging in violent or nonviolent political activities)
0 None reported
1 Surveillance, e.g., domestic spying, wiretapping, etc.
2 Harassment/containment, e.g., saturation of police/military presence; militarized checkpoints targeting members of groups; curfews, states of emergency
3 Nonviolent coercion, e.g., arrests, show-trials, property confiscation, exile/deportation
4 Violent coercion, short of killing, e.g., forced resettlement, torture
5 Violent coercion, killing, e.g., systematic killings, ethnic cleansing, reprisal killings
-99 No basis for judgment
(All Minorities at Risk (AMAR) Phase I Codebook August 2016, 36)

84. Discuss the semantic transformations of Vahanen’s dimensions of his democracy concept. Think carefully about threshold issues.

The selected threshold values [for democracy] of Competition (30 percent) and Participation (10 percent) are arbitrary, but I believe that they are suitable approximations for distinguishing more or less autocratic systems from political systems that have crossed the minimum threshold of democracy. Because both dimensions of democracy are assumed to be equally important, a country must cross both threshold values if it is to be classified as a democracy. It is not enough to define a threshold value of democracy solely for the ID [Index of Democracy]. In the case of the ID, I have used 5.0 index points as the minimum threshold of democracy, which is clearly higher than the ID value 3.0 produced by the minimum threshold values of Competition and Participation. Countries that have reached all three minimum threshold values (30 percent for Competition, 10 percent for Participation, and 5.0 index points for the ID) can be regarded as democracies. It should be emphasized, however, that it is also possible to define threshold values differently, by raising or lowering them. (Vahanen 2000, 257)

85. The World Bank has developed its own database of “Political Institutions” (Beck et al. 2001). This database of 108 variables includes an index/scale similar to those in concepts and measures of democracy:
The database also contains two indexes that characterize the competitiveness of elections in countries, one for executive elections and one for legislative elections. The core of the two indexes is the number of parties that could and did compete in the last election. Building on work by Ferree and Singh (1999), we scale countries as follows:

1. No executive/legislature
2. Unelected executive/legislature
3. Elected, one candidate
4. One party, multiple candidates
5. Multiple parties are legal, but only one won seats (because other parties did not exist, compete, or win seats)
6. Multiple parties competed and won seats (but one party won 75 percent or more of the seats)
7. The largest party received less than 75 percent of the seats. (Beck et al. 2001, 166–67)

Do a semantic transformation of this ordinal scale. What the gray zone would be in this transformation? Should you transform some categories to the same level, e.g., zero? (i.e., the scale is not even ordinal)?

Answer:
As a start, it seems like only levels 6 and 7 represent some form of democracy. One could argue that the gray zone is probably levels 5 and 6, with 5 being dubious. Many concepts of democracy would probably code levels 1–4 as equivalently nondemocratic.

86. One issue in dealing with concepts deals with kinds of scales, ordinal, nominal, interval, and ratio. Is the polity measure of democracy interval or ratio? If it is ratio what is the zero point?
One could argue that –10 is the zero point, completely nondemocratic, like absolute zero as in a semantic transformation. It seems plausible that the polity concept of democracy does have a possible zero point (i.e., completely nondemocratic). However, most people treat the measure as interval when changing the scale by adding 10 (which is standard practice).

87. Give two reasons why the following definition of “extreme” case does not fit with a fuzzy logic approach to concepts or one defended by Social science concepts and measurement.

Extremity (E) for the ith case can be defined in terms of the sample mean (\(\bar{X}\)) and the standard deviation (s) for that variable: 
\[ E_i = \frac{|X_i - \bar{X}|}{s}. \]
This definition of extremity is the absolute value of the Z-score (Stone 1996, 340) for the ith case. This may be understood as a matter of degrees, rather than as a (necessarily arbitrary) threshold. Since extremeness is a unidimensional concept, it may be applied with reference to any dimension of a problem, a choice that is dependent on the scholar’s research interest. Let us say that we are principally interested in countries’ level of democracy—the dependent variable in the exemplary model that we have been exploring. The mean of our democracy measure is 2.76, suggesting that, on average, the countries in the 1995 data set tend to be somewhat more democratic than autocratic (by Polity’s definition). The standard
deviation is 6.92, implying that there is a fair amount of scatter around the mean in these data. Extremeness scores for this variable, understood as deviation from the mean, can then be graphed for all countries according to the previous formula. These are displayed in Figure 3. As it happens, two countries share the largest extremeness scores (1.84): Qatar and Saudi Arabia. Both are graded as −10 on Polity’s twenty-one-point system (which ranges from −10 to +10). These are the most extreme cases in the population and, as such, pose natural subjects of investigation wherever the researcher’s principal question of interest is in regime type. (Seawright and Gerring 2008, 301).

Answer:

(1) This procedure depends on the sample and not on the meaning of the concept. For example, 2.76 is well outside the zone of what polity considers a democracy. Notice that given the positive mean of the data, all the extreme cases are nondemocracy.

(2) An extreme case based on conceptual meaning would typically mean either −10 or positive 10 on the polity scale, i.e., the extreme positive or negative pole.

88. Bühlmann, M. (et al. 2012) use “best practices” to determine the maximum value on the indicators for their Democracy Barometer. However, they claim that they do not need to worry about the minimum values (negative pole) because they are only looking at “established democracies.” Can they get by without exploring worst practices?

Terminology, definitions, and semantics

89. Sometimes scholars use a wide array of adjectives to describe what appears to be the same concept. For example, “negative peace” (Boulding 1978), “precarious peace,” (George 2000), “adversarial peace,” (Bengtsson 2000), or “pre-peace” (Bayer 2005), “conditional peace” (George 2000), or “cold peace,” (Miller 2001) all seem to refer to the same phenomenon. Discuss if this is in fact the case. Is one adjective to be preferred to the others? Are there some adjectives that are not good?

90. There is a very large literature in American politics on the concept and measurement of racism, and it is a topic that arouses much controversy: “Is white opposition to policies driven by racial prejudice or is it grounded in race-blind ideological principles?” (Feldman and Huddy 2005, 168; this is first sentence of the article). There are many aspects of this debate of interest regarding concepts. (1) One interesting aspect is the variety of terminology, such as “overt prejudice,” “new racism,” “racial resentment,” “old-fashioned racism,” and “symbolic racism.” In addition there are related concepts such as “prejudice” and “discrimination.” In particular the use of adjectives is of interest: discuss the issues and implications of the use of the concept “symbolic racism” (e.g., Tarman and Sears 2005) as a particular use of adjectives. Notice as well the relevance of the negation guideline given in the use of adjectives. (2) The negation guideline suggests that much can be gained by focusing on the
opposite concept. Much of the debate revolves around the concept of a "principled conservatism" which argues against policies such as affirmative action but which is not racist (Sniderman et al. 1996). One could focus on that concept instead of directly on racism. For example, if one is principled then those principles should apply across groups by gender and sexual orientation and not just race. It is important to think about causal relationships in concepts. One might suggest that, at least in the case of the south, principled conservatism is a consequence of racism, given that race was a central component of Southern culture and politics for centuries. See Quillian (2006) for a nice survey of the political science, sociology, and psychology literatures.

91. Alvarez, Przeworski, Cheibub, Vreeland, and Gandhi who are all active working together on the concept and data on regime type have a clear preference for the term "dictatorship" as the opposite pole to democracy. In contrast, scholars using the polity data prefer the term "autocracy" or "authoritarian." Is there anything of theoretical or conceptual importance at stake in this terminological difference?

92. Discuss this definition from Eurostat. What do they really mean?

   Qualitative Indicator: The indicator is derived from a quality assessment, e.g. in the form of a yes/no question or a question about category (e.g. gender, nationality). (Eurostat 2014 Towards a harmonized methodology for statistical indicators, 20)

   Answer: They probably mean a nominal indicator. However, "yes/no" can often translate into [0,1] with an underlying scale, which might exist depending on the "category."

93. How does the adjective "civil" function here? Is it a subsetting one, a causal relationship, or what?

   The next (and possibly most important) strand of the evolution of the understanding of peace is the civil peace. According to this approach, every individual in society has the capacity to mobilize for peace from a variety of different perspectives, whether for disarmament, for international cooperation, or against violence, discrimination, and oppression. It relates to the historical phenomena of social direct action for political, economic, and identity reasons, of citizen advocacy and mobilization, in the attainment or defence of basic human rights and values. It is also related to pacifism in its main forms, where civil action is non-violent in principle. It has been strongly influenced by a wide range of social mobilization dynamics. Without the civil peace and its social forms of mobilization, international and constitutional frameworks would not be able to connect with ordinary people in order to represent their interests, identities, needs, and aspirations. (Richmond 2014, epub 116)

   Answer:

   It seems to function as a causal relationship, i.e., as cause of peace rather than concept of peace. It is about how actions of individuals can create peace.
94. Terrorism is a big policy as well as scientific issue. There are a large variety of different adjectives attached to the concept “terrorism” including international, transnational, state-sponsored, state, domestic, political, revolutionary. In making sense of this one must consider how the adjective is modifying the terrorism concept. Related to this is that many of the adjectives involve the negative pole of a dichotomous concept, e.g., state versus nonstate. Analyze the conceptual issues involved in defining terrorism, with or without adjectives. Discuss how the decisions one makes can have a significant impact on case selection. For a survey of different concepts of terrorism see for example Crenshaw (1972), Gibbs (1989), Held (2003).

95. Following the work of Esping-Andersen, much of the work on the welfare state deals with the “commodification” of work. Here one needs to think about the negation: why use de-commodification instead of, for example, noncommodification. Also, the name of the concept implies a process instead of a state, does this matter? For a review see Knijn and Ostner (2002).

96. The word used or the meaning of the word-concept shift can easily shift over time. Discuss how this has been the case of the polity concept of “anocracy,” which is part of the three core concepts of the polity measure, democracy, anocracy, and autocracy. Go to the original Gurr (1974) and see what he meant in the 1970s and contrast that with what anocracy means today.
Answer:
Anocracy for Gurr (1974) largely meant “anarchy”, for example, see his table 2. From the Greek roots anocracy means “absence of rule.” For Gurr it means “absence of power or control” (p. 1487, footnote 11). Today anocracy means a hybrid or mixed regime, one with both democratic and authoritarian characteristics.

97. It is quite common, and probably standard practice, in the literature on “quality of democracy” to use only “consolidated democracies.” Lijphart expresses the standard view quite clearly. What is the role of the adjective “consolidated”? Does it mean “good” or “high quality” democracy?
The same question arises in the literature on peace: “Peace consolidation generally refers to a process leading towards a self-sustaining peace” (UN 2010 “Monitoring peace consolidation United Nations practitioners’ guide to benchmarking, Annex A”)

Measuring the quality of democracy is the third step in a three-step judgment. We have to begin by establishing (1) that the countries we are interested in are sufficiently democratic in terms of regime and institutional characteristics to justify being called democracies, and (2) that their democracy is firmly established. . . . I am not sure where exactly the minimum threshold for institutional democracy should be drawn. My main argument is that there must be such a minimum, below which a country does not qualify as a democracy, and that it does not make sense to discuss the quality of democracy in countries whose institutions and rules are not sufficiently democratic. (Lijphart 2011, 18)
According to Schedler, for instance, ‘consolidating democracy means reducing the probability of its breakdown to the point where [we] can feel
reasonably confident that democracy will persist.’ (Schedler 1998, 95). In a similar spirit, according to Gasiorowski and Power (1998, 740) a democracy is consolidated when it is ‘sufficiently durable that a return to non-democratic rule is no longer likely’; and according to Acemoglu and Robinson (2005, 30) ‘a democracy is consolidated if the set of institutions that characterize it endure through time.’ Empirically, questions about consolidation are addressed by examining what distinguishes democracies that achieve such durability from those that do not. In large-N research, this is frequently accomplished by specifying a temporal criterion that is used to identify consolidated democracies. Prominent examples of such criteria are Huntington’s two-turnover test and Gasiorowski’s twelve-year threshold. . . . In this article, I propose a new empirical approach to democratic consolidation that avoids some of the weaknesses of existing research and leads to new findings about the dynamics and covariates of democratic consolidation and breakdown. . . . More precisely, following existing literature, consolidation is hypothesized to take place when a large, durable, and statistically significant decline in the risk of authoritarian reversals occurs at a well-defined point during the lifetime of a democracy. . . . The results that follow strongly suggest that democratic consolidation indeed occurs: A large and durable decline in the risk of authoritarian reversals is estimated to take place between the seventeenth and twentieth year of a democracy’s existence. Consolidation reduces the annual risk of breakdown from 1 in 33 for transitional democracies to 1 in 200 for consolidated democracies. In other words, consolidation comes close to eliminating the possibility of a return to dictatorship. (Svolik 2014, 715)

98. Terminology is often very important and leads projects in certain directions. The “Minorities at Risk (MAR)” implies that the minority in question is not in the government. Could one pursue a “Majorities at Risk” project. For example, “the MAR coding scheme does not fit countries with ruling minorities or complex coalitions of ethnically defined elites, as for example in Nigeria, India, or Chad, where ethnic conflict will be pursued in the name of excluded majorities (rather than minorities) or ethnic groups that share power (and are thus not “at risk”).” (Cederman et al. 2010, 91).


100. The minorities at risk project (MAR) has receive much criticism for selecting minorities at risk. The new AMAR deals with this by adding “All” to MAR (Birnir et al. 2014). Discuss why this is bad terminology.

Answer:

Presumably the previous version, MAR, had all minorities at risk, what they have done is in fact dropped the “risk” part of the definition, it is now a dataset on minorities tout court.
Scaling

101. Matching has become a very popular way to do statistical analyses with observational data. Matching relies on comparing treatment and control groups that are well matched on key control variables or covariates. To do this means putting the covariates on the same scale to allow one to evaluate the “distance” between observations based on the covariates. In the terms of the concept book this is scaling and then aggregation. Discuss the interesting procedure whereby aggregation is fixed by the distance measure and weighting is done via scaling. This contrasts with the procedure proposed in the concept book where scaling is done separately and weighting is done via the aggregation technique. What is the implicit weighting principle for decreasing the importance of a covariate in the quote below? What is the substantive implication in terms of the logging example?

Adjusting Imbalance Metrics for Relative Importance. Although one can always define a data set that will produce large differences between any two imbalance metrics, in practice the differences among the choice of these metrics are usually not large or at least not the most influential choice in most data analysis problems. Although we describe continuous and discrete metrics below, more important is the choice of how to scale the variables that go into it. Indeed, every imbalance metric is conditional on a definition of the variables in $X$, and so researchers should think carefully about what variables may be sufficient in their application. . . . For example, consider data designed to predict the Democratic proportion of the two-party vote in a cross-section of Senate elections (“vote”), with incumbency status as the treatment, controlling for covariates population and the vote in the prior election (“lagged vote”). Clearly, lagged vote will be highly predictive of the current vote, whereas the effect of population will be tiny. Thus, in designing an imbalance metric, we want to be sure to match lagged vote very well, and should only be willing to prune observations based on population when we see gross disparities between treatment and control. For example, if we used Euclidean distance, we could code vote and lagged vote on a scale from 0 to 1 and population on a smaller scale, such as log population. In computing the Euclidean distance between two observations, then, population differences would count as equivalent to relatively small vote differences. (King et al. 2017, 478, 479)

Answer:
The example suggests that shrinking the scale, e.g., logging, reduces its impact. Logging makes some – not all – big real differences seem much smaller. But note that it also makes some small real differences seem larger.

102. Discuss the claim about bounded versus unbounded scales in Morris’s The measure of civilization how social development decides the fate of nations. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How a serious critique of that approach is this?

It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country’s score in each annual report, but
because the maximum possible score is always 1.0, the HDI does better at charting a nation’s relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

103. Discuss the semantic transformation implied in this nonlinear categorization scheme.

Following guidelines for ICEWS aggregation, we create seven count variables for conflict and cooperation among various parties: counts of high-hostility events (intensity: –10 to –8); moderate-hostility events (intensity: –7 to –4); low-hostility events (intensity: –3 to –1); neutral events (intensity: 0); low-cooperation events (intensity: 1 to 4); moderate-cooperation events (intensity: 5 to 6), and high-cooperation events (intensity: 7 to 10). We use data from all 1,122 municipalities for 20 years from 1993 to 2012. (Campbell et al. 2017)

104. Discuss the UNDP’s semantic transformation and aggregation arguments:

Societies can subsist without formal education, justifying the education minimum of 0 years. The maximum for expected years of schooling, 18, is equivalent to achieving a master’s degree in most countries. The maximum for mean years of schooling, 15, is the projected maximum of this indicator for 2025. … Having defined the minimum and maximum values, the dimension indices are calculated as: Dimension index = (actual value – minimum value) / (maximum value – minimum value) For the education dimension, equation 1 is first applied to each of the two indicators, and then the arithmetic mean of the two resulting indices is taken. Using the arithmetic mean allows perfect substitutability between mean years of schooling and expected years of schooling. Many developing countries have low school attainment among adults but are eager to achieve universal primary and secondary school enrollment. (UNDP 2018, technical-notes, 2; emphasis is mine)

105. Should power be considered a ratio variable?

Power, however, is like love; it is “easier to experience than to define or measure.” Just as one cannot say “I love you 3.6 times more than her,” (Beckley 2018, 8)

Answer:

Power almost has to be considered a ratio variable because one wants to say that a country is twice as powerful as another. See, for example, the literature on the 3:1 rule of relative power. So one should be able to say that one person is 3.6 times more powerful.

106. Discuss whether the conflict–cooperation scale in this dataset (Boschee, E. et al. 2015) is interval or ratio. The ICEWS dataset is a product of automated text analysis of more than 30 million news stories over the past 13 years, with global coverage. The news stories are classified into event categories based on the CAMEO (Conflict and Mediation Event Observations) taxonomy, which provides information about the type of event and its characteristics relative to other events. Each event category in the CAMEO classification has a numerical
value describing its intensity, ranging from negative 10 to positive 10, using Goldstein’s (1992) Conflict-Cooperation Scale. The negative values signify the hostility of an event, with –10 capturing higher hostility. The positive values represent the cooperative nature of the event, with 10 capturing higher cooperation and 1 capturing lower cooperation.

Answer:
This depends mostly on whether the zero in the middle is a true zero or not.

107. Discuss how could one use the hierarchy to give weights to each dimension, is each a necessary condition condition (in theory or practice) for the next highest level:

Even though the features in Table 1 are in theory independent of one another, they tend to cluster in practice, suggesting a hierarchical ordering of four dimensions: third-party review, third-party ruling, judges, and standing. The first question is whether the treaty provides for independent third-party review. Among pacts with some system of review, the next issue is whether rulings are directly binding in international law. Among pacts with binding rulings, those with standing tribunals are more legalistic than those with ad hoc arbitrators. Finally, tribunals with jurisdiction over claims by individuals, treaty organs, and states alike are more legalistic than those accessible only by states. In terms of remedy, the most legalistic pacts provide rulings with direct effect in national law, but the presence or absence of sanctions—though still significant—is a less meaningful indicator of legalism, with unilateral measures always available to states seeking to enforce third-party rulings in the decentralized international system. (McCall Smith 2000, 143)

Answer: This subset hierarchy allows one to use Guttman scaling techniques.

108. There are two questions. (1) Does the concept/measure have a zero point? (2) If so where is the zero point? It is an open question whether many concepts or measures of democracy are interval (i.e., no zero point) or ratio (zero point).

[W]hile some regimes are more democratic than others, unless offices are contested, they should not be considered democratic. The analogy with the proverbial pregnancy is thus that while democracy can be more or less advanced, one cannot be half-democratic: there is a natural zero-point. (Alvarez et al. 1996, 21)

Discuss these two issues considering the quote above.

109. Often scholars pay no attention to the distinction between interval and ratio variables. Much of international conflict work implicitly works with ratio variables because of the contrast of, for example, between friends and enemies or satisfied and dissatisfied states. Often this means scales where –1 is enemy and 1 is friend, and where 0 is the dividing line. For interval measures there is no real zero. One can always rescale interval measures so that they do include zero, but that does not make the measure a ratio one. Discuss whether the widely used S measure (Signorino and Ritter 1999) is interval or ratio.

Answer:
The original scale of the S measure is \([0, 1]\). This is then transformed to \([-1, 1]\) in the data used in applications. The zero in \([-1, 1]\) does not appear to be the zero needed to distinguish between, say, friends and enemies. Also, the data are very skewed towards 1: the median is often in .75 range, hence there are very few negative values in most situations. See Goertz (2008) for an extended discussion. To make the \([-1, 1]\) S into ratio variable would require some theoretical and methodological argument for determining the zero point.

110. A very standard problem in measure creation is to combine the data on two or more secondary-level dimensions. Often this involves dimensions that have different scales and then some way of structuring them together mathematically. For example:

we devise a new way to use the information of civil war duration and deaths for the two-stage regression analysis. The duration and casualties of civil wars are correlated, and it is theoretically inappropriate to estimate equation (2) for only one of them. Therefore, we combine the two characteristics of civil war into one, Civil War Destructiveness Index, by adding their standardized values. Before standardization, civil war duration is measured by the maximum duration in number of days of the war. Civil war deaths are measured by the total number of deaths experienced by the nation, divided by population. (Kang and Meernik 2005, 101)

Discuss (1) the importance of differences in variances between the two component parts in constructing the overall measure, (2) whether skewness in the data matter (often social science data are often heavily skewed), (3) the alternatives to standardizing the two scale, and (4) whether one should multiply or add the component parts.

111. An important aspect of all measure construction is standardization. Often this is done based on population, such as GNP/capita. Time is another common standardizing device, e.g., actions/year. Discuss the advantages and disadvantages of using time or the log of time to evaluate the extent of violations of the laws of war in Morrow and Jo (2006, figure 3). One might feel that to compare the number of violations one needs to take into account the length of the war.

112. The left-right scale ("rule") of the Party Manifesto Dataset (Budge et al. 2001, Klingemann et al. 2006) is one of the most widely used measures in comparative politics. The rule scale is used to situate parties on the left-right continuum based on party manifests. To do this, statements in the party manifesto are categorized according to 56 possible issue categories. The data report the percentage of all statements comprised by each category. If a category is not mentioned, the dataset reports that 0% of the statements is concerned with that category. A party’s position is derived by adding the percentages of 13 categories seen as being left and subtracting the percentages of 13 categories seen as being right.

Analyze how zero has two different uses in the coding and measurement of policy positions. Explain how this produces problems in the final measure.
In particular discuss how aggregating the zeros on the different 26 policy positions becomes zero in the final left-to-right policy scale.

Answer:

Zero is used to indicate missing data. That is when one of the 56 issues categories is not present in the party manifesto it is given a value of zero. Zero also means the middle of the −100 to +100 scale. These are two fundamentally different things. See Rehm (2008) for an extensive discussion of this and other conceptual and measurement problems.

113. Zero points on scales have important roles. For example, in interpreting Cox regression model the impact of independent variables is often assessed vis-à-vis a “baseline” model where the independent variables have value zero. Discuss how this matters using the democracy and other key variables, such as level of development, in the Hegre et al. (2001) analysis of the “democratic civil peace.”

114. Londregan and Poole (1996) transform the polity scale of −10 to 10 to −∞ to ∞. They justify this because

The 21-point GOVTYPE [i.e., polity] scale is nearly continuous. However, because it has a “floor” of −10, and a “ceiling” of 10, we need to be careful not to use a statistical mode that predicts values outside this range. A standard way to ensure this does not happen is to apply a logistic transform to the variable. Let S denote a score on the 21-point scale. Take the following variant of the logit transform of S: T(S) = \ln(S + 10.5) − \ln(10.5 - S). This converts scores to a truly continuous scale. A value of S at 10.5 would correspond to a T(S) of ∞, an S of −10.5 to a T(S) at −∞, while a value for T(S) of 3.71 corresponds to a regime type of 10. (Londregan and Poole 1996, 7)

Consider the following questions: (1) does the fact that polity stops at 10 really mean that a country cannot be more democratic? (2) The transformation used implicitly suggests that the main shift from authoritarian to democracy occurs at zero on the −10 to 10 scale. Is this appropriate? When does polity code the transition to democracy? (3) What the curve look like in the T(S) function? Does that correspond to a good concept of democracy?

Answer:

The concept book suggests that the concept of democracy should extend beyond 10 of the polity scale. Hence, there is perhaps no imperative reason to do the transformation in first place; that model predicts values greater than 10 is not problem since the polity concept extends beyond 10.

The transformation means that the model is predicting increasing marginal effects around zero, which may not be justified. It seems like the transition to democracy really occurs more toward 6 on the polity scale, and that is the official point where scholars dichotomize the polity measure. The Londregan and Poole transformation assumes a strongly decreasing marginal relationship between around 6–7, where in reality this might be where it is strongly increasing.
115. Are there situations where you really need ratio rather than interval data? For example, in conflict studies where one wants to distinguish between friends and enemies.

116. Sometimes the use of zero becomes a critical cut point. Bennett argues for a continuous specification of the democracy–autocracy continuum. He says that “In the specification advanced below, I avoid arbitrary cutoffs for democracy and autocracy” (2006, 320). Then he does the following “The first interactive similarity [of regime type] variable is constructed by multiplying the initiator’s dem [polity democracy score] score by the target’s dem score. Given dem scores ranging from –10 to +10, this multiplicative similarity variable will be high (large and positive, to a maximum of +100) for jointly autocratic and jointly democratic pairs. Similarity will be in a middle range, around 0, for pairs of mixed regimes (a dem score for either state near 0 produces a joint similarity score near 0). Similarity takes on its lowest values (to a minimum of +10) for pairs of dissimilar states (with a +10 and –10 dem score combination). If it is similar regimes that do not fight rather than jointly democratic pairs, then this variable will have a negative effect on conflict while other variables do not” (2006, 321). How might be argued to have used an arbitrary cut point with the zero?

Answer:

The construction of the interaction terms assumes that the polity measure is ratio. It could be argued that using the polity zero is arbitrary, particularly since the dichotomous coding of democracy for polity uses 6 or 7. One would need to make the argument, particularly in a fuzzy logic perspective, that zero is really the half-in, half-out point on the polity scale.

117. A very common problem is confounding “not applicable,” “missing,” and “zero,” by using the same codes, often zero. Discuss some common examples of this such as: (1) trade where very small levels of trade might be missing or might be zero (e.g., Oneal and Russett (1999) assume missing data to be 0 in the trade relationship; see also Gleditsch 2002); (2) COW militarized dispute data on fatalities, if you can find no reports on fatalities, should you code that as missing or zero. See also exercises 118, 222, 324, and 112.

118. A problem with zero codings is that it might seem like a real zero, but in fact is used for other purposes. Vreeland discusses in detail the anocracy coding of countries using the polity data: “PARCOMP [one of the five parts of the polity measure] ranges from repressed to competitive (numerically, from –2 to 3). What does the middle capture? PARCOMP is coded 0 for “not applicable situations, or, more specifically, the coding “is used for polities that are coded as Unregulated, or moving to/from that position (Gurr 1989, 14)” (2008, 406). Discuss the different things that zero means for PARCOMP.

Answer:

The PARCOMP scale goes from –2 to 3, this suggests that zero is a middle ground. In fact, it is also used to code “not applicable” which is certainly not in that middle ground.
119. Scholars often do not pay much attention to the difference between interval and ratio variables. Often methodologists and philosophers say that the higher the scale type the better, i.e., ratio is better than interval. This can conflict with the desire to log variables because of skewness. Using foreign direct investment data describe how scholars can implicitly convert ratio data into interval data.

Answer:

Raw foreign direct investment (FDI) data (e.g., in constant US dollars) can be positive or negative, hence there is a true zero which is meaningful and hence we have ratio data. Since one cannot take the log of negative numbers these most be converted into all positive numbers and thus the true zero in effect disappears making the data interval level.

120. *Social science concepts and measurement* takes a realist view of concepts that emphasizes what the theorist *means* by the concept. Many scholars will log variables because of skewness and many see skewness as a problem. Explain why an realist and one concerned by the meaning of concepts would object to Fearon’s comments:

\[ F = 1 - \sum_{i=1}^{n} p_i^2 \]

where \( p_i \) are the population shares of each ethnic group. ... And its empirical distribution—summarized in Figure 2—is not highly skewed. (footnote: Cox (1997) and others sometimes prefer to use the “effective number of ethnic groups” (or political parties’ vote or seat shares), which is \( 1 = (1 - F) \). Thus, a country with \( n \) equal-sized groups has an “effective number” of \( n \) groups, with departures from equal shares shrinking the effective number continuously. Although the interpretation is “nice,” this measure is highly skewed, at least for ethnic fractionalization, so that it tends to exaggerate the influence of very diverse countries like Tanzania when used as an explanatory variable). (Fearon 2003, 209)

Answer:

She would answer that the methodological cart is driving the conceptual and theoretical horse.

121. *The Bell Curve* (Herrnstein and Murray 1994) was a very controversial book. It turns out that the data on “intelligence” used were in fact not bell-shaped, but skewed to the right (i.e., quite a few very high scores). This skewness was “fixed” so that the data actually analyzed were bell-shaped. Discuss how this will influence their statistical results particularly given that they want to show that intelligence is a cause of outcomes such as income. It is important to note that it is completely standard procedure to assume a normal distribution in standardizing IQ data, for example, to make them comparative across nations or time.

Answer:

See Fischer et al. (1996) chapter 2 for a discussion.

122. Logging variables means that some variation in the data become much more important than other. When using raw FDI inflows instead of as a percentage
of GDP, FDI is usually logged (Chan and Mason 1992; Wei 2000; Globerman and Shapiro 2003), though not always (Oneal 1994; Li and Resnick 2003). Explain which countries become more important with logged FDI.

Answer:

Logging compresses large values and expands low values (i.e., values near zero). So when using logged values the model is trying to explain more the variation in low levels of FDI than high levels.

123. Discuss the extent to which (if any) the two secondary-level factors—civil liberties and political liberties—of the Freedom House measure of democracy are equally weighted in the final measure. See http://www.freedomhouse.org/research/survey2005.htm.

Answer:

There is slightly unequal weighting because there are different numbers of indicators used in each secondary-level dimension and because the basic level is the sum of the two secondary levels.

124. Discuss how the political rights, civil rights and electoral democracy concepts and measures are made and how they relate to each other in the Freedom House framework. See http://www.freedomhouse.org/research/freeworld/2004/methodology.htm.

Answer:

From the Freedom House web site:

In addition to providing numerical ratings, the survey assigns the designation “electoral democracy” to countries that have met certain minimum standards. In determining whether a country is an electoral democracy, Freedom House examines several key factors concerning how its national leadership is chosen.

To qualify as an electoral democracy, a state must have satisfied the following criteria:

1) A competitive, multiparty political system. 2) Universal adult suffrage for all citizens (with exceptions for restrictions that states may legitimately place on citizens as sanctions for criminal offenses). 3) Regularly contested elections conducted in conditions of ballot secrecy, reasonable ballot security, and in the absence of massive voter fraud that yields results that are unrepresentative of the public will. 4) Significant public access of major political parties to the electorate through the media and through generally open political campaigning.

The electoral democracy designation reflects a judgment about the last major national election or elections. In the case of presidential/parliamentary systems, both elections must have been free and fair on the basis of the above criteria; in parliamentary systems, the last nationwide elections for the national legislature must have been free and fair. The presence of certain irregularities during the electoral process does not automatically disqualify a country from being designated an electoral democracy. A country cannot be listed as an electoral democracy if it reflects the ongoing and overwhelming dominance of a single party or movement over the course of numerous national elections; such states are considered to
be dominant party states. Nor can a country be an electoral democracy if significant authority for national decisions resides in the hands of an unelected power, whether a monarch or a foreign international authority. A country is removed from the ranks of electoral democracies if its last national election failed to meet the criteria listed above, or if changes in law significantly eroded the public’s possibility for electoral choice.

Freedom House’s term “electoral democracy” differs from “liberal democracy” in that the latter also implies the presence of a substantial array of civil liberties. In the survey, all Free countries qualify as both electoral and liberal democracies. By contrast, some Partly Free countries qualify as electoral, but not liberal, democracies.

Consider the “Competitiveness of political participation” dimension of the polity measure of democracy (Jaggers and Gurr 1995, 472) given in the table below. Consider how the negative pole (autocracy) and positive pole (democracy) are coded. Think about the universally used measure “Democracy minus Autocracy” and the scale values that produces. What is the asymmetry between the positive and negative poles? How would you make it symmetric? The same question can be asked of the “Constraints on the Chief Executive” dimension. Ignore the substantive nature of the dimensions and just look at how things are scored.

Table 1: The Polity numeric measure of democracy

<table>
<thead>
<tr>
<th>Authority Coding</th>
<th>Democracy Weight</th>
<th>Autocracy Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Competitiveness of Political Participation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) competitive</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>(b) transitional</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>(c) factional</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>(d) restricted</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(e) suppressed</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: Jaggers and Gurr 1995, 472.

Once you do the Democracy-Autocracy operation, this dimension then goes, \(-2, -1, 1, 2, 3\). Notice then that democracy can achieve a maximum of 3 while autocracy can only get a \(-2\). Since there are five levels one could have done \(-2, -1, 0, 1, 2\).

The concept book argues that basically all concepts should be considered continuous. However, there is an extensive literature on freedom and rights that argues that they are dichotomous (see Dowding and Hees 2003 for a continuous view). This may seem counter-intuitive since most people talk about “more or less” freedom; so the implicit view of most is that liberty and freedom are continuous. Notice that in the second quote he talks about “possible” and “impossible” as dichotomous. Contrast this with the Possibility Principle.
which clearly sees possibility as continuous. Notice that in Kramer’s view “overall liberty” is continuous, but individual liberties are dichotomous.

Are particular liberties similar to overall liberty in being scalar or partitive? In other words, can each particular freedom exist to varying extents? Can somebody be free-to-\(\varphi\) to a certain degree, and be less free-to-\(\varphi\) or more free-to-\(\varphi\) than somebody else? With only one minor qualification, the current essay will maintain that the answers to these questions are negative. The existence of any particular liberty, as opposed to the extent of anybody’s overall liberty, cannot vary cardinally or ordinarily. This essay sides firmly with Ian Carter, then, who writes that ‘[t]he freedom to do \(x\) is not a matter of degree; one either is or is not free to do \(x\)’. Carter aptly oppugns ‘the claim that specific freedoms are a matter of degree (i.e., that one can be more or less free to do \(x\))’ (Carter, 1999, 228, 233, emphasis in original). Hillel Steiner robustly espouses a similar view: ‘The notion of degrees of freedom to do an action is superfluous, misleading and descriptively imprecise’ (Steiner, 1983, 78). (Kramer 2002, 231)

In being non-scalar, ‘free’ and ‘not free’ in discussions of particular actions and types of actions are on a par with ‘possible’ and ‘impossible’. When \(P\) has been able to perform \(X\) and has now become able to perform \(X\) in some novel manner, we cannot correctly state that his performance of \(X\) is now more possible than before. ‘Possible’ applies in an all-or-nothing fashion, as does ‘impossible’. Instead of declaring that the performance of \(X\) by \(P\) is now more possible than previously, we ought to declare that the performance of \(X\) by \(P\) is now possible in more ways than previously. (Kramer 2002, 242)

127. Critical to the construction and analysis of concepts is the underlying scale or continuum from the positive to negative poles. One way to argue for an underlying continuum is if the data have the structure of a Guttman scale (Guttman 1944). The CIRI measure of human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. If the data fit the Guttman requirements then each category is a subset of the lower level categories. For example, if a state does (2) then does (1); if it does (3) then it also does (2) and (1); if it does (4) then it does (1)–(3). Cingranelli and Richards (1997) argue that their data on physical integrity rights has basically this structure. Draw a Venn diagram illustrating what this looks like.

Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a given year; their final score is the sum of the scores of the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

Addition is one way to structure the concept. Another is the maximum. To make this simpler, assume that one can only score 0 or two on each dimension. Provide a rescaling of the four dimensions of the CIRI scale so that when you use the maximum to aggregate you would get exactly the same final score as the CIRI data if the data fit perfectly the Guttman requirements.
One problem is that the data do not exactly fit the Guttman requirements. For example, there are cases with “extrajudicial killing” but no “political imprisonment.” How would your proposed rescaling and use of the maximum deal with these cases? Do you think it is better or worse than using the sum?

Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How does your proposal using the maximum deal with or not this critique?

Answer:

Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of eight. With the new scoring system it would also have a score of eight. Note that things would get more complicated with a 0, 1, or 2 scale, instead of the 0 or 2, but the logic would be the same.

The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

128. Often the line between nominal, ordinal and interval is blurred once one begins to do mathematical operations like addition or multiplication. Explain why the polity measure is interval, not ordinal. For example, see this discussion:

One way to categorize this growing corpus of indicators is by the type of scale used to measure the key concept (democracy)—binary, ordinal, or interval. Binary indices include the democracy–dictatorship (“DD”) index produced by Przeworski and collaborators (Cheibub, Gandhi, & Vreeland, 2010) and an index produced by Boix, Miller, and Rosato (2013, hereafter “BMR”). Ordinal measures include the Political Rights (“PR”) index and the Civil Liberty (“CL”) index, both produced by Freedom House (2013), along with the Polity index drawn from the Polity IV database (Marshall, Gurr, & Jaggers, 2013)” (Skaaning et al. 2015, 2).

Answer:

Once one begins to add (or subtract) one is implicitly converting the binary or ordinal variables into interval ones. This is classic measurement theory.

129. The AMAR “rebellion” scale (Birnir et al. 2018) has as its lowest rebellion level “political banditry and/or sporadic terrorism.” Often zero on the scale (like in many dichotomous variables) is potentially a causally heterogeneous category. Discuss to what extent theoretically zero is conceived of as “no mobilization” as opposed to nonviolent mobilization such as protests or nonviolent action. Would the causal mechanisms leading to these two different forms of zero be the same or different?

130. As discussed in the (Goertz 2020) chapter on scaling, it is often not so clear what kind of scale is being used. Most often the problem is most severe
between interval and ratio scales. However, it can occur as well between nominal and ordinal concepts.

Gerring discusses Table 7.3 in his measurement chapter in the section on scaling:

Limiting himself to the concept of electoral contestation (a dimension of the larger concept of democracy), Gerardo Munck defines a four-part nominal scale including categories for authoritarianism, semi-authoritarianism, semi-democracy, and democracy. Here, each category is distinguishable and clearly ranked relative to the concept of theoretical interest. . . . The unconstrained nominal-scale indicator is also more likely to approximate the virtues of an interval scale, where neighboring categories are equidistant from each other. . . . A final option for the four-point scale in Table 7.3 may also be envisioned. Insofar as the first category comprises a true zero – no contestation whatsoever – the key dimension of electoral contestation may be redefined as a ratio scale (an option that Munck endorses).

(Gerring 2012, 170, 171)

Discuss the tension between a “nominal” scale with no underlying conceptual dimension (by definition of nominal scale) versus the common notion of a autocracy–democracy continuum.

Concept structure, aggregation and hybrid concepts

Morris’s wide ranging book The measure of civilization how social development decides the fate of nations critically uses a key measure of the concept of “ social development.” Below is an scaling–aggregation decision he faced. Discuss the advantages and disadvantages of these two aggregation techniques, sum then log versus log then sum. Morris focuses on the differences at the low end of the index, why would in general sum of logs show greater differences? Notice that the log in general accentuates differences on the low end.

The simplest way to bring out the variations in premodern social development scores is by representing the data on a log-linear graph. I did this at several points in Why the West Rules. For now, adding up the scores
on the four traits and then calculating the logarithm of the sum. This has its uses, but if we instead calculate the separate logarithms of the four traits and then add up the logs to produce a single score, we end up with a graph that is no longer dominated by the high energy capture scores. The two upper lines in figure 7.6 show the log of the sum of the traits for West and East, while the two lower lines show the sum of the logs of the traits. As can be seen, summing the logs of the traits produces curves that are more sensitive to even quite small pre-twentieth-century changes in organization, war making, and information technology. The sum of the logs, while less intuitively obvious, is more sensitive to small premodern changes in organization, war making, and information technology. (Morris 2013, 250)

132. Discuss this philosophical claim that “natural kinds” cannot be formed via the logical OR (disjunction). This means that many social science concepts cannot be natural kinds.

I considered the question of whether a natural kind could be structured disjunctively, and I argued that disjunctive categories do not correspond to natural kinds (see section 1.4). The main consideration I put forward for thinking so was that disjunctive categories are not projectible. It is not that philosophers have privileged access to the a priori or conceptual truth that natural kinds cannot be disjunctively structured. Rather, given that the whole point of discovering natural kinds is to discover sets of co-instantiated properties that are inductively linked to a series of other properties, disjunctive categories do not enable us to achieve this goal. Disjunction undermines projectibility, so it defeats the purpose of having natural kinds in the first place. Moreover, this conclusion is vindicated by the fact that disjunctive kinds are rarely if ever attested in actual scientific practice, and where they are, they appear to be challenged by scientists for sound reasons. (Khalidi 2013, 203)

133. In the literature on HDI and similar measures the question arises about equal weighting of the dimensions. Can making all the dimensions necessary be counted as equal weighting?

Another way of justifying equal weighting was based on normative ideas. In particular, Haq claimed that the weights be determined ‘on the simple premise that all these choices were very important and that there was no a priori rationale for giving a higher weight to one choice than to another’ (Haq 1995b: 48). Similarly, Jahan explained: ‘The variables [which should be read as “dimensions”] are given equal weights not because of simplicity, but because of the philosophical reasoning that all the components included in the HDI are equally important and desirable in their own rights for building human capabilities’ (Jahan 2002: 155). (Hirai 2017, 85)

134. People often push for “integrated” development. Does suggest an additive concept structure or something else? Take the various measures of well-being or poverty as examples. See also the Women’s Peace and Security Index: “The Sustainable Development Goals (SDGs) were adopted as an “integrated and indivisible” set of goals, and we sought to respect that principle by giving equal weight to each of the three dimensions in the WPS Index.” (Georgetown University’s Institute for Women, Peace and Security, PRIO 2017, 57)
Discuss the numerous scaling and aggregation decisions in the OECD approach to measuring well-being. For example:

Box 1.1. The OECD approach to measuring well-being The OECD framework for measuring well-being was introduced in How’s Life? 2011. It builds on a variety of national and international initiatives for measuring the progress of societies, as well as on the recommendations of the Stiglitz, Sen and Fitoussi Report (2009) and the input provided by the National Statistical Offices (NSOs) represented in the OECD Committee on Statistics and Statistical Policy. Conceptually, the framework reflects elements of the capabilities approach (Sen, 1985; Alkire and Sarwar, 2009; Anand, Durand and Heckman, 2011), with many dimensions addressing the factors that can expand peoples choices and opportunities to live the lives that they value including health, education and income (see OECD, 2013a). This approach to measuring current well-being has several important features: It puts people (individuals and households) at the centre of the assessment, focusing on their life circumstances and their experiences of well-being. It focuses on well-being outcomes aspects of life that are directly and intrinsically important to people rather than the inputs and outputs that might be used to deliver those outcomes. For example, in the education dimension, measures focus on the skills and competencies achieved, rather than on the money spent on schools or the number of teachers trained. It includes outcomes that are both objective (i.e. observable by a third party) and intrinsically subjective (i.e. those where only the person concerned can report on their inner feelings and states), recognising that objective evidence about peoples life circumstances can be usefully complemented by information about how people experience their lives. . . . Taking these stocks as the primary measurement focus is in line with the recommendations of the Stiglitz, Sen and Fitoussi Report (2009) as well as several other recent measurement initiatives, including the UNECE-Eurostat-OECD Task Force on Measuring Sustainable Development (United Nations, 2009), the UNU-IDHP and UNEPs Inclusive Wealth Report (2012), the Conference of European Statisticians Recommendations on Measuring Sustainable Development (UNECE, 2014) and several country initiatives (e.g. FSO, 2015; Statistics New Zealand, 2011).

Discuss the UNDP’s semantic transformation and aggregation arguments:

Societies can subsist without formal education, justifying the education minimum of 0 years. The maximum for expected years of schooling, 18, is equivalent to achieving a master’s degree in most countries. The maximum for mean years of schooling, 15, is the projected maximum of this indicator for 2025. . . . Having defined the minimum and maximum values, the dimension indices are calculated as: Dimension index = (actual value – minimum value)/(maximum value – minimum value) For the education dimension, equation 1 is first applied to each of the two indicators, and then the arithmetic mean of the two resulting indices is taken. Using the arithmetic mean allows perfect substitutability between mean years of schooling and expected years of schooling. Many developing countries have low school attainment among adults but are eager to achieve universal primary and secondary school enrollment. (UNDP 2018, technical-notes, 2; emphasis is mine)
137. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the “one-drop rule”?

The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the census’s history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

138. Walter (1997, 340–41) gives “at least” three conditions that “must hold” for a third-party guarantee to be effective in ending civil wars. Are each of these necessary? Is there substitutability between them?

139. Examine Hartzell and Hoddie’s (2003) concept of “power sharing.” How does substitutability play a big role in this concept?

Answer:

“An innovation of this study is to suggest that power sharing should instead be considered a continuous variable that ranges in value from zero to four with each increment representing an additional dimension of power sharing (political, territorial, military, and economic) specified in the rules governing the society. The greater the overall number of power-sharing dimensions specified, the more likely that peaceful relates among collectivities will endure.” (Hartzell and Hoddie 2003, 321) This clearly indicates a family resemblance structure.

“A second reason that the inclusion of multiple dimensions of power sharing in a negotiated civil war settlement proves advantageous to the prospects of long-term peace is that it serves as a source of protection against the failure to implement any single power-sharing provision of the settlement.” (Hartzell and Hoddie 2003, 321). This clearly indicates that substitutability is core to power-sharing.

One might test to see if any of the four measures is necessary. Also, one might look at different forms of substitutability in addition to the straight additive measure.

140. Substitutability places a key role in a couple of aspects of Gibson and Howard’s (2007) analysis of the scapegoating of Jews in Russia in the 1990s. They note that Jews have not been scapegoated in the 1990s. (1) How generally would substitutability play a key role in thinking about scapegoating in general and why some plausible groups might not be targeted. (2) Use substitutability theory to suggest changes in the approach to the psychological variables, dogmatism, xenophobia, etc. used as part of the explanation of scapegoating.
(1) It may be that there are multiple possible choices from scapegoats, i.e., they are substitutable. So, for example, failure to scapegoat Jews is because there is some substitute for them: “Finding whom to blame is not necessarily easy or even possible, but for many Russians it likely makes more sense to blame the Communists, the nomenklatura, the legacy of Communism, oligarchs, the West, etc., than to blame Jews.” (2) It may be that you need only one of these psychological attitudes, e.g., dogmatism OR xenophobia OR etc. to exhibit scapegoat behavior.

141. Examine substitutability claims in Doyle and Sambanis’s analysis (2000) of the effectiveness of UN peacekeeping.

Answer:

“All strategies should address the local sources of hostility, the local capacities for change, and the (net) specific degree of international commitment available to assist change. One can conceive of these as the three dimensions of a triangle whose area is the ‘political space,’ or effective capacity, for building peace. This suggests that the dimensions substitute for one another, that is, more of one substitutes for less of another; less extreme hostilities substitute for weak local capacity or minor international commitment.” (Doyle and Sambanis 2000, 781; see also the statistical model of table 4, p. 791)

However, “Thus, we theorize that the PB [peacebuilding] process is captured by: PB = IC [International Capabilities]\*NLC [Net Local Capacity].” (Doyle and Sambanis 2000, 782)

142. Multiple necessary conditions are sometimes used in a theory to eliminate multiple ways of getting a zero on the dependent variable. Frequently the language of constraints is used, e.g., the absence of the necessary condition constrains the actor into a different path. The multiple necessary conditions thus eliminate substitutable options. Discuss how this works in Doner et al. (2005). Also, explain how the structure of their theory is very similar to Downing’s (1992) on the emergence of democracy in early modern Europe?

143. Bennett (2006) and Werner (2000) uses different measures (and hence concepts) of “similarity” between two countries’ regime types. Ignoring that Werner uses 5 polity indicators, compare the two similarity measures using the information in equation for the generalized mean:

\[
\min(a, b) \leq (a \times b)^{1/2} \leq \frac{a + b}{2} \leq \left(\frac{a^2 + b^2}{2}\right)^{1/2} \leq \max(a, b)
\]

In general will the Bennett measure be larger or small than Werner’s? What would a table or figure comparing these two measures look like?

144. Brancati (2006) presents a very interesting argument for using the maximum. It combines case selection with the use of the maximum as an aggregation technique. What is the error she makes about the maximum as an integration technique?
The analysis presented in this article is an ordered logit analysis since the dependent variables in this study are categories of conflict and rebellion ordered from low to high forms of conflict and rebellion. This type of analysis does not assume that the categories of intercommunal conflict and antiregime rebellion are equally spaced, although it does assume that the effects of the explanatory variables are the same for all categories of the dependent variables. For this analysis, I aggregate the group level data to the national level using the maximum level of antiregime rebellion among “at-risk” groups in a country per year, and the maximum level of intercommunal conflict among “at-risk” groups in a country per year as my dependent variables.

I aggregate the data to the national level using the maximum value of conflict or rebellion in a country per year because this method of aggregation overcomes group-based selection bias in the MAR dataset. Since the MAR data set does not measure antiregime rebellion and intercommunal conflict for all groups in a country but only “at risk” groups, any measure that aggregates across groups, such as the median level of conflict or rebellion in a country, is biased by the absence of groups not deemed “at risk.” Using the maximum level of conflict or rebellion in a country overcomes this problem because it does not aggregate across groups. (Brancati 2006, 673)

Answer:
The maximum is an aggregation technique. Using the maximum means that groups not in the dataset are assumed to have lower values than ones in the dataset.

145. The concept book shows that the extension can go up with intension (i.e., adding secondary-level dimensions) for the family resemblance structure. Construct an example where it can go down.

Answer:
To do this create an example where the new dimension is very rare empirically.

146. If the extension of a family resemblance concept can go either up or down with intension (see the previous question) then can you provide a general rule from when each direction (i.e., up or down) is likely to happen?

Answer:
If the added dimension is empirically very common then extension is likely to go up. If the added dimension is empirically rare then extension is likely to go down.

147. Frequently ideal types involve a long list of characteristics. Why does that naturally lead to the use of family resemblance m-of-n procedures when constructing measures? Fearon (2003) might be an example.

Answer:
It is very unlikely that any given object will actually have all the characteristics listed as part of the ideal type. Hence, it is natural that one only requires a certain minimum number. One can perhaps generalize that when long lists of characteristics are given that family resemblance structures will be used, but
that when the number is small, necessary and sufficient condition structures are more likely.

148. How does Sartori view the question of sufficiency in conceptualization and concept structure? See also Collier and Levitsky (1997) for the same question. Answer:

He focuses his attention on the minimal necessary conditions and does not explicitly make claims about sufficiency. In one footnote he explicitly rejects sufficiency ideas. Collier and Levitsky make no claims about sufficiency and focus only on the minimum requirements for democracy.

149. It was suggested in the concept book that hybrid concepts – with some necessary and some facilitating conditions – might be good for idea types. Is this what Diamon and Morlino are suggesting?

The five essays that follow are part of a collaborative effort, launched at a conference at Stanford University, to elaborate and refine the concept of democratic quality and to apply it to a series of six paired country comparisons. We asked each author to discuss a particular dimension of the quality of democracy such as freedom, the rule of law, vertical accountability, responsiveness, and equality (our own list, and by no means exhaustive). We wanted each author to explain how the dimension in question relates to other dimensions in our framework, to suggest possible indicators for measuring the dimension, to identify ways in which this element of democratic quality is subverted in the real world, and to offer (where possible) policy recommendations. Our full framework features eight dimensions: the five outlined above, plus participation, competition, and horizontal accountability. Other dimensions might include transparency and the effectiveness of representation. The different aspects of democratic quality overlap, however, and we choose to treat these latter two as elements of our principal dimensions. We attempt here to identify some of the ways in which the different elements of democracy not only overlap, but also depend upon one another, forming a system in which improvement along one dimension (such as participation) can have beneficial effects along others (such as equality and accountability). At the same time, however, there can be trade-offs between the different dimensions of democratic quality, and it is impossible to maximize all of them at once. In this sense at least, every democratic country must make an inherently value-laden choice about what kind of democracy it wishes to be. Talk of a “good” or “better” democracy implies knowing what democracy is. At a minimum, democracy requires: 1) universal, adult suffrage; 2) recurring, free, competitive, and fair elections; 3) more than one serious political party; and 4) alternative sources of information. If elections are to be truly meaningful, free, and fair, there must be some degree of civil and political freedom beyond the electoral arena so that citizens can articulate and organize around their political beliefs and interests. Once a country meets these basic standards, further empirical analysis can ask how well it achieves the three main goals of an ideal democracy—political and civil freedom, popular sovereignty (control over public policies and the officials who make them), and political equality (in these rights and powers)—as well as broader standards of good governance (such as transparency, legality, and responsible rule). (Diamond and Morlino 2004, 20–21)
The *Diagnostic and statistical manual of mental disorders* (DSM-V) provides many nice examples of how important decisions are made using mostly family resemblance concept structures. Choose some of the diseases on the list and analyze their structure.

In *Resort to Arms*, Small and Singer (1982, 210) defined a civil war as “any armed conflict that involves (a) military action internal to the metropole, (b) the active participation of the national government, and (c) effective resistance by both sides.” This appears as a list concept. How can you ascertain, given that data are being collected, if the necessary and sufficient condition or family resemblance concept structure is being used? The same question can be asked of Sambanis (2004) who has an “operational definition” of civil war.

**Answer:**

Typically it is easy to see the necessary and sufficient condition structure even though it is not explicitly stated. If you examine how the data are actually collected almost always failure on one item of the definition means that the case is excluded. This means that each item is in fact a necessary condition.

Examine Hartzell and Hoddie’s (2003) concept of “power sharing” and the resolution of civil wars (one only need read pp. 318–21). What is the structure of the concept (necessary and sufficient condition or family resemblance)? How does substitutability play a big role in this concept?

**Answer:**

“An innovation of this study is to suggest that power sharing should instead be considered a continuous variable that ranges in value from zero to four with each increment representing an additional dimension of power sharing (political, territorial, military, and economic) specified in the rules governing the society. The greater the overall number of power-sharing dimensions specified, the more likely that peaceful relates among collectivities will endure.” (Hartzell and Hoddie 2003, 321) This clearly indicates a family resemblance structure.

“A second reason that the inclusion of multiple dimensions of power sharing in a negotiated civil war settlement proves advantageous to the prospects of long-term peace is that it serves as a source of protection against the failure to implement any single power-sharing provision of the settlement.” (Hartzell and Hoddie 2003, 321). This clearly indicates that substitutability is core to power-sharing.

Discuss the King and Murray’s measure of “human security”. (1) Compare their analysis to Nussbaum’s view of human well-being (2) Evaluate their claim that they do not need to justify their weights because they are not using any.

According to our definition, a person is in a state of generalized poverty whenever he or she dips below the pre-defined threshold in any of the component areas of well-being. Our dichotomization of each component of well-being is based on the belief that there is a qualitative difference in
life experience above and below the threshold. For example, the difference between not having enough food and nutrition to survive and having enough food is fundamentally different from the difference between having enough food to survive and having food that also tastes especially good.

A key advantage of our definition of generalized poverty is that it does not require a set of weights to be developed to equalize the different domains of well-being. Since a person missing even one of these essential elements for any part of a year would be considered impoverished, the only arbitrary element in the definition is the threshold for each domain of well-being. Moreover, the policy world has much experience with choosing threshold values for income and many other areas (such as to decide whether individuals qualify for certain programs). Although these thresholds are arbitrary and can be improved in theory in some ways (at the cost of simplicity), they are frequently used because they are fairly accurate reactions of peoples life experiences and are simple to use. In addition, small changes in these thresholds do not always produce as large changes in population-based indexes as weights would in an aggregate well-being index.

For example, we treat both (1) being tortured and (2) being tortured and starving, as impoverished and unacceptable conditions. Condition (2) may be harder to remedy than (1), but we do not have to decide how much worse (2) is than (1) in order to decide that the person is experiencing a state of generalized poverty. Similarly, few would argue that persons to be tortured four times in the next year are secure no matter how high their income. Rather, being tortured in the next year will put them in a state of deprivation or generalized poverty. The prospect of this outcome makes them insecure today. Of course, for analytical purposes other than defining human security, definitions of generalized poverty that include trade-offs between the level of achievement in one domain of well being versus another may be appropriate, particularly when individuals freely choose to balance some domains of well-being against others. Since we do not need to create and justify weights in combining domains, we can include as many other domains as the international community agrees on. For example, we can include education as a domain of well-being, even though it was once not considered an essential element for having a minimal level of well-being. (King and Murray 2002, 594–95)

Answer:
Clearly they take a necessary condition view of the “human security” concept. They are in fact weighting each dimension, and could use fuzzy logic weights if they wanted.

154. Analyze in three-level terms Vanhanen’s concepts of “ethnic heterogeneity.” In particular, contrast the issue of redundancy (or substitutability) at the indicator level and the secondary level.

The measurement of ethnic division therefore is based on three types of ethnic groups: (1) ethnic groups based on racial differences, (2) ethnic groups based on linguistic, national or tribal differences, and (3) ethnic groups based on stabilized old religious communities. Thus we have three operationally defined indicators to measure three dimensions of ethnic division. In each dimension, the level of ethnic division will be measured
by the percentage of the largest ethnic group of the country’s total population. Together the three percentages measure the relative degree of ethnic homogeneity, and the inverse percentages measure the degree of ethnic heterogeneity. The three inverse percentages of ethnic heterogeneity are combined into an Index of Ethnic Heterogeneity (EH) by summing the three percentages. This index is used as the principal operational substitute for the hypothetical concept “ethnic division.” (Vanhanen 1999, 59)

Contrast that with his three-level concept of “ethnic conflict.” What is the three-level structure here? In particular, in terms of combining the two secondary-level dimensions of “violent ethnic conflict” and “institutionalized ethnic conflict” what is the role of redundancy?

Answer:

Vanhanen justifies his procedure:

Is it justified to combine the three dimensions of ethnic heterogeneity into an index in this way or do the dimensions differ too greatly from each other? My argument is that different forms of ethnic division strengthen each other and that, it is justified to sum the percentages. For example, if the population of a country is divided into separate ethnic groups not only by language or nationality, but also by clear racial or religious division, the degree of ethnic heterogeneity is higher than in a country in which the population is divided into separate groups only by language or religion. Besides, I am not especially interested in the affects of any particular ethnic division, I am interested in the total level of ethnic heterogeneity. (Vanhanen 1999, 59)

If one thinks that redundancy is acceptable because each factor reinforces the other, one might think of multiplication instead of addition at the secondary level. Also, one wonders why the argument he uses at the secondary level cannot also be used at the indicator level where he applies the maximum, thus signaling complete substitutability.

When combining his two secondary-level factors of ethnic conflict he adds them. But since he argues that they are two indicators of the same thing, the mean would be more appropriate.

Because the two scales of institutionalized and violent ethnic conflict are intended to measure the same phenomenon from two different perspectives, it is reasonable to combine the scores of the two dimensions into a combined Index of Ethnic Conflict (EC). It is done simply by adding the scores. This means that the value of EC can vary between 0 and 200. This index is used as the principal empirical indicator of ethnic conflict in this study. (Vanhanen 1999, 62)

155. Describe Kramer’s (2002) view of the concept of “liberty” or “freedom” as a three-level concept, where the levels are (1) overall liberty, (2) individual liberty, and (3) token liberty. What is the structural principle used at each level?

Answer:
The secondary-level dimensions of individual liberties which combine via family resemblance to constitute overall liberty. This produces a continuous basic-level concept, e.g., 3/5 particular liberties defines .60 as the level of overall liberty. Token liberties – data/ indicator level – are specific occasions where individuals can have or not (dichotomously) liberty. Kramer argues that if an individual token has liberty then the secondary-level dimension has liberty (p. 240).

156. The concept of customary international law has provoked much discussion. Traditionally the concept has two secondary-level dimensions “state practice” and opinio juris (“the practice is required by, or consistent with, prevailing international law” (Kirgis 1987, 146). Thus international law faces the same problem as concept-builders in thinking about how to structure the relationship between these two defining dimensions. Discuss Kirgis’s approach, it is necessary condition or family resemblance?

Answer:

Kirgis clearly argues that there cases where if the weight of one secondary-level dimension is larger enough it can substitute for lack of state practice. Human rights law would be an example.

157. “Any definition that requires a combination of characteristics captures only a subset of most of the identities that comparative political scientists classify as ethnic” (Chandra 2008, 413). What is the implicit assumption about concept structure that Chandra is making if this statement is to be true.

Answer:

She is clearly talking about a necessary condition structure: “Although I have discussed each characteristic individually, most definitions of ethnic identity incorporate various secondary characteristics in addition to the primary characteristic of descent. Any definition that requires a combination of characteristics captures only a subset of most of the identities that comparative political scientists classify as ethnic. This is because, as shown above, each characteristic, taken singly, captures only a subset of these identities. Because any one of the defining characteristics discussed above captures at best a subset of the classification that we started with, any combination of characteristics will capture a still smaller subset. Figure 1 illustrates this point” (Chandra 2008, 412–13).

158. Fearon (2003) has proposed an important concept of “ethnic group.” Here is the core of his concept. What is the structure used for this concept?

Another approach to definition—in several ways more useful for the purpose of constructing a list by countries—is to employ the idea of “radial categories” advanced by linguists and cognitive scientists (Lakoff, 1987; see Collier and Mahon, 1993, for a discussion with respect to political science). In practice, people may understand the meaning of a concept X by reference to prototypical cases. Less prototypical cases may not share all the features of a prototype, and yet still be validly classed as Xs, at least in some circumstances. For example, the prototypical ethnic group has the following features: 1. Membership in the group is reckoned primarily by
It appears that he is using an implicit family resemblance structure in that not all seven characteristics must be present to be coded as an ethnic group.

Chandra and Wilkinson have this to say about Fearon: “In an improvement over the other two data sets, Fearon (2003) does provide a conceptual justification for his count. He attempts to include groups that fulfill as many of the following prototypical criteria as possible: First, membership is reckoned primarily by descent. Second, members are conscious of group membership. Third, members share distinguishing cultural features. Fourth, these cultural features are valued by a majority of members. Fifth, the group has a homeland. Sixth, the group has a shared history as a group that is “not wholly manufactured but has some basis in fact. Seventh, the group “is potentially stand alone in a conceptual sense – that is, it is not a caste or caste-like group” (p. 201). . . . Nor is it clear how many prototypical criteria a group must satisfy to be included or how a coder should decide between multiple candidate groups on multiple dimensions that fit the prototypical criteria. Why, for instance, was the category Jat (included in Atlas Narodov Mira but not in Fearon’s work [2003]), which appears to meet the first six criteria but not the seventh, not chosen over the category Punjabi, which appears to meet the first, third, fourth, and fifth criteria but not the second, sixth, and arguably, seventh? Some of the groups included in Fearon’s count do not meet several of the conditions. Hindi speakers, for instance, are not a group in which
members are conscious of group membership, share distinguishing cultural features that are valued by a majority of members, and have or remember a homeland” (2008, 531).

159. Critical to the construction and analysis of concepts is the underlying scale or continuum from the positive to negative poles. One way to argue for an underlying continuum is if the data have the structure of a Guttman scale (Guttman 1944). The CIRI measure of human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. If the data fit the Guttman requirements then each category is a subset of the lower level categories. For example, if a state does (2) then does (1); if it does (3) then it also does (2) and (1); if it does (4) then it does (1)–(3). Cingranelli and Richards (1997) argue that their data on physical integrity rights has basically this structure. Draw a Venn diagram illustrating what this looks like.

Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a given year; their final score is the sum of the scores of the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

Addition is one way to structure the concept. Another is the maximum. To make this simpler, assume that one can only score 0 or two on each dimension. Provide a rescaling of the four dimensions of the CIRI scale so that when you use the maximum to aggregate you would get exactly the same final score as the CIRI data if the data fit perfectly the Guttman requirements.

One problem is that the data do not exactly fit the Guttman requirements. For example, there are cases with “extrajudicial killing” but no “political imprisonment.” How would your proposed rescaling and use of the maximum deal with these cases? Do you think it is better or worse than using the sum.

Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How does your proposal using the maximum deal with or not this critique?

Answer:

Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of eight. With the new scoring system it would also have a score of eight. Note that things would get more complicated with a 0, 1, or 2 scale, instead of the 0 or 2, but the logic would be the same.

The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

160. Pevehouse in an important study of the impact of regional organizations on democratization wants to measure the “democraticness” of a regional organization. This is defined as the number of member states that are democracies.
Of course a given country is likely a member of numerous regional IGOs, the question is then is how to aggregate to give one number for the variable: “I use only the most democratic organization to measure each state’s IO involvement (versus an average of all IOs) since it should take only one membership to supply any of the causal mechanisms posited by my theory” (Pevehouse 2005, 70). Discuss this rationale for using the maximum as an aggregation or structural principle.

Answer:

One might argue that it is the “net effect” of democratic and authoritarian IGO memberships. Also, it is possible that multiple memberships in IGOs might have an increased effect. Hence one could test variables such maximum authoritarian membership in regional IGOs, or sum of democratic IGO memberships.

161. Abbott, Keohane, Moravcsik, Slaughter, and Snidal (2000) published a very influential article on “The concept of legalization.” This article could serve as a good example for the discussion of concepts. (1) Draw a diagram that represents the three-level structure of their concept. The three secondary-level dimensions are quite clear and the indicators are given in various tables. (2) What is the structure that they implicitly use to connect secondary-level dimensions and indicators? (3) Thinking about the negation of a concept is important; what about the negations given in figure 1 as well as table 1? (4) “Delegation” (table 4) might be separated into two, or even three, separate dimensions? Are “dispute resolution,” “rule-making” and “implementation” part of the higher level concept of “delegation?” (5) In the delegation dimension the concept of “binding” appears frequently, but is also central to the “obligation” dimension. Is this a problem?

Answer:

It seems like the implicit structuring used at both levels is the family resemblance one. There is not much sense that any of these are necessary. There are potentially other negations that could be used in figure 1, such as “nonbinding” for obligation. For precision it is interesting that the positive pole uses the term “rule” while the negative pole uses the term “norm.” Delegation is problematic because it includes both implementation and arbitration. In terms of figure 1, “international court” suggests third-party dispute settlement mechanism and so the opposite might be bilateral agreements. Much of the problem with delegation is that this is where most of the organizational dimensions of legalization appear; but it is not obvious that one would lump administrative, monitoring, and sanctioning organizations together with courts (i.e., one would not do this in an analysis of domestic governance and legal systems.) Also missing is a potential dimension about who makes international rules, e.g., states, United Nations, IGOs, etc.

162. Brown (2010) proposes a measure for measuring IGO delegation and then applies it to two IGOs. Table 1 has the standard three-level concept structure. He defends a additive index of all the indicator-level variables. Discuss
alternative aggregation procedures. Different indicators have different maximum possible values which implies a weighting scheme. The additive index assumes substitutability and equivalence, theoretical and empirical, of the various indicators. Discuss this weighting and aggregation scheme.

Bühlmann et al.’s (2012) Democracy Barometer uses an interesting graphical device to show how countries score on the nine key components of the concept, reproduced here as figure 1. Propose how one could use this to give a measure from zero to 1.0 of the overall qualitative of democracy. Could you extend this to 9-dimensional space?

Figure 1: Measuring the quality of democracy

**Figure 2** Shapes of democracy, three countries compared. IL: Individual Liberties; RL: Rule of Law; PS: Public Sphere; CO: Competition; MC: Mutual Constraints; GC: Governmental Capability; TR: Transparency; PA: Participation; and RE: Representation.

Source: Bühlmann et al. 2012.

Answer:
The overall measure would be percentage of the polygon covered in gray. An extension to nine-dimensional space would be the volume of the convex surface defined by the nine points (one can visualize this with a cube).

164. Call (2010) discusses the literature on the important and large concept of “failed state.” He argues that there are three large secondary-level dimensions to this concept conceptualized by three “gaps”: (1) capacity gap, (2) security gap, and (3) legitimacy gap. He illustrates this via a triangle with each dimension at a corner. What is an alternative way to illustrate this? What are its advantages and disadvantages vis-à-vis a triangle figure?

Answer:
A perhaps more natural way to illustrate this is via a cube. If one tries to locate a country within a triangle it will be very difficult to illustrate its distance from each corner, because one is trying to force three dimensions into two. Also, the three-dimensional version of a triangle is often a pyramid. This is not good because one cannot be high (or low) on one dimension and low on another.

165. Skaaning et al. make the claim: “Our claim is that once a minimal level of inclusion has been attained—sufficient to constitute an electorate and hence the precondition for an election, as discussed—further increases in suffrage are irrelevant unless and until elections are competitive. This argument will be taken up below.” (Skaaning et al. 2015, 2). Discuss how aggregation and substitutability are explicitly or implicitly involved conceptualization and measurement of this “lexical” index.

Typologies

166. Sometimes concepts are basically defined via typologies. This raises the question of there is some kind of basic level concept. The idea of a positive or negative pole makes little sense if the typology is really of different kinds, i.e., apples and oranges. As the following two questions also illustrate, a key assumption in concepts is that they are causally homogeneous: they have the same causes and effects. The dominant Correlates of War (see also the ATOP alliance data set, Leeds et al. 2002) defines “alliance” (Singer and Small 1966; Gibler and Sarkees 2004) in terms of three types of alliance: (1) defense, (2) neutrality-nonaggression, (3) entente. Discuss whether these three types of alliance have potential problems of causal heterogeneity. For a concrete example of this criticism see Long’s (2003) analysis of the alliance-trade relationship.

Answer:
Many have had their doubts about whether the neutrality-nonaggression alliances belong. Typically in the liberal peace literature the alliance variable excludes them. See the ATOP coding manual (Leeds 2005) for additional reservations.

167. Compare Rawls and his typology of war with UCDP.
Rawls distinguished between different types of wars in order to define the principles that would best apply in each case. The initial typology found in his lecture notes proposes nine kinds: 1. Wars between existing via states (WW I + II) 2. Civil wars (of social justice) within via states or society (French Rev); 3. Wars of secession of minorities within region: American Civil War. 4. Colonial Wars of secession (from Empire): Algerian War; American Rev War? 5. Wars of intervention (humane intervention) 6. Wars of national unification (War of Roses; Tudors) 7. Wars of conquest, of Empire (Wars of Rome). 8. Wars of Crusade, religious or secular 9. Wars of national liberation (in present sense); guerilla wars (Armitage 2017, 335)

168. Concepts with adjectives often came become typologies. For example, gender quotas come various forms, “party,” “legislative,” “constitutional,” “reserved seat” etc. which could generate “types of gender quotas” (Krook 2014). The goal of gender quotas is to raise the number of women in parliament. As such the typology is really a series of causal mechanisms designed to raised the number of women in parliament. Discuss why the rule of mutually exclusive categories for typologies is particularly a bad idea in this concept.

Answer:

Clearly a country could, and maybe should, adopt multiple mechanisms for raising the number of women in parliament. In addition there might be synergies between these mechanisms when used together.

169. The well-known UCDP data projects have the goal of basically mapping all forms of armed conflict, both domestic and international. Discuss the extent to which their typology is mutually exclusive and exhaustive. Are there some forms of armed conflict that are not covered by their scheme? Note that this is a tricky and complicated issue.

Organized violence 1989–2017 For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as ‘organized violence’. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535; see https://www.pcr.uu.se/research/ucdp/definitions/)
Answer:
For example, one form of conflict not covered by UCDP datasets is that between organized crime and the government. Armed conflict between organized crime groups is covered.

Collier et al. 2008 use the following example of a typology. Critical in thinking about typologies is the scale type of the the row and column variables. This table clearly has ordinal variables. How would you change the format to reflect better the ordinal nature of the scale to make it similar to the familiar Cartesian coordinate system?

Figure 2: Typology dimensions and orientation

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ambiguity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Administrative Implementation</td>
<td>Political Implementation</td>
</tr>
<tr>
<td>High</td>
<td>Experimental Implementation</td>
<td>Symbolic Implementation</td>
</tr>
</tbody>
</table>

*Source: Adapted from Matland (1995).*

Answer:
To make the ordinal nature clearer and to make it similar to the Cartesian system, one should flip “high” and “low” on the Y-axis (ambiguity variable) so that low is near the origin.

Traditional typology rules have the core mutually exclusive and exhaustive rules. This is an issue in concepts of ethnic group. Suppose there are two ethnic concepts “white” and “black.” In fuzzy set one could be a .4 member of white AND .4 member of black. Discuss the extent to which various datasets on ethnicity (e.g., Fearon 2003; Chandra 2008) assume coding of ethnic groups is exclusive. Discuss how this plays a central role in the widely used ethnic fractionalization index: 

\[ F = 1 - \sum_{i=1}^{n} p_i^2 \]

where \( p_i \) are the population shares of each ethnic group.

Answer: Because fuzzy logic does not think strict dividing lines are good practice, it naturally will allow, and even encourage, partial membership in multiple ethnic groups. Because it is not probability based, there is no need for things to add up to one.
The fractionalization index assumes that membership in ethnic groups is (1) zero or one, and (2) that you cannot be a member of two ethnic groups at the same time. In short the sum of the population shares must be one.

Statistical issues, latent variables, etc.

172. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when $X_1$ is a subset of $X_2$. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace: “This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies” (Owsiak and Vasquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in $X_1$ is taken up in $X_2$? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if $X_1$ is a perfect subset of $X_2$ and one introduces an interaction term, $X_1 \times X_2$?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as $X_1$ less than or equal to $X_2$ for all observation or vice versa).

173. Discuss the causal figures in Multimethod research, causal mechanisms, and case studies: an integrated approach, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

174. Discuss Pahre’s claim that necessary conditions violate common statistical assumptions: “Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that ‘if two units have the same value of the key explanatory variable, the expected value of the dependent variable will be the same’ (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition $x \leftarrow y$; in other words, x may occur with or without y” (Pahre 2005, 131).

175. Draw a figure describing the three-level character of the Hicks-Kenworthy (2003) concept of the welfare state. Compare the theoretical discussion of the dimensions of the welfare state with the results of the principal components statistical model. What are the possible options for the secondary-level dimensions?
176. For German speakers. Opp (2005) is a very rare methods textbook that has devoted significant attention to concepts – most textbook only look at measurement. It is also a rare textbook that devotes a whole chapter to teaching logic – most research design and methodology textbooks do not discuss this at all. In Opp’s discussion of concepts (chapter 4) he distinguishes between (1) “Analytische operationalisierung” and (2) “empirische operationalisierung.” Discuss the extent to which (2) corresponds the the standard latent variable model of concepts and (1) corresponds to what the concept book calls ontological concepts.

177. Use the following quote to discuss the importance of predictive power in terms of evaluating a measurement model or concept. Contrast this with an approach which stresses the semantics and content of the test items.

“The g factor is an artifact of linear correlation analysis. A theorem of Suppes and Zanotti (1981) informs us that for any vector of test scores from an achievement test, it is possible to construct a scalar latent factor such that, conditional on the factor, test scores are independent. The g factor exists for any vector of binary, finite-valued, or countably valued random variables. The g of conventional psychometrics is a product of mathematical conventions in factor analysis. A g also exists to account for correlations among test scores. That is a mathematical theorem of no behavioral consequence for psychometrics or for finance, another field addicted to factor models. The value of g in predicting behavior is the real test of its importance. There is much evidence that it has predictive power. (Heckman 1995, 1105)”

178. Recently statisticians have become much more concerned with problems of “unit homogeneity.” Here is Henry Brady defining the idea:

We shall make the transformation of $Y_B(1, 0)$ into $Y_A(0, 0)$ in two steps which are depicted on Table 10.9. If $A$ and $B$ are identical and $Z_A$ and $Z_B$ [Z is the treatment] are identical as well (footnote: By saying that $Z_A$ and $Z_B$ have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven’t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1, 0) = Y_A(0, 1)$, [Identicality of units and treatment or Unit Homogeneity]. That is, $A$ and $B$ are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on $B$ is the same as the impact of $Z = 0$ and $Z = 1$ on $A$. (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:

Holland (1986) says “unit homogeneity” means that units are prepared carefully “so that they ‘look’ identical in all relevant aspects” (Holland 1986, 948). To be “identical” means that the concept must produce identical units, $A$ and $B$ in the Brady analysis. If a study looks at all “states” then the conceptualization of state means that there are no casually important differences in “stateness.” Treatments, $Z$ in the Brady analysis, also are assumed to be homogeneous.

179. A key issue in the three-level framework is whether the same data-indicators can be used for different secondary-level dimensions. See how this occurs
in “Measuring the rule of law”: “A careful examination of the nine factors reveals two facts. The first is that there is a partial overlap among sub-factors; that is, a sub-factor can simultaneously belong to different factors at once. This is simply to reflect the fact that various rule-of-law dimensions partially overlap in practice” (Botero and Ponce 2011, 16).

180. Concept generalization across groups is a core methodological issue. Discuss the sense of “meaning” in the abstract and how it relates to generalization of a concept across groups.

Measurement invariance assesses the (psychometric) equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups or across repeated measurements. . . . Appropriate and proper comparison of a construct between groups or across times, therefore, depends first on ensuring equivalence of meaning of the construct. The untoward consequences of measurement noninvariance can be readily illustrated in the study of depression in men and women. Suppose frequency of crying, weight gain, and feelings of hopelessness are indicative of the severity of depression in women, but only feelings of hopelessness are indicative of the severity of depression in men. If the three indicators are combined into a scale to compare depression in women and men, mean differences on the scale may mislead because crying and weight gain have little relation to depression in men. In this example, men may score lower than women on the depression scale because they cry less and gain less weight. (Putnick and Bornstein 2016, 71)

181. Discuss whether a causal interpretation should be given to the latent variable “coup risk.”

My key independent variable is coup risk. . . . But why should we believe that all three items – civil-society strength, political legitimacy and recent coups – tap the coup risk equally well? . . . I then modeled the logit of a latent coup risk, logit(\(q_{it}\)), as a linear function of several willingness and capability indicators (\(z_{it}\)). Specifically, \(z_{it}\) is a vector of country-year characteristics that impact the military’s willingness and capability to organize a coup and thus are plausible sources of variation in the probability of coups. (Sudduth 2017, 9)

Normative issues

182. Explain why this is a normatively problematic conceptualization of “disability.” Explain why being black might be a disability under this conceptualization.

In the ICF, problems with human functioning are categorized in three interconnected areas:

- impairments are problems in body function or alterations in body structure – for example, paralysis or blindness;
- activity limitations are difficulties in executing activities – for example, walking or eating;
• participation restrictions are problems with involvement in any area of life – for example, facing discrimination in employment or transportation.

Disability refers to difficulties encountered in any or all three areas of functioning… Disability arises from the interaction of health conditions with contextual factors – environmental and personal factors as shown in the figure below … The ICF is universal because it covers all human functioning and treats disability as a continuum rather than categorizing people with disabilities as a separate group: disability is a matter of more or less, not yes or no. However, policy-making and service delivery might require thresholds to be set for impairment severity, activity limitations, or participation restriction. (WHO-World report on disability 2011, 5)

183. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the “one-drop rule”?

The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the census’s history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

184. The concept of “transitional justice” must be normative. How are Olsen et al. dodging this issue by including “amnesty” as justice? “Still others value amnesty as a mechanism that acknowledges wrongdoing, but releases societies from the trap of the past. Rather than making normative assumptions about the appropriate form of justice, our dataset recognizes that societies adopt a variety of mechanisms to engage their violent pasts and allows scholars to include or exclude specific mechanisms based upon their own operationalization of transitional justice.” (Olsen et al. 2010, 805)

Answer:

It might be better then to call the dataset the “Transitional practices” dataset. Also, the justification for amnesty is almost never normative but pragmatic.

185. Braybrook (1987) discusses what he calls the Principle of Precedence. The Principle of Precedence means that when Alan needs something that Brenda wants but does not need, then meeting Alan’s need is prima facie morally preferable to satisfying Brenda’s desire. This suggests that needs have precedence over luxuries. Explain the moral philosophy involved. What does this suggest in terms of an aggregation principle? Discuss the moral philosophy behind this principle.

Answer:
The weakest link principle then suggests that for any good you look at each person’s need-luxury score on the good and give it to the person with the minimum (i.e., most needy).

**Global indices, barometers, etc.**

186. The Possibility Principle is closely related to scope conditions. The welfare state literature only looks at wealthy countries. Much of the literature on well-being focuses on poor or destitute countries (Dasgupta 1990). When Dasgupta examines the relative destitution of countries he excludes the well off ones. He uses the Borda Rule (Dasgupta and Weale 1992) to aggregate a variety of indicators of well-being, such as infant mortality, literacy, per capita income civil and political rights. The Borda Rule ranks countries on each dimension then sums the ranks (see Fine and Fine 1974 for an extensive and technical discussion of its formal properties). Explain why this concept of human destitution or well-being is sensitive to the exclusion of wealthy countries. For example, the comparison of India and China will significantly be affected by the inclusion or exclusion of nonpoor countries.

Answer:

As Dasgupta and Weale note, the Borda Rule is sensitive to the population chosen because the distance between a pair of countries depends on the population. In particular, if an alternative population puts countries in between a given pair their overall rankings (and hence the comparison between the two) can change. Particularly with regarding to civil and political rights the inclusion of wealthy countries can change the comparison because these are scattered throughout the ranking.

187. A very important theoretical and policy issue is the evaluation and counting of the number of “poor” people in the world. The World Bank is very involved such efforts (e.g., Chen and Ravallion 2001) and concludes, controversially, that amount of poverty has decreased over the last 20–30 years. Typically the World Bank model uses an “international poverty line,” typically an income of $1 US per day. Incomes in local currencies must then be converted into US dollars to evaluate the extent of poverty. Purchasing power parities (PPP) are used to do this. These are general figures based on a “basket” of goods and the cost of that basket in different countries, and are often used to make GNP/capita comparisons. These PPP values then provide the conversion rates of local currency into US dollars which make world-wide evaluation of poverty possible. Discuss how when the basket of goods consists of a wide range of goods that it might be problematic when the goal is to evaluate poverty.

Answer:

See Reddy and Pogge (2005) for an extensive discussion. Basically, PPP data use a wide variety of goods that are irrelevant for assessing poverty. PPP depends on the relative price of things like, say, SUVs, which are not relevant
for assessing poverty. The relative price of SUVs influences PPP data which may distort the relative costs of things that are really important to poor people, like food. For example, it may be that bread is relatively more expensive in poor countries than in rich (this is often true in poor neighborhoods in the United States compared to middle-class ones). If this is generally true between countries then PPP underestimates the extent of poverty in the world. Hence one would ideally like a PPP that only uses goods that are consumed by poor people, such as food, housing, and medical care.

For the mathematically inclined, one can easily use Reddy and Pogge (2005) to devise an exercise that uses basic calculus (i.e., derivatives) to make the point of the exercise.

**Gender**

188. In the intersectionality literature on often hears that say, black-women, are “qualitatively distinct.” Discuss what this might mean. Might this be more a causal claim than a conceptual claim? To what extent is intersectionality about hybrid concepts?

189. Concepts with adjectives often came become typologies. For example, gender quotas come various forms, “party,” “legislative,” “constitutional,” “reserved seat” etc. which could generate “types of gender quotas” (Krook 2014). The goal of gender quotas is to raise the number of women in parliament. As such the typology is really a series of causal mechanisms designed to raise the number of women in parliament. Discuss why the rule of mutually exclusive categories for typologies is particularly a bad idea in this concept.

Answer:

Clearly a country could, and maybe should, adopt multiple mechanisms for raising the number of women in parliament. In addition there might be synergies between these mechanisms when used together.

190. Examine a concept–measure for democracy. Which secondary-level dimension is most likely to be affected by gender considerations? Answer:

Almost all concepts of the democracy have “participation/voting” as a secondary-level dimension. Almost universally, the concept involves principles like participation in general, but at the indicator level it is almost always white male participation. Paxton (2000) discusses extensively the mismatch between concept and measure:

Although there remains some debate about the definition of democracy, most scholars would agree that a democracy has at least three components: competition, participation, and civil liberties. I focus on the second component of democracy – participation – which is the source of the operational mismatch in current research. Participation is central to most definitions of democracy. For example, Diamond, Linz, and Lipset (1990:6–7) explain: “democracy . . . denotes . . . a highly inclusive’ level of political participation in the selection of leaders and policies, at least
through regular and fair elections, such that no major (adult) social group is excluded.” Modern theoretical definitions of democracy stress numbers of people rather than “types” of people. Most authors today consider the criterion for democratic inclusion to be “all native-born adults” (Paxton 2000, 93-94).

When conceptualizing government action through a gendered lens there are two different approaches to identifying policies that take an explicitly feminist approach and assessing the degree to which public policy is gendered. In assessing government action that takes an explicitly feminist frame, in other words that promotes women’s rights and strikes down gender hierarchies, a series of subsectors have been identified, e.g., blueprint, political representation, equal employment, reconciliation, family law, body politics, reproductive policy, and sexuality and violence to name some of the major subsections. Conceptualizing gender-specific public policy is a much broader endeavor and involves identifying policies that have a formal gendered content or goal but may not be necessarily feminist, e.g., target men and women’s roles but may not be feminist. The study of gendered public policy may also involve conceptualizing how policy implementation and impact may affect men and women’s lives and roles.

Questions and Issues:

- How are feminist policies differentiated from gender-specific policy e.g., reconciliation policy versus family policy?
- When is a feminist policy, actually feminist?
- Given the Western-focused nature of feminism, can feminist policies be studied outside of the Western countries?

Selected References:


Pollock, Mark and Amy G. Mazur (eds). 2009. Gender and Public Policy in Europe. special issue of Comparative European Politics. 9(1).

First introduced as a concept by the United Nations in the 1980s and then through the European Union via the Treaty of Amsterdam in the 1990s, gender mainstreaming has come to be identified with the systematic introduction of gender equality across all policy areas both in terms of formal policy statements and policy implementation. Gender mainstreaming is associated with specific policy instruments that compel governments to systematically introduce gender equality statements and principles in budgets, sectoral plans, etc.
Seldom found as a policy approach in the USA, gender mainstreaming tends to mean introducing gender components in curricula in the US context.

Questions and Issues:

- To be considered gender mainstreaming does this policy approach have to go beyond formal statements to actually promoting gender-based equality?
- To what extent does this concept travel outside of the European Union to countries, like the USA, where gender mainstreaming in policy is not a key approach.

Selected References:


Thanks to Kara Ellerby for this exercise. Gender quotas are considered a “fast track” policy aimed at increasing women’s representation within the legislative electoral process. They have gained immense popularity over the last thirty years, and are currently present in over one-hundred countries. Conceptually, quotas are best understood by two key characteristics: the degree to which they mandate women’s access to public office (is it specified in the Constitution or did political parties create their own quotas?) and at which stage of the electoral process the quota mandates women’s representation (does it specify how women are listed on candidate lists or are there reserved seats for women in parliament?). Quotas vary greatly within these dimensions, but there seem to be patterns among regions. Gender quotas entail important discussions about sex/gender distinctions, descriptive and substantive representation issues and whether or not they really empower women.

Questions and Issues:

- What are the goals of gender quotas? Is it, to get more women in, provide women (and which women?) with more political empowerment, or both? How do secondary dimensions shift if scholars distinguish different goals?
- How does one theorize/operationalize a ‘successful’ gender quota?
- Are gender quotas meant to be a short-term or long-term policy?
- Is the use of ‘gender’ rather than ‘sex’ depoliticizing the very issues quotas are in place to address?
Table 2: Gender quota concept

<table>
<thead>
<tr>
<th>Quota Type</th>
<th>Key Features</th>
</tr>
</thead>
</table>
| Party Quota  | Adopted voluntarily by political parties  
(1) Set out new criteria for party candidate selection.                                                                                         |
|              | (2) Affect composition of party lists in PR electoral systems and candidates eligible for particular seats in majoritarian systems.             |
|              | (3) May entail internal party sanctions for non-compliance.                                                                                   |
| Legislative Quota | Mandated by national parliaments  
(4) Set out new criteria for party candidate selection.                                                                                       |
|              | (5) Affect composition of party lists in PR electoral systems and candidates eligible for particular seats in majoritarian systems.             |
|              | (6) Usually entail sanctions for non-compliance.                                                                                                |
| Soft Quota   | Adopted voluntarily by political parties.  
(7) Set out informal targets and recommendations in relation to party candidate selection.                                                        |
|              | (8) Set out new criteria for membership of internal party bodies.                                                                          |
|              | (9) Rarely entail sanctions for non-compliance.                                                                                               |

Source: Krook, M. et al. 2014

In the table above discuss how adjectives are used to modify “gender quota.”  
If one were to develop a single concept of gender quota instead of a typology, what would be the secondary-level dimensions based on the key features?  
Can you draw a diagram of the unified gender quota concept?

Selected References:

“Sexual violence,” “gender violence” consists of the operation of adding an adjective to a standard concept. The concept focuses on forms of sexual violence that occur particularly often to women, such as rape or domestic violence, although more recently gender violence has been associated with homosexuals and transgendered groups. For this concept there is some variation on names. The United Nations, as well as the feminist literature on the topic, in most of its documents uses “violence against women.”
Questions and Issues:

- When should the notion gender violence be used as opposed to violence against women?

Selected References:


195. Gendering the concept of institutions is highly complex and involves the premise that institutions as organizations in society and in the state are inherently gender-biased. For some this is a question of research for others it is an underlying assumption. Bringing the study of gender into institutional analysis also implies that feminist agents of change, what some call feminist engagement with the state, will attempt to change the underlying gendered/patriarchal nature of institutions. The ultimate outcome in the process of these feminist engagements is that institutions reflect cultures of gender equality and account for the improvement of women’s right in relation to men. The study of gender and institutions is often conducted in the context of seeking to gender new institutional scholarship. Questions and Issues:

- To what degree is this a discrete analytical concept – e.g., gendered institutions – rather than an approach to institutional analysis?

Selected References:


196. A central concept to feminist analyses of politics, interestingly many empirical feminist analyses conducted in the purview of political science do not actually use the term, opting instead for notions of gender-bias, gender-based discrimination etc. In addition, while for many feminist theorists, patriarchy is an operating assumption of any organization or structure, for analysts who seek to conduct empirical studies of gender and politics the presence of patriarchy or not, is a question for future research. Nonetheless, the notion of patriarchy contains conceptual and theoretical ramifications that are important to the systematic study of gender and politics. While there are many articles
on patriarchy and works that discusses patriarchy there are few that provide operational research definitions of the concept. Walby (1989) provides some useful perspective on operationalization, measurement and traveling.

Questions and Issues:

- If patriarchy is an underlying assumption of every organization, how then do researchers operationalize and measure it?
- To what degree does the notion of patriarchy travel outside of western contexts?

Selected References:


Analyzing power is core to the gender and politics literature given the notion that politics is traditionally organized to organize out gendered concerns and power relations are biased against women see patriarchy, for example. At the same time and as a result of its omnipresent nature, it is quite difficult to operationalize power from a gendered perspective for empirical research. For those who are interested in reading further on power and feminist analysis and an excellent overview of feminist theories of power we suggest the classic books by Judith Squires and Connell on the topic. Given the degree to which patriarchy is interlinked to power also see the above entry on patriarchy.

Selected References:


An oft-nebulous and contested concept in the study of comparative politics, social capital is related to the study of norms and values about civic engagement, quite often in a cross-national perspective. Gendering the concept, therefore, faces the same issues of imprecision and poor operationalization. Nonetheless, at least for feminist analysts, bringing in the interface between the social construction of men’s and women’s identities in relation to one another with the development of participatory engagement is an empirical and theoretical imperative. We list several works that place gender on the agenda of comparative work on social capital.

Selected References:


199. Linked to the concept of democracy, this concept is another example of an established concept engaging with gender. Here, the notion of regime transition from and to democracy undergoes gendering. Thus a more complex view of democratization is dealt with here than in the chapter on democracy in the book, with a more global frame that provides conceptualization that goes beyond Western democracies. Women as objects and actors are brought into the conceptual mix as well as taking on seriously gender as an analytical concept in the drivers and dynamics of transition. This gendered take dialogues with the large “transitiology” literature in comparative politics. Whether this body of feminist literature provides a structured approach to the concept of gendered transitions remains to be seen given the early stages of the concepts genesis. Nonetheless, the selection of literature cited here provides a road map.

Selected References:

200. For many obvious reasons war has been seen as a highly gendered activity. War, and more generally “national security” have been a focus of feminist scholars, not surprisingly within international relations. The organization of militaries, the conduct of war, discourses of national security, and many other topics have been the focus of a gender analysis.

It is useful to consider the nongender concepts of war which usually contain two core sets of characteristics (1) nature of the actors, usually one of the actors must be a state, and (2) severity of militarized violence (typically numbers of soldiers killed in battle). One issue is how or should one include gender the basic concept of a war. Much work is concerned with gender and the causes of war, the conduct of war, or the impact of war. As an exercise it is useful to survey the literature to see which of these has been the subject of a gender analysis, and how the gender analysis is conducted. One development has also been to move away from war to a more general analysis of “human security” (http://www.hsrgroup.org/). This has important implications for a gender analysis since it naturally brings in other issues such as gender violence during wars, health, etc.

Selected References:


201. “Women’s interests” is central to the analysis of gender and politics. How do researchers define women’s interests as opposed to men’s; in other words a gendered approach to the notion of interests. More than dialoging with the concept of interests in nonfeminist political science which is quite diffuse, the notion of women’s interests, sometimes referred to as gender interests, was developed separately in the literature on gender and politics. From Sapiro’s first treatment of the issue to Molyneux’s efforts to differentiate between strategic and practical interests, conceptualization of women and/or gender interests involves identifying what, if any are the common political interests of men as opposed to women. A central component of developing operational definition of gendered interests has been the diversity of women and men as groups-across race, class, ethnicity, sexual orientation, religious affiliation etc. Thus, conceptualization of women’s interests have increasingly become more sophisticated over-time, taking on the reality of intersectionality in political mobilization.

Selected References:


Case studies, case selection

NOTE: the next section on the Possibility Principle is all about case selection as well. See also the multimethod section for case selection exercises.
202. There exists a lot of debate about rules for case selection in multiple case study research. Discuss these principles of case selection:

For the purposes of differentiating different theories and visualizing the operation of mechanisms singled out in our theory, we have chosen our cases according to seven principles. Three of the seven principles (2, 3, 7) are drawn from Goertz and Mahoney, Gerring, and Weller and Barnes. purposes here (1, 4, 5, 6).

The seven principles are as follows: (1) Cases must have been well documented, so that their basics are not subject to disputes allowing potential accusations of cherry-picking within-case facts. The exception here is the true negative case of Gabon as a “dog that did not bark.” This case has not received much scholarly attention. (2) Cases must include both positive and negative cases with real possibility of positive outcomes (for example, conflict) according to one’s own or others’ competing theories. (3) Cases must include pathway cases (or typical cases) that exemplify the variations of the key explanatory variable(s), thus allowing clear differentiation of competing theories regarding them. (4) Cases must include pathway cases that exemplify the variations within the operation of the core mechanisms, thus allowing clear differentiation of competing theories regarding core mechanisms. (5) To compare the strengths and weaknesses of quantitative and qualitative methods, cases should include those that can be easily miscoded or misidentified in quantitative exercises, even with GIS datasets. (6) Cases should come from different geographical and cultural backgrounds. (7) Together, the cases examined should provide a fairly generable picture about factors and mechanisms across the population or universe of cases. (Tang et al. 2017, 369–70)

203. Explain how Howard is using the Avoid Overdetermination Rule in selecting her cases:

Choosing the population based on American-initiated solutions is the best way to test for internally driven hypotheses of US foreign policy, since such decisions are a result of dynamics within the United States and not of multilateral deliberation. The vast majority of US foreign policy literature assumes or argues that internal US dynamics drive decisions. In other mediation efforts in ethnic conflicts (e.g. Kosovo, Macedonia, Afghanistan, and Kenya), the United States worked very closely with international partners. Thus, multilateral, coalition behavior, not internal, domestic decisions, often drove foreign policy. It is necessary to choose cases where the leading hypotheses “most likely” determine the outcomes. In other words, the domestic audience-focused, unilaterally driven population would have the most likely cases of decision points where the leading hypotheses would hold. (Howard 2017, 725)

204. Discuss case selection in Stasavage’s influential book Public debt and the birth of the democratic state France and Great Britain, 1688–1789. Clearly he is using the very popular paired comparison, in terms the causal mechanism book a (0,0) case compared to a (1,1) case. There are lots of nondemocracies in this. Why choose France? What about choosing more (1,1) cases to generalize?

Representative political institutions may improve a government’s ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide
support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

205. Selecting on the dependent variable is very common in trying to figure out the causes of something. Discuss how the World Bank is selecting on the dependent variable to figure out the causes of violent conflict.

This chapter explores the accumulation and intensification of risks and opportunities in critical spaces, called arenas of contestation. These arenas involve what groups care about in their relationships with each other and with the state and thus what they tend to fight over—access to power, land, and resources, equitable delivery of services, and responsive justice and security. These four broad arenas are by no means an exhaustive list, but they have been selected because they have consistently recurred in violent conflict in various contexts. ((World Bank 2018, 141, Pathways for peace, chapter 5)

206. Gerring and Cojocaru 2017 describe what they call “algorithmic” case selection, often using a regression model. Discuss if the italicized elements below are necessary and required or not.

Algorithmic case selection follows a set of rules executed in a sequence of steps, which we envision as follows. 1. Define the research question and the population of theoretical interest.
2. Identify a sample of potential cases. Ideally, this sampling frame should be representative of the population of interest.
3. Measure relevant features of the cases—for example, D, X, Y, and/or Z—across the sample.
4. Combine diverse indicators of D, X, Y, and/or Z into indices, if necessary.
5. Construct a causal model, if required. (Gerring and Cojocaru 2016, 411; emphasis is mine)

207. When doing multiple or comparative case studies when there are multiple hypotheses involves complicated and potentially tricky case selection decisions. Rocabert et al. 2019 have three hypotheses regarding international parliamentary institutions and IGOs and ends up doing two case studies. Draw the table of all the possible values, presence–absence for all variables in the three hypotheses and then locate the case studies within this list of all possible cases
one might choose based on this table. They chose (1) ECOWAS and (2) PIF. Are those the best choices?

208. Discuss the extent to which Alter is using the Avoid Overdetermination Guideline in her selection of cases: “I generally pick hard cases, situations where important interests are at stake, where the policies leading to law violations are politically entrenched, and where law operates in places and ways that are counter to the expectations of international relations scholars and domestically focused lawyers.” (Alter 2014, 24)

209. Shannon 2009 is a case study of the US invasion of Panama in 1989. If one were to do a large-N qualitative analysis starting from this article what would be the basic case selection criteria using figure 2. Should you consider case selection further down the tree? Why is that important?

Answer:

The top of the decision tree is those leaders who are motivated to violate the norm is the minimum case selection rule. These are the cases where we can see the norm potentially in action.

210. Selecting on the dependent variable is legitimate for necessary condition hypotheses. Explain why.

Answer:

For a clear discussion see Dion (2003) or Harvey (2003). The short answer is that a necessary condition hypothesis says that we should see X present when Y occurs. Hence the test starts by looking at when Y occurs, i.e., selecting on the dependent variable. See Harvey for a discussion of how this has been ignored in the literature on selection and deterrence (e.g., Achen and Snidal 1987). Dion (2003) uses Skocpol as one of his main examples.

211. Causal mechanisms and scope decisions can be closely related. Bush’s main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.
212. Collier and Collier (1991) is a classic of comparative historical analysis. Central to their work is the concept of labor incorporation. Discuss their case selection on this variable and whether it should be a partially a scope condition.

Because their interest is in the effects of labor incorporation, they [Collier and Collier 1991] consider only the effects of variation in the productive condition: all eight of their cases saw labor incorporated into politics after a major social crisis in the early 20th century. Their interest is in the ways in which different “values” of the productive condition shape divergent regime outcomes over subsequent decades. Since the permissive condition of the rise of labor bounds the theory of regime dynamics that they develop, it acts as a scope condition on the proposed relationship between the mode of labor incorporation and regime outcomes. (Soifer 2012, 1581–82)

213. Discuss Huber and Stephens’s (2012) case selection strategy, particularly its longitudinal variation on X and Y. Did they choose cases that were (1,1) cases at some point in their history? Note as well that their core theory involves an interaction term.

Of the various purposive strategies, we follow Ragin’s (1987, 2000, 2008) advice to focus on the positive outcome cases, so we select all cases that had high values on the dependent variable as of 1980. This fits Goertz’s (2008) admonition to choose diversely among positive outcome cases and Lieberman’s (2005, 444) to select for wide variation in the explanatory variables among the cases that fit the theory according to the quantitative analysis. (pp. 49–50)

214. Ostermann (2007) looks at compliance with policy regarding wood harvesting in protected areas in a park which lies on the border between India and Nepal. She is interested in the effect of state policies on compliance. Discuss her case selection discussion in light of the Avoid Overdetermination Guideline. Discuss in particular the key confounding variables of “weak state presence” and “poverty” which can explain compliance or noncompliance.

The Terai region in the South of Nepal and the North-Indian state of Bihar is an excellent place to test compliance theories in adverse circumstance for two main reasons: (1) it fits squarely within a least favorable case research design—the states on both sides of the border are weakly institutionalized, at least in this region; and (2) the populations on either side of the border are generally poor and uneducated. Indeed, for much of the period leading up to this study, the states on both sides of the border had largely retreated from this region in response to armed Maoist movements. Case selection was also part of an effort to use design to gain methodological traction on persistent questions in the socio-legal compliance literature. Chitwan and Valmiki are contiguous and divided only by an open international border that was not defined by geography or culture. The same culture and dominant ethnic group—the Tharus—prevail on both sides of the border, and most people in this region are poor, fairly uneducated, and speak the same language. (Ostermann 2006, 107)

215. The medium-N paradigm means exploring cases in border areas of the (1,1) zone as secondary case studies. For the democratic peace this would be borderline democracies or borderline peace (in the joint democracy is sufficient
for peace version of the democratic peace). An interesting exercise to see what
the cases scholars have used when critiquing of the democratic peace, e.g.,

216. Norris in Making democratic governance work is interested in the impact of gov-
ernance and/or democracy on a variety of outcomes, e.g., economic growth,
welfare, and conflict. She also uses case studies to illustrate her findings and
approach. In the various chapters on various dependent variables she uses
paired comparisons. Discuss how she decides about these pairs focusing on
the problem that she always has two independent variables – governance and
democracy – in each chapter. For example, does she use the Avoid Overde-
termination Guideline?

217. Copeland in an influential study of trade and militarized conflict makes claims
about the democratic peace. Explain how his case selection (all major power
crises and wars) does not include key cases for evaluating the democratic
peace.

The trade expectations approach also helps us understand exactly what
is at stake in the ongoing debate about the democratic peace. As we saw
in chapter 2, it may well be the case that the correlation between mutual
democracy and peace is really reflective of an economic peace as opposed
to a political one. Democracies are unlikely to fight each other, in other
words, not because they respect each other’s normative values or because
their legislatures pull illiberal leaders back from the brink, but rather
because democracies generally have open liberal economic foundations
and thus are able to signal their commitment to open-door economic
policies into the future. (Copeland 2015, 434)

Answer:
By selecting only major-power crises and wars he does not include the non-
barking dogs, where democracies are at high levels of peace.

218. Seawright (2016) argues for “extreme on X” case selection in multimethod
research. Extreme on X could mean very high or very low or distant from
the mean. Contrast that with “good” causal mechanism case selection as
argued in Multimethod research, causal mechanisms, and case studies: an integrated
approach, i.e., some extreme cases are good cases and others are not good cases.

219. Horowitz table 2.4 gives a list of all possible case studies: “The major military
innovation cases evaluated in this study are the following: early twentieth-
century battle fleet warfare, mid-twentieth-century carrier warfare, nuclear
weapons, and suicide terrorism. As table 2.5 shows, this selection of [four, one
per chapter] cases maximizes variation on the two key independent variables:
financial intensity and organizational capital. It also allows for significant time
variation, over a period of a century, and cases that focus on both nation-states

If you were going to do only two case studies and followed the Avoid Overde-
termination Guideline which cases would you choose?
The Possibility Principle


220. Should one assume that all countries are at risk of a coup or apply the Possibility Principle in the study of coups? See for example “The data are in time-series cross-sectional format and the unit of analysis is the country-year. The data include 200 countries for the period 1968–2003. (Sudduth 2016, 7)

221. A very difficult conceptual and methodological problem is dealing with norms, particularly norms that have the logical structure of rights, e.g., human rights. This problem has exercised me a great deal and in fact resulted in a book (Goertz *International norms and decision making*; a large part of that book was devoted precisely to analyzing the concept of a norm). (1) Discuss the concept of a right (contrasting it with prohibitions and obligations). (2) The positive and negative poles are interesting here in terms of behavior: do you focus on the positive exercise of rights (e.g., voting) or on the absence of the exercise of the right. This has theoretical consequences because in the former case you focus on what factors in addition to rights lead to the use of rights, while in the latter you focus on factors that eliminate the right itself in practice. (3) Rights are about possibility, relate this to the Possibility Principle of case selection and necessary conditions in general. Dowding and Hees (2003) is a very nice discussion of many of these issues. For measurement see also Rosenbaum (2000) and Carter (2000).

222. Morrow and Jo illustrate how the Possibility Principle can lead to specific coding decisions: “Violations are not possible for some issues in some cases because the sides either lacked the capability to carry out such violations or no fighting of the type in question occurred during the war. All the fighting in the Hungarian-Allies War of 1919 was on land, making violations of conduct on the high seas impossible.... These cases receive missing value codes of –9 for all five measures of compliance” (Morrow and Jo 2006, 99). Discuss whether these cases should be coded as “missing” or something else.

Answer:

It would seem to confuse cases where the data is really missing with cases where the coding scheme does not apply.

223. Explain how Chandra and Wilkinson’s (2008) apply Posner’s (2004) principle of *politically relevant* ethnic groups in their two measures ECI and EVOTE.

Answer:

Both these measures involve how ethnic group distinctions are politically mobilized in some manner other another, i.e., they have been made politically relevant by some actor(s): “At the broadest level, we can imagine the term ethnicity as encompassing two families of concepts – the structure of ethnic identities and the practice of ethnic identification. Ethnic structure refers to
the distribution of descent-based attributes—and, therefore, the sets of nominal identities—that all individuals in a population possess, whether they identify with them or not. Ethnic practice refers to the act of using one or more identities embedded in this structure to guide behavior. In other words, it refers to the set of activated identities that individuals employ in any given context. The set of activated ethnic identities for any given country is typically a subset of the identities contained in the ethnic structure” Chandra and Wilkinson (2008, 523).

224. McAdam and Boudet explore why environmental social movements arise. Discuss how they use the Possibility Principle to select communities “at risk” of an environmental social movement.

That is, instead of selecting successful movements for study, we chose to research communities communities “at risk” for mobilization by virtue of their shared exposure to the “threat” of environmentally risky energy projects. (McAdam and Boudet 2012, 179)

Because all large, potentially controversial, infrastructure projects are required to file an EIS, these records provide a population of communities at risk for mobilization. Thus our cases were drawn from the CSA Illumina’s Digests of Environmental Impact Statements, which contains all EISs from the Federal Register. The population of cases was limited to proposals for new energy infrastructure projects in which a Final EIS was completed between 2004 and 2007 (N = 49). We chose projects that had completed a Final EIS in to ensure that the window for potential collective action would be closed once the study commenced. After drawing a random sample of twenty cases, three of them (15% of the sampling frame) were dropped because it was not feasible to conduct subsequent data collection due to the difficulty in attaining local newspaper records. The sample was then supplemented with three California projects for which data had already been. The final sample consists of twenty communities responding to eighteen projects in twelve states regarding proposals for LNG terminals (13 cases), nuclear-related projects (2), a hydroelectric project, a wind farm, and a cogeneration project to supply electricity to an existing oil refinery. (McAdam and Boudet 2012, 36).

225. Howell and Pevehouse in their analysis of Congressional impact on decisions to use military force rely on a dataset of “opportunities to use force” (2007, see Appendix B). For example, “we identify a wide variety of opportunities to exercise force abroad. The database inventories the following kinds of crises: attacks on United States embassies and consulates; instances when United States ambassadors, consuls, or military personnel are killed; hijackings that include human casualties; stateside attacks perpetrated by foreign groups; civil wars; interstate armed clashes” (p. 248). Discuss the extent to which the Possibility Principle is used to constitute a “possible use of military force by the President of the USA.”

226. Fordham (2011) looks at the creation of major powers. Explain how using the Possibility Principle would significantly change the case selection research design.

Answer:
He uses all states as potential major powers, clearly there are only a small number of states that are potential major powers.

227. One case selection issue deals with cases some individuals – in this case states – are eligible to sign a treaty (dependent variable) while others are not. Only countries in the region are eligible to sign those treaties. The question is what do in this situation. Alcañiz includes in her study global treaties and regional treaties. Here is Alcañiz’s solution (2012, 325): “the dummy variable Regional that takes the value of 1 if a treaty is open to a restricted menu of members or 0 if it is open to all countries. it is important to highlight that the universe of countries for regional treaties is restricted to eligible members. Valid subjects for the Tlatelolco treaty, for example, will include only countries of Latin America to which the treaty was open to signature.” Is this the best or correct solution?

Answer:
The dummy variable is not correct. For example, all survival models assume that $0 < P(Y = 1) < 1$. Notice that the probability is assumed to be greater than zero, however, we know that the probability that European countries will sign a Latin American treaty is zero.
The simple solution is just to remove the observations of countries not eligible to sign a treaty.

228. While the Possibility Principle is most often used to choose cases, it can also be used to create measures and variables. Describe how Mansfield and Pevehouse use the Possibility Principle in their study of the impact of IGOs on democratization. What role is the Possibility Principle playing in indicator or variable construction?

Our argument centers on the demand for IO membership by democratizing countries. We have implicitly assumed that the supply of IOs is constant across countries. However, not all countries are eligible to join every IO. Consequently, we reestimate the three original models after adding a variable measuring the number of possible organizations that state $i$ could enter in year $t$: This variable is the ratio of the number of IOs that state $i$ participates in to the number of IOs it could participate in, where we assume that this state has the opportunity to join every universal organization (e.g., the United Nations) and every organization in its geographical region (Manseld and Pevehouse 2006, 154). In no case is this new variable statistically significant, nor does including it alter the coefficients of the other variables in the model. (Mansfield and Pevehouse 2008, 279).

Answer:
They assume that every country can joint every universal IGO, similarly they assume that every country in a region can join a regional organization. In practice they may not really be true. For example, can Israel join the Arab League?
The role of the Possibility Principle is to standardize the IGO variable. So instead of the raw number of organization that a country is a member of, it is the number of memberships divided by the number of possible memberships.
229. One potentially interesting new variant on the Possibility Principle uses negative cases on the dependent variable to defining negative cases. If the dependent variable is defined using the necessary condition structure then one can covert the AND to OR as per the Possibility Principle. Then take as negative cases all cases with at least one necessary condition present on the dependent variable. Discuss this idea. (Thanks to Sverre Bodung for suggesting this idea.)

230. Just before he died David Freedman wrote a critique of Possibility Principle (Freedman 2009). Freedman discusses the Rule of Inclusion in terms of a bivariate, $2 \times 2$ table. He states that the Rule of Inclusion is bad practice because it leads to including no cases in the $(X_1 = 0, Y = 0)$ cell. The Rule of Exclusion is typically a bivariate case selection rule, but the Rule of Inclusion is a multivariate principle. Show, using the Skocpol data, how when one has multiple independent variables (i.e., 2 or more) that in general there will be cases of $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$. Explain which cases one will in fact not be included.

Answer:
Using the Rule of Inclusion in the multivariate setting will almost always give cases in the $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$ cells. This occurs because cases of, say, $(X_1 = 1, Y = 0)$ will sometimes have value 0 for $X_2$. In short, following the Rule of Inclusion one typically has cases of $(X_1 = 1, X_2 = 0, Y = 0)$ and $(X_1 = 0, X_2 = 1, Y = 0)$. What one does not have are cases of $(X_1 = 0, X_2 = 0, Y = 0)$. You can see this by looking at the universe of all cases used to test Skocpol's theory which clearly do have cases of $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$.

231. Kroenig has as his main hypothesis “Hypothesis 1: The more powerful a state is relative to a potential nuclear recipient state, the less likely it will be to provide sensitive nuclear assistance to that state” (2009, 116). He uses an “all cases” research design to test this: “Potential nuclear recipients are all nonnuclear weapon states in the international system” (p. 177). Using the Possibility Principle discuss how this research design decision will include many cases which support his hypothesis, but which might be consider irrelevant for such a test.

Answer:
Very small and poor countries might be considered as irrelevant for this test. Since the main hypothesis is about relative power, these cases combined with nuclear suppliers (i.e., powerful countries) will be cases in support of his hypothesis, e.g., France-Fiji.

232. What constitutes “potential mediators” in the study of international conflict mediation, e.g., Frazier 2006?

233. Describe how the Possibility Principle is used in a number of studies that test the influence of rules of war, such as Morrow (2007), Valentino et al. (2006), Downes (2006).

Answer:
Only cases where the state had the capability to violate the rules of war are used. For example, Downes (see the Appendix of his article) only includes states that had the capability to target civilians of the adversary during the war.

234. Collier and Sambanis’s (2005) *Understanding civil war: evidence and analysis* chose negative cases based on “high risk.” What is definition of high risk? Could you choose cases in gray zone on the dependent variable (they use a dichotomous dependent variable)?

235. Describe the concept structure used to construct politically relevant dyads as a population in conflict research. Once you understand the structure could you easily expand the definition of politically relevant dyads?

Answer:
The family resemblance structure is used because it uses a rule of the general form “If any of the criteria is met then include the observation.” In this case the list is (1) major power status or (2) contiguity. You could just extend the list. For example, Quackenbush (2006) adds alliances to the list.

236. There exists a large literature on diversionary war (see Levy 1989 for a survey). Discuss how the Possibility Principle is used or not to (or should be and is not) to decide on the population of cases to analyze. Notice that some use the concept “opportunities” to use force as the unit of analysis (e.g., Meernik 1994, 2000). See Foster (2006) for an explicit look at “opportunities” as a function of major power status. See also the question regarding Andy Bennett’s decisions about opportunities to intervene in his analysis of USSR behavior. Also, contrast the case selection in Ostrom and Job (1986) with that of Meernik (1994) and make the analogy to the politically relevant versus all dyad debate. See Mitchell and Prins (2004) for another way to think about opportunities for diversion.

Answer:
Foster wants to explore the impact of major power status on diversionary use of force. It is hard to find it but a footnote in table 1 indicates he used politically relevant dyads. Hence, he is using the same variable to choose populations as he is looking at in the main statistical analysis.

Ostrom and Job base their analyses on time periods, i.e., quarterly data setup. Underlying this is the view that the US always has an opportunity to use force. They propose and Meernik follows up on the idea that a dataset on opportunities to use force would be a better research design. Notice that the impact of domestic politics (the key finding in Ostrom and Job) disappears in the Meernik study.

237. The problem of relevant case selection arises in the literature on the intervention of third parties in ongoing conflicts. If it is unlikely that minor powers get involved in conflicts in distant regions it is probably even less likely that they intervene in a partisan fashion in ongoing international or civil war. The Possibility Principle eliminates cases from the population under (statistical) analysis. An alternative approach is to model-in a separate equation-those
states that are likely to intervene at all and then look at issues of which side.

One way to think about this is “mixture cure” models:

Thus we estimate, instead of standard survival models, “mixture cure” models (e.g., Spoto 2002; see the online appendix at http://journalofpolitics.org/ for technical details and additional references). These models contain an additional binary regression component to model the probability of the event occurring eventually and to correct the hazard/duration model portion. Estimating the mixture cure model with competing risks confers two major advantages: (1) We avoid having to address the above-mentioned problematic assumption of standard survival models by first defining, then analyzing only “relevant” states. This approach would invariably have resulted in exclusion of cases of intervention by “nonrelevant” states. (2) We are able to ascertain simultaneously what factors affect the probability of intervention, the hazard/timing of the intervention (given the decision to intervene), and the side on which the intervention occurs. (Findley and Teo 2006, 833)

Note that this is similar in spirit to the zero-inflated Poisson models (Chin and Quddus 2003) which attempt to separately model the zero occurrences from the nonzero ones. Discuss these statistical alternatives which separately model the “irrelevant” cases.

Answer:

A key thing to keep in mind in the application of the mixture cure model is any given variable which appears in both equations must considered carefully since the “total” impact is the combination of the two. This may be hard to disentangle because in the variable only influences the hazard rate in the selection equation.

Notice that the Findley and Teo argument also implicitly refers to the problem in the conflict literature where using politically relevant dyads as a population selection device (as opposed to a zero-case selection device) excludes the “impossible happens” cases.

It is important to realize that these statistical approaches as well as the Possibility Principle apply a theoretical or empirical model. One needs to consider which model is better. When the impossible happens is clearly error in the Possibility Principle model. However, it may be the case that zero-case model in the mixture cure performs poorly in modeling the hazard rate.

King and Zeng (2007) have presented a large-N argument about counterfactuals. Basically the article deals with the perils of extrapolation and a technique for deciding when a counterfactual is an extrapolation outside the data. Basically, counterfactuals within the data are more reliable than extrapolations outside it: a bit of conventional wisdom. One conclusion from this methodology is that one should collect more data to increase the region covered by the data; a conclusion not too surprising from a statistical point of view. Use the Possibility Principle cube to think about these issues.

(1) It is quite possible for counterfactuals within the boundaries of the data to be quite distant from actual data points. Illustrate this in the cube.
(2) Use the cube to show how counterfactuals outside the data might be closer to actual data points than points in the data cube.

(3) Use the democracy and war literature to illustrate how interpolation might be more problematic than extrapolation.

(4) Applying the Possibility Principle usually means reducing the size of the sample space, while King and Zeng implicitly argue for increasing it. However, both argue that causal or counterfactuals are best made with contrasting cases that are “close” to each other. Discuss.

Answer:

For (1) and (2) put points at the corners, and a few outside the corners and leave the middle of the cube empty.

The democratic peace literature is paralleled by a small “autocratic” peace literature. But there is a fair amount of evidence that hybrid or transitional regimes are more war prone. Hence interpolation of causal effects may not work in this case. Also, extrapolation to higher quality democracy is probably going to be more valid than interpolation to gray zone democracies.

For (4) the key issue hidden by the King and Zeng article is that one wants to make contrasts with cases that are nearby. Data in the convex hull may be more distant than data outside it. The Possibility Principle says that increasing the convex hull may be counter-productive if the contrasting cases are too far away.

239. The Possibility Principle can play an important role in concept/measure construction when the number of zeros has a large impact on the resulting measure. Explain how the S measure of dyadic interstate preferences (Signorino and Ritter 1999) is sensitive to this when using all dyads versus politically relevant dyads. Why do all the many politically irrelevant cases have an influence?

Answer:

See Sweeney and Keshk (2005) for an extensive discussion. Basically, all the politically irrelevant dyads have similar preferences and hence one gets a huge number of cases with high preference similarity (e.g., the vast majority of cases have a value near the maximum of 1.00).

240. Sometimes the Possibility Principle is related to truncation on the dependent variable which appears in case selection. For example, the dependent variable could be US foreign direct investment (Oneal 1994) or terrorist actions (Li 2005; Enders and Sandler 2005). Sometimes the cross-sectional cases are only those countries with a positive value on the dependent variable, for example, those that actually got foreign direct investment or had a terrorist act over the period in study. Explain how the Possibility Principle relates to this kind of potential selection bias. In practice which would be worse the US foreign direct investment case or the terrorism case?

Answer:

Implicitly the Possibility Principle is used because countries that, for example, did not receive US direct investment are perhaps assumed to be those
where it was in some sense impossible. US legislation forbids investment in some countries, e.g., Cuba. Of course, there might be cases where there is no economic value in such investments but those cases one would want to include. (Note that FDI raw data often have positive and negative values for inflows and outflows. Some authors take the log which then typically means throwing out all the negative and zero cases.)

The issue is probably less severe in the literature on terrorism. As a matter of history (or historical accident if you prefer) there have been terrorist incidents in almost all countries of any size (e.g., about 120 according to Li 2005). Thus the exclusion of countries who have not had a terrorist incident is not likely to have much practical effect. However, if it were the case that only 60 countries had had terrorist incidents then this issue would become important. Some authors include all cases and then try model the zero cases separately with techniques like zero-inflated Poisson (Chin and Quddus 2003).

241. The Possibility Principle can be used in indicator construction as well. A major issue in the political and economic analysis of countries is their degree of ethnic fractionalization. A large literature in economics (e.g., starting with Easterly and Levine 1997) deals with its impact on economic growth. Clearly ethnic variables have played a large role in the analysis of civil wars and genocide. How might the Possibility Principle be used to think about which ethnic groups to include in an index of ethnic fractionalization in ways similar to the way politically relevant dyads are used to study international conflict?

Answer:

Posner (2004) develops a measure of ethnic fractionalization—what he calls Politically Relevant Ethnic Groups (PREG)—based on the logic of the Possibility Principle and politically relevant dyads:

The crux of the problem is that the Atlas data from which the ELF index was calculated enumerates dozens of groups in each country that may be culturally or linguistically distinct from their neighbors but that are irrelevant as independent political actors. . . . Let me be clear: my assertion is not that the many ethnic groups included in the Atlas (and, often, in the newer measures) are unimportant per se. Rather, my claim is that these groups are unimportant for the explicitly political mechanism that the researchers are trying to test. To capture the contribution that a country's ethnic heterogeneity makes to its policymaking process requires an index of fractionalization that reflects the groups that are actually doing the competing over policy, not the ones that an ethnographer happens to identify as representing distinct cultural units. (Posner 2004, 853)

242. The Possibility Principle is closely related to scope conditions. The welfare state literature only looks at wealthy countries. Much of the literature on well-being focuses on poor or destitute countries (Dasgupta 1990). When Dasgupta examines the relative destitution of countries he excludes the well off ones. He uses the Borda Rule (Dasgupta and Weale 1992) to aggregate a variety of indicators of well-being, such as infant mortality, literacy, per capita income civil and political rights. The Borda Rule ranks countries on each dimension then sums the ranks (see Fine and Fine 1974 for an extensive
and technical discussion of its formal properties). Explain why this concept of human destitution or well-being is sensitive to the exclusion of wealthy countries. For example, the comparison of India and China will significantly be affected by the inclusion or exclusion of nonpoor countries.

Answer:

As Dasgupta and Weale note, the Borda Rule is sensitive to the population chosen because the distance between a pair of countries depends on the population. In particular, if an alternative population puts countries in between a given pair their overall rankings (and hence the comparison between the two) can change. Particularly with regarding to civil and political rights the inclusion of wealthy countries can change the comparison because these are scattered throughout the ranking.

243. Fox and Lawless (2004, 2005) study the phenomenon of “nascent political ambition” in American politics. They need to gather data on those who “might” run for office. Explain how they use the Possibility Principle in designing their large-N survey (see Maetas et al. 2006 for another approach).

Answer:

To form their population for sampling they look at the professions of people who have already run for office or already hold office. In accordance with the Possibility Principle they use those that have run or hold office to determine the population of those who might run for office: “Despite the theoretical importance of studying nascent political ambition, a number of methodological and sample design issues make conducting an empirical investigation quite difficult. The complexity of assembling a national sample of potential candidates, alone, explains why most research on political ambition and candidate emergence focuses on declared candidates and office holders. In an attempt to overcome these difficulties, we developed the “eligibility pool approach,” which we carried out in the Citizen Political Ambition Study. We drew a national sample of 6,800 individuals from the professions and backgrounds that tend to yield the highest proportion of political candidacies: law, business, education, and political/community activism.” (Fox and Lawless 2005, 647)

Note also that the other approach to sampling in this area the “reputational approach” starts by starting with office-holders (i.e., people who have been candidates AND successful) as a place to get names of “prospective” or “visible” candidates (e.g., Stone and Maisel 2003).

244. Many theories of coalition government formation rely on formal models where a key variable is the bargaining power of the potential coalition partners. Ansolabehere et al. (2005) describe a method for calculating “voting weights” based on a calculation of all possible minimal winning coalitions (p. 555; see also Warwick and Druckman 2006). Explain how you could use the Possibility Principle to improve this method of calculating weights? (See Waters 1998 for a philosophical discussion and example-using the analogy with copper as used in the book-of how one would think about the about the minimum winning coalition problem)

Answer:
“All possible minimal winning coalitions” includes many coalitions that are very unlikely to happen (e.g., between extreme left and right parties). It would be better to include only those minimal winning coalitions with any real likelihood of forming. Notice that doing so would require using theory and data about what kinds of coalitions actually do form (i.e., a theory of the positive outcome). For example, Warwick and Druckman mention this common critique: “A common critique leveled against measures such as these [of bargaining power] is that they assume that all coalitions are equally likely to occur.” (Warwick and Druckman 2006, 645).

245. Explain the two variables in Singh and Way’s quantitative analysis of nuclear proliferation (2004) that could be used with the Possibility Principle. What in their discussion and empirical results would support such a view? Note that the IAEA has category of “nuclear latent states”; does this concept invoke the Possibility Principle?

Answer:

They clearly argue that technological/industrial capacity is a necessary condition for nuclear weapons acquisition: “there are numerous examples of states that have had the technical capacity to build nuclear arms for several decades but have never attempted to do so. However, we argue that it is an important starting point because no nation can build nuclear weapons without attaining a minimal economic/technological capacity” (p. 862).

“Every country to acquire nuclear weapons, with the exception of Pakistan, was above the threshold embodied in the [industrial capability] index.” (p. 875 footnote 30)

They also note empirically that the existence of a rivalry is basically a necessary condition: “This is not surprising when one recalls that of the countries to acquire nuclear weapons, all but two (South Africa and France) are coded as participating in an enduring rivalry at the time of acquisition. By alternate but defensible coding rules, one could make the case that both of these exceptions were involved in enduring rivalries.” (p. 875)

Also, the empirical results show that the industrial variable and the enduring rivalry variable have extremely large effects, with rivalry having from a 380% to 740% impact and industrial capacity 560% to 2,340% effect. No other variables are even close to this.

246. In the large-N study of US foreign aid – or foreign aid in general – one must choose a population of cases. This involves deciding whether to use all countries or some subset that “might” receive foreign aid. What might be a criterion to exclude a set of countries that almost never get foreign aid (e.g., Lai 2003; see also Kosack and Tobin 2006 who contrast development aid with foreign direct investment in a world-wide sample). See also Carey (2007) for another approach to choosing cases based on who already receives aid.

The same question can be asked in studies of U.S. arms exports. Examine Blanton (2005). Are developed countries excluded? Should they be included or excluded? What effect might that have on the results? Where can you see the logic of the Possibility Principle in her argument?
Answer:
OECD countries almost never receive foreign aid. More generally, aid is not given to rich countries. One does not give foreign aid to Sweden and Switzerland.

In Blanton all developed countries are excluded. So this means including Saudi Arabia but excluding Israel. Arms transfer certainly do occur to developed countries so it is not completely clear (other than the convention of excluding them in these studies) why they should be excluded. The section “A two-stage model” clearly uses the logic of the Possibility Principle. This also illustrates the close connection of the Possibility Principle to Heckman selection models.

247. In the large-N quantitative conflict literature some researchers prefer to use directed dyads instead of the more common undirected dyad approach. The non-directed dyad only includes one case per dyad-year, e.g., US-USSR. The directed dyad includes two, US-USSR and USSR-US. One advantage of the directed dyad is that it allows you to look at dispute/war initiation. For example, Bennett and Stam (2004) argue that “Given that we want to study initiations, the correct population to employ includes cases where there could have been MID initiation.” (p. 55). One problem they then have is what to do in with directed dyad BA when A has already initiated a dispute against B. Discuss whether it is “possible” then for B to initiate a dispute with A (in the same year) and thus whether one should keep (or not) the BA directed dyad in the population.

Answer:
It turns out that there are about 150 cases where two countries have multiple disputes within the same year (Goertz, Jones and Diehl 2005). Bennett and Stam argue (p. 55–56) that one should exclude the BA directed dyad when a AB dispute exists. However, if there does happen to be a BA MID they do include it in the analysis. Notice that this is the opposite of the way politically relevant dyads work, since the impossible-but-happens cases are excluded.


Answer:
Negotiated settlement is sufficient condition for no genocide. Military victory is necessary condition for genocide.

249. In the geometric representation of the Possibility Principle Mahoney and Goertz used a cut-off of .50. We think that in practice, qualitative researchers will tend to chose much high cutoffs. Discuss whether you think this is true. Can you think of relevant examples? If it is true why do qualitative researchers do this?

250. Some researchers believe that one should look at “revolutionary situations” instead of revolutions:
We can clear a good deal of conceptual ground by means of a simple distinction between revolutionary situations and revolutionary outcomes. Most significant disagreement about the proper definition of revolution falls somewhere along these two dimensions. (Tilly 1978, cited in Walton 1984, 11; see figure in Tilly 1978)

Explain how the idea of a revolutionary situation relates to the Possibility Principle.

Answer:

The quotation from Tilly suggests that one might use the concept of “revolutionary situation” to define cases where the outcome of (successful) revolution is possible. Hence, in a study of revolutionary outcomes, appropriate negative cases might be revolutionary situations that did not culminate in revolutionary outcomes.

See Stinchcombe for a related point: “Rather than explaining the occurrence of revolution, a sociological theory ought to try to explain the occurrence of a ‘revolutionary situation.’ Whether or not a change in the ruling powers of a society takes place by means of violence depends both on the predisposing characteristics of the social structure and on concrete military and political situations at given historical times” (1965, 169).

251. The cube in Figure 2 (Mahoney and Goertz 2004) presents three zones with disconfirming cases. But in qualitative research the zone near the right-back-bottom (1,1,0) corner is often the most important one for falsifying theories.

Why?

Answer:

Necessary condition hypotheses are disconfirmed by cases where the outcome takes place (located in the top half of the cube). Sufficient condition hypotheses are disconfirmed by cases where the outcome does not happen (located in the bottom half of the cube).

In qualitative research, analysts are often quite familiar with many or all positive cases. As a result, they may be less likely to build theories that are obviously disconfirmed by positive cases. By contrast, they may be less knowledgeable about nonpositive cases, such that these cases may be more likely to disconfirm their theories. The right-back-bottom (1,1,0) corner is where all of these disconfirming nonpositive cases will be located.


Answer:

Bennett’s outcome is Soviet/Russian intervention, and he must decide which nonpositive cases are relevant for theory testing. He uses the Rule of Inclusion to inform this choice. Bennett (1999) uses this logic when selecting cases to test a theory of the causes of Soviet and Russian military intervention. He identifies five factors that provide “opportunities” for intervention, such as
the presence of a pro-Soviet/Russian faction or a low level of U.S. threat. He then considers as relevant only those countries or territories where one or more of these factors provided the opportunity for intervention; he does not consider the vast majority of countries or territories, because Soviet or Russian intervention was simply not possible in these cases.

253. Explain how the Possibility Principle is crucial to the large-N study of war. For example, Vasquez in his ISA presidential address: “Of course, many such dyads do not go to war simply because they do not have any conflict with each other or the opportunity to engage in militarized confrontation. Therefore, the analysis will be confined to only dyads that actually engage in the threat or use of force, what has been called involvement in a militarized interstate dispute (MID)” (Vasquez 2004, 2).

Answer:

The militarized dispute dataset (Jones et al. 1996) was originally developed so that scholars would be able to study a set of cases where war was possible. The Singer and Small war data set existed but there was no set of negative cases to compare it against. Hence, the militarized dispute data set helped scholars identify appropriate negative cases.

254. Read Barbara Harff’s (2003) article on genocide. Explain how she (informally) uses the Possibility Principle to select her cases. Does she primarily use the Rule of Inclusion or the Rule of Exclusion to eliminate irrelevant cases? See also later work using the same basic ideas, e.g.:

We follow Harff (2003) and Colaresi & Carey (2008) in restricting the sample to country-years experiencing political upheaval since almost all mass killings occur during or immediately after episodes of political upheaval. Like Colaresi & Carey (2008), we only include country-years with ongoing state failure, defined as revolutionary wars, ethnic civil wars, adverse regime transitions, or genocides or politicides, as coded in the Political Instability Task Force (PITF) data (Marshall, Gurr & Harff, 2013). To ensure that our estimation results are not simply an artifact of this decision, we also examine the full sample including all country-years during the sample period for those countries for which data exist (1955–2010). Results remain similar. ((Hong and Kim 2019, 535)

Answer:

Harff uses the rule of exclusion to remove cases where genocide is impossible. In particular, she excludes all cases that lack a causal factor (political upheaval) that is almost always necessary for genocide.

Harff (2003) notes that almost all genocides (i.e., 36 out of 37) occur during or immediately after political upheavals. Accordingly, she excludes cases like France and Canada that lack political upheaval when testing her theory of genocide. These politically stable cases have such a low probability of experiencing genocide that their inclusion would distort inferences about other cases where the outcome of interest is possible. For example, one may have a theory of genocide that highlights ethnic divisions as a key independent variable. Under the Rule of Inclusion, contemporary Canada could therefore
be considered a relevant case. However, under Harff’s (2003) exclusion criterion, Canada is irrelevant because its value on the political upheaval variable eliminates it from the analysis.

255. The quantitative conflict literature uses “politically relevant dyads” to apply the Possibility Principle to select cases. Mahoney and Goertz suggest that the total population is the union of negative cases (selected with Possibility Principle) and the positive cases. However, the conflict literature uses a different procedure to determine the total population of cases. What is that procedure? How does that influence what the researcher does when the “impossible happens?”

Answer:
The quantitative literature on politically relevant dyads adopts the following procedure: (1) use the Possibility Principle to distinguish relevant and irrelevant cases, (2) treat the relevant cases as the whole population of interest, and (3) select the positive cases as a subset of this population. As a result of this procedure, a mistake in step one can cause positive cases to be coded as irrelevant, such that the impossible happens.

256. Does the state failure (http://www.cidcm.umd.edu/inscr/stfail/) project use the Possibility Principle to select cases? If so, how? If not, should it use that principle?

Answer:
The project does not exclude cases where state failure is impossible, and thus the project does not employ the Possibility Principle. The project might reach different results if the Possibility Principle were applied. For example, the application of this principle might lead to the exclusion of many or most of the advanced industrial democracies. In turn, with the new population of relevant cases, different results could emerge.

257. What would be the set of appropriate negative cases for the following Boolean theory?

\[ Y = A + B + C + D \]  

(2)

Does this contradict the Possibility Principle? Why does this example not fit the general kind of problem that typifies case selection in small-N studies? Hint: Which combinations of A, B, C, and D predict that Y does not occur?

(Thanks to Charles Ragin for this exercise.)

Answer:
This is a question dealing with the logic of the Possibility Principle. The problem here is that if we followed the Possibility Principle for selecting negative cases, we would not have any negative cases at all since the Possibility Principle excludes all cases where A–D are absent (i.e., not-A AND not-B AND not-C AND not-D, this is a hint in the question). But this is exactly the set of cases we need to test the Boolean theory \( Y = A + B + C + D \). This situation arises because A–D are individually sufficient for Y. However, models like \( Y = A + B + C + D \) are extremely rare in the qualitative-comparative literature.
258. Skocpol considers the Cuban Revolution (1959) to be irrelevant for directly testing her theory. Do you think she excludes this case using the Possibility Principle or scope conditions? Why?

Answer:

Skocpol must be using scope conditions – not the Possibility Principle – to exclude the Cuban Revolution, given that the Possibility Principle excludes only cases where the outcome of interest is absent.

259. Explain how the cube (Mahoney and Goertz 2004) provides a visual summary of disconfirming observations for necessary and/or sufficient condition hypothesis.

Answer:

With a necessary condition hypothesis, theory disconfirmation occurs when the necessary condition is absent, but the outcome of interest is present. Thus, disconfirming cases are located in the top portion of the cube where only one necessary condition is present (i.e., the (1,0,1) and (0,1,1) corners).

With a sufficient condition hypothesis, theory disconfirmation occurs when the sufficient condition is present, but the outcome of interest is absent. Thus, disconfirming cases are located in the bottom portion of the cube where the two jointly sufficient causes are present (i.e., the (1,1,0) corner).

260. Seawright (2002; see Clarke 2002 and Braumoeller and Goertz 2002 for responses) argues that one should use “all cases” for testing hypotheses concerning necessary or sufficient conditions. Explain how the Possibility Principle and the Raven paradox enter into this debate.

Answer:

Much of the debate centers around the issue of which nonpositive cases are relevant for theory testing. Seawright suggests that all cases in an “appropriately defined universe” are relevant to testing a proposition about causal sufficiency, even negative cases that lack the hypothesized sufficient cause. He shows that the inclusion of all cases can substantially enhance statistical significance by increasing the number of confirming observations. By contrast, Clarke argues that including all cases will lead one to confirm a proposition through irrelevant observations, in much the same way that “most ravens are white” might be confirmed by observing yellow pencils and blue books. Braumoeller and Goertz’s argument likewise suggests that, when testing a hypothesis about a sufficient cause, cases that lack both the cause and the outcome are irrelevant, since the hypothesis does not imply anything about the number or proportion of these cases that should be present.

261. Find the parameter that can change dramatically depending on whether one applies the Possibility Principle or not in King and Zeng’s (King and Zeng 2001a) discussion of statistical estimation in rare events data. For a similar problem see also King and Zeng (2001b).

Answer:
“Prior correction requires knowledge of the fraction of ones in the population, \( \tau \). Fortunately, \( \tau \) is straightforward to determine in international conflict data since the number of conflicts is the subject of the study and the denominator, the population of countries or dyads, is easy to count even if not entirely in the analysis.” (King and Zeng 2001a, 144).

In King and Zeng (2001b) one should note similarly that relative risk involves calculations using baseline probabilities (\( Pr(Y = 0) \) or \( Pr(Y = 1) \)) so that \( Pr(Y = 1) \) will be much lower with all dyad design.

262. The welfare state literature typically excludes less-developed countries. Is this because of scope conditions or the Possibility Principle? If it is the Possibility Principle, is it the Rule of Inclusion or Rule of Exclusion?

Answer:

To avoid these problems, many analysts of welfare states include only OECD countries (see Pierson 2000; Amenta 2003 for recent reviews). Typically, they justify the exclusion of poorer countries through the use of scope conditions. However, they exclude only negative cases, and so they are really employing the Possibility Principle, not scope conditions. In particular, they use the Rule of Exclusion to eliminate countries that possess a condition sufficient for the absence of welfare state development—namely, poverty. Indeed, the finding that economic wealth is related to welfare state development among all countries but not among rich countries is what we would expect if all cases are homogeneous (i.e., if scope conditions do not apply).

263. Fuzzy-set analysis is often problematic when there are many observations near the origin (i.e., observations whose values on all variables approach zero). When observations are near the origin it is extremely easy for them to pass tests of sufficiency since \( X \) is likely to be less than \( Y \). How does the Possibility Principle aid fuzzy-set methodology by dealing with these cases?

Answer:

This geometric interpretation of the Possibility Principle offers a solution to an important problem that arises in fuzzy-set analysis. The problem involves what to do with cases that are near the origin (i.e., cases near the (0,0,0) corner). As Ragin notes (2000, 250–51), when testing whether variables are causally sufficient for an outcome, observations with a zero for all the independent variables will always satisfy causal sufficiency and thus artificially inflate the number of cases where the theory works (this dilemma is the Raven Paradox). See also the exchange between Osherson and Smith (1981) and Zadeh (1982). The former is a classic in the cognitive psychology literature on categorization and Zadeh is the inventor of fuzzy-set logic.

264. Do Foran (1997) or Goodwin (2001) apply the Possibility Principle in their studies of (social) revolution?

Answer:

Foran’s (1997) Boolean analysis of social revolution selects as negative cases only state-periods that have a positive value on at least one of his five major
independent variables. Goodwin (2001) likewise selects as negative cases only state-periods where at least one key independent variable is present.

**Multimethod research**

265. This case illustrates the importance of the Avoid Overdetermination Rule. An alternative to the importance of nuclear weapons is clearly conventional weapons superiority. In many circumstances it is likely the case that nuclear weapons are accompanied by conventional weapons superiority. How many, if any, of the cases satisfy the Avoid Overdetermination Rule? Discuss what this means in the statistical analyses if essentially the nuclear weapons variable is almost colinear with the conventional weapons variable.

Conclusion. The chapter critically assessed ten apparently successful cases of nuclear coercion. These cases should have provided the clearest evidence in favor of the view that nuclear blackmail works. However, none of them unequivocally supports the nuclear coercionist school. In each case, there is some doubt – often considerable doubt – about whether nuclear weapons provided states with coercive leverage. Indeed, several of these crises provide strong support for nuclear skepticism theory, highlighting the challenges that countries face when attempting to employ nuclear signals for coercive purposes. It is worth noting that the cases that provide the strongest support for the nuclear coercionist view – including the Cuban missile crisis – happened in the early days of the Cold War. The world has changed dramatically since that time. It is hard to argue that cases from the 1950s and 1960s tell us much about the role of nuclear weapons in world politics today. There is scant evidence that nuclear blackmail has worked since the collapse of the Soviet Union. The nuclear crises of the last quarter-century illustrate the coercive limits, rather than the virtues, of nuclear weapons.

266. An issue in multimethod research is the location of the case study vis-à-vis the independent variable in the statistical analysis. Lee (2018) chooses a case study which scores very low on her independent variable of rivalry (at best a low level rivalry). “I complement the quantitative analysis with a qualitative examination of Malaysia’s role in undermining domestic sovereignty in the 1970s in Mindanao, a region of the southern Philippines.” Given that this is a low level rivalry how might she defend her case selection?

Answer:

She could use an extrapolation justification: if it works for weak rivalries it should work for more serious rivalries. See figure 8.2 in Goertz 2017.

267. Grewal and Kureshi (2019) is an example of multimethod research. The main hypothesis involves an interaction term. Discuss the selection of the two case studies in light of the issues posed by an interaction term in the theory-statistical analysis.

268. Law and norms are often considered constraints. What would be good case studies to study this constraint mechanism?
The threat that legalistic trade dispute settlement poses to the discretion of political leaders is threefold. First, it may constrain their ability to manage the unforeseen costs of adjustment, making it more costly to provide relief or protection to specific groups injured by trade liberalization. Second, it may limit their general policy autonomy across a range of domestic regulations, which it judges against treaty commitments to eliminate nontariff barriers to trade. A third and final consideration is that the delegation of authority to third parties may constrain their ability to pursue trade policy bilaterally, a strategy with distinct political advantages. On all three counts, political leaders in trading states are risk-averse regarding the impact of dispute settlement on policy discretion. Other things being equal, they do not want to cede veto power over domestic policies to appointed trade law experts or judges, because the political price of doing so may be high. (McCall Smith 2000, 143)

269. Gerring and Cojocarū describe Mansfield and Synder’s work as follows. Using the discussion of large-N qualitative testing in *Multimethod research, causal mechanisms, and case studies: an integrated approach* explain what is wrong with their characterization.

In a setting where the relationship between X and Y is well established—perhaps as a result of cross-case analysis (the researcher’s or someone else’s)—the pathway case is usually focused specifically on causal mechanisms (M). An example is provided by Mansfield and Snyder’s (2005) research on regime transitions and war. The authors find a strong relationship between democratization and bellicose behavior in their large-n cross-national analysis. To ascertain whether their hypothesized causal mechanisms are actually at work in generating this relationship, they look closely at countries where the posited covariational pattern between X and Y clearly holds, that is, where democratization is followed by war. (Gerring and Cojocaru 2016, 406)

Answer:

Their core hypothesis is an interaction between democratization and weak institutions, not democratization alone. They do not find a strong relationship between democratization alone and militarized conflict.

270. Should one assume an interaction term or just a straight additive model, or a Boolean model for the data in the table below? Table 4 is a stripped-down version of their table 2 (p. 88) which gives the two core independent variables: (1) federal versus centralized state and (2) multiparty system versus two-party system. I use their primary dependent variable “macrocorporatism” (the other one they discuss is “sector coordination”). There is no statistical analysis in this article; the empirical analysis is carried out via four case studies, one from each cell in table 4.

Answer:

As in almost all two-variable models – i.e., $X_1 \text{ AND } X_2 \rightarrow Y$ – the $(1,1,1)$ cell is clear as is the $(0,0,0)$ cell. The critical feature that will influence the decision between a QCA interpretation against an additive linear one is what happens with the cases where one independent variable is present and the other is absent, i.e., $(1,0)$ or $(0,1)$.  

100
Table 3: Macrororporatism

<table>
<thead>
<tr>
<th>Number of Parties</th>
<th>Level of State Centralization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Centralized</td>
</tr>
<tr>
<td>Multiparty</td>
<td>Denmark macrocorporatism=0.72</td>
</tr>
<tr>
<td>Two party</td>
<td>Britain macrocorporatism=-0.14</td>
</tr>
</tbody>
</table>

Source: Martin and Swank 2011, table 2.

In set theoretic models the value of the (1,0) and (0,1) cells would be close to that of (0,0) because of the necessary condition relationship. The additive model suggests that when one independent variable is present and the other absent we should see intermediate results, somewhere between the extreme values on the dependent variable. Hence, a signal that the additive relationship is at work is when the dependent variable is significantly greater than the (0,0) but yet significantly less than the (1,1) dependent variable value. This is in fact what we find in table 4. The two off-diagonal cells have values in between those in the on-diagonal cells.

It is pretty clear that a set theoretic interpretation of these data would not be appropriate because of the intermediate values of the dependent variable when only one independent variable is present. However, it is very difficult to make any strong conclusions regarding whether there should be an interaction term.

271. In a very influential article Alesina and Spolaore (1997. On the size and number of nations. Quarterly Journal of Economics) give a formal model. Discuss it in terms of game theory multimethod research. For example, how plausible are their assumptions? What is the causal mechanism and how would you explore that in case studies? You can explore this in more detail in their book on the topic.

Answer:

For example, they do not discuss war which is historically probably the most important determinant of the size of nations.

272. Copeland (2015) has at the core of his book an interaction; positive levels of economic interdependence (often trade) interacted with expectations about future economic conditions. Discuss the situation where one of the interaction terms is basically a subset of the other. For example, what case studies we be appropriate? Is the superset variable (positive economic independence) more like a scope condition?

To determine whether the liberal prediction or realist prediction will prevail, we must introduce an additional causal variable—namely, a state’s expectations of the future trade and investment environment. When a
dependent state has positive expectations about this future environment, it is more likely to see all the benefits of continuing the current peace and all the opportunity costs of turning to war. Economic interdependence would then be a force for peace. Yet if a dependent state has negative expectations about the future economic environment—seeing itself being cut off from access to foreign trade and investment, or believing that other states will soon cut it off—then the realist logic will kick in. Such a state will tend to believe that without access to the vital raw materials, investments, and export markets needed for its economic health, its economy will start to fall relative to other less vulnerable actors. If this economic decline is anticipated to be severe, the leaders of the dependent state will begin to view war as the rational lesser of two evils—that is, as better than allowing their state to fall to a point where rising states can attack it later or coerce it into submission. (Copeland 2015, 2; basic hypothesis of the book)

273. Seawright (2016) argues for “extreme on X” case selection in multimethod research. Extreme on X could mean very high or very low or distant from the mean. Contrast that with “good” causal mechanism case selection as argued in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, i.e., some extreme cases are good cases and others are not good cases.

274. Bush (2011) is one of the relatively rare statistical multimethod articles, which chooses a case study based on the statistical results. Her case study is Afghanistan:

> A case study can check the results’ robustness through process tracing . . . . Since the preliminary large-N analysis generated satisfactory results for my theory, I chose Afghanistan—a country that was well-predicted by the statistical model (footnote: Lieberman 2005, 442–43. The average residuals for Afghanistan from Model 4 in Table 2 were –0.03) (Bush 2011, 126)

Discuss the pro’s and con’s of this as a good on-line case study – in the qualitative logic sense as well as in the statistical multimethod sense.

275. Collier and Sambanis (2005) conducted a large multimethod project on civil war and case studies. Explore their rationale for selecting cases to analyze since they cannot look at all countries for all time periods.

276. When doing statistical multimethod work scholars often choose “theoretical on-line” cases, i.e., cases where the theory is working (aka (1,1) cases, assuming a positive relationship). Construct an example where the theoretical on-line case is one which is not on the regression line, and which has a very large positive residual vis-à-vis the regression line.

Answer:

The ideal theoretical on-line case would lie in the top right hand corner (i.e., $X = 1$ and $Y = 1$, assuming variables [0,1]). Draw a regression line with a modest slope (e.g., .25). Cases in the top right corner will have large positive residuals.
Theories and concepts, e.g., two-level theories

277. Two level or SUIN models raise questions about what to include in the statistical analysis. An example on the independent variable side is Uzonyi’s concept of a “big political opportunity” which takes three forms that are expressed in his three hypotheses. Should he include each of the three forms separately in a regression analysis or just the higher variable formed by the logical OR?

There are generally three types of domestic unrest that may lead to the onset of genocide or politicide. The first, which is captured by those scholars focusing on big opportunities for government mass murder, are militarized threats to a country’s political leadership. . . . Coups provide a similar logic. . . . The third type of threat a government faces is non-militarized unrest. Non-militarized domestic unrest, such as protests, riots and strikes, can spark political violence if the regime perceives that such unrest will grow into greater instability. (Uzonyi 2016, 318–19)

278. Ohlson (2008) “Understanding causes of war and peace” has implicit two-level theory. One might think that the basic level variables are the same, his “Triple-R” triangle for both causes of war and causes of conflict resolution, but that the secondary-level variables are different for each dependent variable. Discuss.

279. Horowitz (2010) chapter 2 presents a two-level theory. For the two core, basic level, independent variables of organizational capital and financial intensity. Draw a figure with ANDs and ORs (or with *’s and +’s for the linear algebra version) that represent the logical structure of his theory.

280. Discuss the Rodrik (2007) figure below. Note that everything is conceptualized as a constraint (see chapters 2 and 4 of Multimethod research, causal mechanisms, and case studies: an integrated approach). Which case studies would be useful? Which ones does he do? Discuss the nature of the arrows – e.g., are they causal – and the two-level structure of the causal mechanism. Could the arrows be casually going up instead of down?

Moving down the branches of the decision tree is tantamount to discarding candidates for the most binding constraint on growth. The overarching lesson from our theoretical analysis is that it is this constraint, once identified, that deserves the most attention from policymakers. (Rodrik 2007, 66)
CHAPTER TWO

Moving down the branches of the decision tree is tantamount to discarding candidates for the most binding constraint on growth. The overarching lesson from our theoretical analysis is that it is this constraint, once identified, that deserves the most attention from policymakers.

COUNTRY EXPERIENCES: IDENTIFYING THE CONSTRAINTS

We now have a framework to think of growth diagnostics. In this section we apply our approach to three countries with three very different growth experiences: Brazil, El Salvador, and the Dominican Republic. The first two countries have had lackluster growth in spite of quite impressive reforms. The last had a sustained period of very rapid growth triggered by rather modest reforms, but more recently has stumbled into a financial crisis from which it has had difficulty extricating itself fully.

Both Brazil and El Salvador made major efforts at dealing with their perceived problems during the 1990s. Brazil returned to democracy in the 1980s, started opening up its economy in the early 1990s, stopped...
Chauvinist elites use symbolic appeals to group myths, tapping into and promoting fear and mass hostility, to mobilize their groups for conflict. A predation-driven security dilemma arises, in which the growing extremism of the leadership on at least one side results in radicalization of the leadership on the other. These propositions together identify a causal mechanism. For the model to be correct, there must be evidence that in each case of ethnic war, the necessary preconditions not only were present but were causally important in enabling chauvinist mobilization and in causing a predation-driven security dilemma. Additionally, the logic requires that an explicitly genocidal policy must be based on explicitly eliminationist myths. (Kaufman 2006, 58)

One INUS model would be: $S_1*S_2*S_3*(S_4+S_5+S_6)$.

“Peace” is overdetermined in the sense that if any of the necessary conditions, i.e., $S_1$–$S_3$ is absent then there is no conflict. Conflict is also absent if $(S_4+S_5+S_6)$ is too small.

283. Compare Geddes’s (1993) figure of Skocpol’s theory to that in chapter 10 of *Social science concepts*. Give two major differences between the two.

Answer:

1. military threat, and substitutability of other causes of state breakdown and (2) problems with interaction term and her interest in “the” effect of military threat.

284. One important question is the extent to which two-level or fuzzy set models/methods differ from statistical ones. Examine Geddes’s (1993) figures 3.5 and 3.6. She finds no linear relationship between labor repression and economic growth. Draw a new horizontal and vertical through these figures which gives a $2 \times 2$ table that nicely supports the hypothesis that labor repression is a necessary condition for high economic growth.

Answer:

One can easily draw horizontal and vertical lines through her scatter plot to show that her data support the necessary condition hypothesis. Just draw a horizontal line at around 4% GDP growth and the vertical line at 1.5 labor repression.


Answer:

Basic level concepts are (1) international capacities, (2) hostility, and (3) local capacity. For secondary-level variables see page 783, in the section “The Explanatory Variables.”

“All strategies should address the local sources of hostility, the local capacities for change, and the (net) specific degree of international commitment available
to assist change. One can conceive of these as the three dimensions of a triangle whose area is the ‘political space,’ or effective capacity, for building peace. This suggests that the dimensions substitute for one another, that is, more of one substitutes for less of another; less extreme hostilities substitute for weak local capacity or minor international commitment.” (Doyle and Sambanis 2000, 781; see also the statistical model of table 4, p. 791)

286. Elman (2005) in his article on typological methods discusses the method of “indexing.” Using his discussion of Walt’s theory of alliance, explain how indexing can be the creation of a higher level variable/concept. How would the indexing he does fit into the varieties of two-level theories?

Answer:

“Threat” is composed of several secondary-level variables such as capability and proximity. By adding up the scores he is clearly using a family resemblance strategy.

287. Combine Walter (1997) and Hartzell and Hoddie (2003) to form a two-level theory of civil war conflict resolution. As basic-level variables use third-party guarantees and power-sharing. See Walter for an argument about the structural relationship between the two. (Note: you might assume the scope conditions that the countries have already started negotiations.)

Answer:

Walter (p. 351) clearly argues that there is substitutability between the two basic-level variables. Walter argues a necessary and sufficient condition view of credible commitments (though she is not completely explicit on this point). Hartzell and Hoddie clearly argue for a family resemblance perspective on power-sharing (p. 321).

288. What are the three “basic level” hypotheses in Skocpol’s States and social revolutions and which one of these is Geddes (2003) basically (though obviously not exactly) testing in her discussion of Skocpol?

Answer:

The three basic-level hypotheses are: (1) peasant revolt is necessary for social revolution; (2) state breakdown is necessary for social revolution; and (3) the combination of peasant revolt and state breakdown is sufficient for social revolution.

Geddes explores whether “external threat” is correlated with the occurrence of social revolution. She operationalizes external threat in terms of invasion and/or loss of territory. She also briefly explores the hypothesis that external threat is necessary for social revolution. Hence, Geddes examines whether one of Skocpol’s secondary-level variables for “state breakdown” is causally related to social revolution. Note that Skocpol herself hypothesizes that external threat affects social revolution by working through the basic-level variable of state breakdown.

Answer:

Sekhon reassesses Geddes’s data from Latin America concerning the relationship between defeat in war and revolution. Contrary to Geddes, he concludes that these data do not rule out the possibility that the two factors are (probabilistically) related to one another. Hence, he concludes that Skocpol’s argument may be correct, if stated probabilistically.

By contrast, chapter 10 evaluates Skocpol’s theory under the assumption that defeat in war is a secondary-level variable that affects social revolution only by working through the basic-level variable of state breakdown. Hence, from this perspective, Sekhon’s reanalysis is not directly testing the theory developed by Skocpol.

290. Burawoy’s (1989) critique of Skocpol shows how important, and how difficult, it is to see the complete causal structure of Skocpol’s argument. His table 1 (p. 768) clearly divides variables into two levels and he correctly identifies peasant revolt and state crisis as the basic-level variables. However, he does not interpret the relationship between the secondary level and the basic level correctly. What is the nature of his error?

The task now is to show that both international pressure and an ‘organized and independent dominant class with leverage in the state’ were necessary ingredients for political crisis . . . So far so good, but note immediately that the contrasting cases [Germany and Japan] do not demonstrate ‘international pressure’ as necessary for the development of a revolutionary political crisis. In the next chapter Skocpol examines the necessary conditions for the second component of revolution: peasant revolt . . . She now has to demonstrate that both political crisis and peasant autonomy were necessary for peasant revolt. (Burawoy 1989, 766)

Answer:

Burawoy assumes that the secondary-level variables are necessary for the basic-level variables. However, the indicator/data-level variables are better characterized as individually sufficient for the basic-level variable.

291. Most and Starr introduced the influential notion of foreign policy substitutability (Most and Starr 1984; see also Most and Starr 1989). They are also well-known for the idea that opportunity and willingness are individually necessary and jointly sufficient for foreign policy action (Starr 1996). Put opportunity and willingness together with foreign policy substitutability to get the structure that Ostrom uses in her model of common pool resource institutions.

Answer:

If one puts opportunity and willingness at the basic level and foreign policy substitutability at the secondary level, one arrives at a two-level model. Cioffi-Revilla and Starr (2003) formally present this kind of model in a completely probabilistic fashion.

292. Interpret the following theories in two-level terms. Some may not be complete two-level models, in that the secondary-level may not always be clearly present for all basic-level factors.


Answer:

The answers below are suggestive hints and avenues to pursue and do not constitute a complete answer to the question.

(1) “The requisites of consensual incomes policy that have been discussed here, socialist participation in govern and centralized trade union structure, coincide only in the Nordic countries, the Low Countries, and Austria…. Without stable social democratic governance, consensual incomes policy is unlikely to be attempted, and without centralized trade unions, it is unlikely to be a viable policy in the medium term.” (Marks 1986, 269)

(2) “A general theoretical argument focuses on four conditions. We argue that the operating system only responds to normative changes when response is “necessary” (stemming from incompatibility, ineffectiveness, or insufficiency to give the norm effect, and when the change is roughly coterminal with
a dramatic change in the political environment (that is, “political shock”). We also argue, however, that opposition from leading states and domestic political factors might serve to block or limit such operating system change.” (Diehl et al. 2003, 71–72)

(3) “Institutional transfer involves three necessary conditions: first, state elites refer explicitly to a model prominent in another place; second, they try to identify the foreign model’s legal framework and the actors that help it function; finally, these elites build a replica of all or part of the model.” (Jacoby 2001, 2)

“For institutional transfer to be effective, Jacoby says, two conditions have to be met: First, institutional transfer has to be supported by active segments of civil society, and second, flexibility in adapting institutions to foreign settings is essential. . . . Similarly, the politics of imitation faces higher obstacles when exact transfer is intended. A functional-equivalency approach, with its built-in elasticity, is preferable.” (Welsh 2002, 211; review of Jacoby 2001)

(4) “As the patrimonialist fate of Hungary and Poland illustrate, however, the mere presence of participatory local government is in itself not enough to ensure the triumph of bureaucratic constitutionalism. It is only the combination of participatory local government with a strong center equipped with independent capacities of rule that, the British case implies, can assure such an outcome.” (Ertman 1997, 324 (last page); note that the other key variable is timing of geopolitical competition, either pre or post-1450).

(5) Kingdon’s model has three “streams” each of which is necessary and which together are jointly sufficient. His discussion is very clear about the different substitutable ways these streams can be filled.

(6) “Democracy is a form of government of a state. Thus, no modern polity can become democratically consolidated unless it is first a state. . . . If a functioning state exists, five other interconnected and mutually reinforcing conditions must also exist or be crafted for a democracy to be consolidated. First, the conditions must exist for the development of a free and lively civil society. Second, there must be a relatively autonomous and valued political society. Third, there must be a rule of law to ensure legal guarantees for citizens’ freedoms and independent associational life. Fourth, there must be a state bureaucracy that is usable by the new democratic government. Fifth, there must be an institutionalized economic society” (Linz and Stepan 1996, 7).

“Having analyzed the necessary conditions for a consolidated democracy and then spelled out the key differences among the four ideal-typical non-democratic regimes, it should be clear that the characteristics of the previous nondemocratic regime have profound implications for the transition paths available . . . Within the logic of our ideal types, it is conceivable that a particular authoritarian regime in its late stages might have a robust civil society, a legal culture supportive of constitutionalism and rule of law, a usable state bureaucracy that operates within professional norms, and a reasonably well-institutionalized economic society. For such a polity, the first and only necessary item on the initial democratization agenda would relate to political society—that is, the creation of the autonomy, authority, power, and legitimacy
of democratic institutions. We argue in chapter 6 that Spain, in the early 1970s, approximated this position” (Linz and Stepan 1996, 55).

(7) Wickham-Crowley (1992) has a two-level model with the following structure:

I. Peasant support
   A. Agrarian structure
   B. Agrarian disruption
   C. Rebellious cultures
   D. Peasant linkage (guerilla leaders have peasant ties, background)
II. Guerilla/Government military strength
   A. Financing
   B. Internal solidarity
   C. Support from outside (e.g., Cuba, US)
III. Regime weakness (see figure 8.1)
   A. Weak cohesion of upper classes
   B. Weak cohesion of middle classes
   C. US occupation.
IV. Existence of guerilla groups (this is added in Boolean analysis in chapter 12)

(8) Burgerman proposes “an interacting set of necessary conditions” for the success of human rights groups in forging peace agreements: the existence of relevant international norms and transnational activism; the existence of elites in the target state that have control over the armed forces and that have concern about their (country’s) international reputation; and the existence of organized domestic groups linking up with transnational activists. Added to these propositions is the sensible if unremarkable condition that “if a major power maintains overriding security or economic interests in the target state, it can inhibit the enforcement of human rights principles and agreements.”

(9) “Clearly, the necessary but not sufficient conditions for major war emerge only in the rare instances when power parity is accompanied by a challenger overtaking a dominant nation. The odds of a war in this very reduced subset are 50 percent. No other theoretical statement has, to our knowledge, reduced the number of cases to such a small set, and no other is so parsimonious in its explanatory requirements” (Organski and Kugler 1989, 179, see their table same page).

“More importantly, the power transition suggests that, during are rare periods when a challenger overtakes the dominant nation, war will be waged only if the potential challenger is dissatisfied” (Organski and Kugler 1989, 51).

In particular, it is clear that there are various ways a country can be dissatisfied.

(10) “We contend that developmental states will only emerge when political leaders confront extraordinarily constrained political environments. Specifically, we argue that political elites will only build such institutional arrangements when simultaneously staring down the barrels of three different guns: (1) the credible threat that any deterioration in the living standards of popular sectors could trigger unmanageable mass unrest; (2) the heightened need for
foreign exchange and war materiel induced by national insecurity; and (3) the hard budget constraints imposed by a scarcity of easy revenue sources. We call this interactive condition ‘systemic vulnerability’” (Doner et al. 2005).

(11) Necessary conditions are very explicit. Two-level nature of theories comes out clearly in discussion and data in Appendix A. Substitutability is clear, particularly in footnote 2, page 372.

293. Reinterpret the following QCA analyses in terms of two-level theories:


Answers:

(1) In the Boolean equation SUCCESS = A*i + A*S*R one can see “i” and “S*R” being easily interpretable in terms of costs of compliance. Hence there is substitute between them. A is a necessary condition already.

(2) For example, “National health insurance emerged and/or endured in three basic scenarios. First, given a receptive societal culture, a relevant leftist party, and a corporatist pattern of interest group activity, NHI existed regardless of the nature of the major political institutions. Second, unitary states with receptive cultures were capable of producing NHI in countries with weaker executives and in countries without a relevant labor party (irrespective of the nature of the other two independent variables). Third, Anglo-American societies with pluralist interest group environments adopted NHI with a dominant executive (regardless of the federal-unitary distinction or the relevance of labor parties).” (Blake and Adolino 2001, 696)

(3) Snow and Cress (2000) in their analysis of the success of homeless social movements find the same sort of pattern:

Six SMOs [social movement organizations] obtained positions on boards and task forces that addressed the homeless issue [dependent variable]. Two pathways led to this outcome. Organizational viability, diagnostic frames, and prognostic frames were necessary conditions for obtaining representation. These conditions were sufficient in combination with either disruptive tactics, where allies were present, or nondisruptive tactics, in the context of responsive city bureaucracies. (p. 1082)

Here too one sees the substitutability of power in the analysis. One needs either “allies” who are influential in the community or a friendly city government to begin with. You do not use disruptive tactics with friends, where you do if the city government is unfriendly.
294. Braumoeller (2003) has a large table giving many theories that use ANDs and ORs. Examine this list for other potential two-level theory candidates.

295. The Nussbaum example illustrates that one can think of many theories of (social) welfare and justice in the same structural terms as this book applies to concepts:

Most theories of justice can also be usefully analyzed in terms of the information used in two different-though interrelated-parts of the exercise, viz. (1) the selection of relevant personal features, and (2) the choice of combining characteristics. To illustrate, for the standard utilitarian theory, the only intrinsically important “relevant personal features” are individual utilities, and the only usable “combining characteristic” is summation, yielding the total of those utilities. (Sen 1992, 73)

Choose some well-known theories of justice or social welfare and describe the dimensions and the structuring principle.

Answer: Sen gives the following examples:

Examples of selection of “relevant personal features” other than utilities include liberties and primary goods (Rawls 1971), rights (Nozick 1974), resources (Dworkin 1981), commodity bundles (Foley 1967; Pazner and Schmidtler 1974; Varian 1974, 1975; Baumol 1986), and various mixed spaces (Suzumura 1983; Wriglesworth 1985; Riley 1987). . . . The selection of personal features has to be supplemented by the choice of a combining formula, e.g., sum maximization (Harsanyi 1955), lexicographic priorities and maximin (Rawls 1971; Sen 1977 “On weights and measures”), equality (Foley 1967; Nozick 1974; Dworkin 1981), or one of various other combining rules (Varian 1975; Suzumura 1983; Wriglesworth 1985; Baumol 1986; Riley 1987). (Sen 1992, 74)

296. Describe the three-level structure of Dasgupta’s concept of human well-being or destitution (Dasgupta 1990; Dasgupta and Weale 1992). What is its structure? What are the relative weights attached at the indicator and secondary levels?

Answer: The structure is very similar to the one used by Schmitter for corporatism, since it uses the sum of the ranks. The secondary-level level variables such as “health,” “education,” “wealth” will not be equally weighted in general unless they have the same number of indicators. More generally, is it reasonable to rank literacy equally with per capita income?

297. Chapter 2 stresses that an important part of concept-building is considering the negative pole of the concept. Nussbaum and Sen focus on that concept of human needs. What is the opposite pole of a human need?

Answer: Classically one contrasted the concept of needs with that of luxury. For example, “[W]hy is it that all societies make some category distinction between
need/necessity and luxury? ... It is not paradoxical to assert that the distinction constitutes a unity, just as it is not mysterious to talk of two gloves, a right-hand and a left-hand, constituting a pair. ... The relation between necessity and luxury is negative or oppositional so that ... luxury specifies the necessity by indicating what it is not” (Berry 1994, 231–32).

298. When deciding who gets a particular good the weakest link principle, chapter 5, can be used with the need-luxury scale to make a decision. Braybrook (1987) calls this the Principle of Precedence. Explain the moral philosophy involved.

Answer:

The Principle of Precedence means that when Alan needs something that Brenda wants but does not need, then meeting Alan’s need is prima facie morally preferable to satisfying Brenda’s desire. This suggests that needs have precedence over luxuries. The weakest link principle then suggests that for any good you look at each person’s need-luxury score on the good and give it to the person with the minimum (i.e., most needy).

Causal issues, causal hypotheses, causal asymmetry, etc.

299. One problematic concept situation is when the conceptualization and coding of the dependent variable involves important causal claims. Discuss this for Chenowith and Stepan’s influential work:

Success and failure are also complex outcomes, about which much has been written. For our study, to be considered a “success” a campaign had to meet two conditions: the full achievement of its stated goals (regime change, anti-occupation, or secession) within a year of the peak of activities and a discernible effect on the outcome, such that the outcome was a direct result of the campaign’s activities. The second qualification is important because in some cases the desired outcome occurred mainly because of other conditions. The Greek resistance against the Nazi occupation, for example, is not coded as a full success even though the Nazis ultimately withdrew from Greece. Although effective in many respects, the Greek resistance alone cannot be credited with the ultimate outcome of the end of Nazi influence over Greece since the Nazi withdrawal was the result of the Allied victory rather than solely Greek resistance. (Chenowith and Stepan 2011, 10)

300. Often with complex concepts or interaction terms it is not clear what the comparison group might be that one needs for causal analysis. The generic “compared to what?” question. For example, “H1: In states that recently experienced a revolution, personalist dictatorship is associated with greater militarized dispute initiation” (Colgan and Weeks 2015, 170). What is the comparison group since the hypothesis involves the interaction between (1) revolution and (2) personalist dictatorship?
301. “Hypothesis 3: Coups are more likely in states involved in civil conflicts” (Piplani and Talmadge 2016, 1375). Conceptual tautology occurs here because some coups are also civil wars. Discuss how you might fix this.

302. Causation is frequently built into concepts. Discuss the example of “environmentally displaced persons,” which is probably built into international legal conventions.

303. Is “Ethnic conflict” another term where causation is built in?

Ethnic conflict refers to a very specific phenomenon. Ethnic conflict encompasses all forms of small- and large-scale acts of violence between and among different ethnic groups (Some definitions of ethnic conflict argue that ethnicity must also play a causal role in the conflict – something difficult to determine in practice. See Sambanis 2001; and Lake and Rothchild 1996). (Brancati 2006, 654)

304. Discuss these conceptualizations of electoral violence in terms of a causal relationship:

At the most abstract level, electoral violence can be understood as any event in which the use of coercive force coincides with the electoral process. A number of more contextually-specific definitions have also been advanced. The United Nations Development Program (UNDP) defines electoral conflict as: “acts or threats of coercion, or physical harm perpetrated to affect an electoral process or that arise in the context of electoral competition.” . . . There are two key components to all these definitions: the temporal link between violence and elections and the causal link between the two. Electoral violence is conventionally understood as violence that takes place during the electoral cycle, including the pre-electoral, electoral, and post-electoral periods. The causal link, which is often more implicit, limits electoral violence to that which is in some way connected to the electoral process, as opposed to violence that takes place during the electoral process but has no direct bearing on the election. (Birch and Muchlinski 2017, 2–3)

Therefore, the data include events of subnational electoral contention only if subnational elections and national elections took place in close temporal proximity and if it is plausible to expect that subnational contentious events influenced national elections or responded to both local and national elections. (Daxecker et al. n.d. Electoral Contention and Violence (ECAV) Codebook)

305. How does the adjective “civil” function here? Is it a subsetting one, a causal relationship, or what?

The next (and possibly most important) strand of the evolution of the understanding of peace is the civil peace. According to this approach, every individual in society has the capacity to mobilize for peace from a variety of different perspectives, whether for disarmament, for international cooperation, or against violence, discrimination, and oppression. It relates to the historical phenomena of social direct action for political, economic, and identity reasons, of citizen advocacy and mobilization, in the attainment or defence of basic human rights and values. It is also related to pacifism in its main forms, where civil action is non-violent
in principle. It has been strongly influenced by a wide range of social mobilization dynamics. Without the civil peace and its social forms of mobilization, international and constitutional frameworks would not be able to connect with ordinary people in order to represent their interests, identities, needs, and aspirations. (Richmond 2014, epub 116)

Answer:

It seems to function as a causal relationship, i.e., as cause of peace rather than concept of peace. It is about how actions of individuals can create peace.

306. There is often unclarity about whether something is a defining feature of the concept or rather the cause or effect of it. Discuss the extent to which emotions are a defining feature of war, a cause of war, or even ineffective of war.

The third defining element of war is, according to Clausewitz, the emotions that accompany and drive the extreme activity of waging war. At minimum, there must exist a drive to coerce the enemy using utmost force, as is evident in Clausewitz’s core definition of war as “an act of violence intended to compel our opponent to do our will” (75). War is fueled by—and gives rise to—intense emotions of hatred and enmity. Clausewitz argues that both the outbreak and the character of war are influenced by the types and intensity of emotions involved. The prevalence in a society of a warlike drive to dominate others is accordingly the third dimension of quality peace. Clausewitz notes that two sides will not prepare for and wage war against each other unless moved by emotions of enmity. But the type and intensity of emotion will vary, and the character of war will accordingly both react to the emotions and help shape those emotions. So, for example, the same political object may “elicit differing reactions from different peoples, and even from the same people over time.... Between two peoples and two states there can be such tensions, such a mass of inflammable material, that the slightest quarrel can produce a wholly disproportionate effect—a real explosion” (81). (Davenport et al. 2018, 262 epub)

307. Two level or SUIN models raise questions about what to include in the statistical analysis. An example on the independent variable side is Uzonyi’s concept of a “big political opportunity” which takes three forms that are expressed in his three hypotheses. Should he include each of the three forms separately in a regression analysis or just the higher variable formed by the logical OR?

There are generally three types of domestic unrest that may lead to the onset of genocide or politicide. The first, which is captured by those scholars focusing on big opportunities for government mass murder, are militarized threats to a country’s political leadership.... Coups provide a similar logic.... The third type of threat a government faces is non-militarized unrest. Non-militarized domestic unrest, such as protests, riots and strikes, can spark political violence if the regime perceives that such unrest will grow into greater instability. (Uzonyi 2016, 318–19)

308. Discuss the claim below that it is better to look at the superset than the subset.

For the sixty-two sovereign states included in the territorial contenders data set, Cunningham et al. identify ninety-seven rebel groups that controlled territory. The territorial contenders data set includes eighty-one
of them (83.5 percent). Of the remaining sixteen rebels identified by Cunningham et al. as territorial, five are included among the ancillary materials of the territorial contender data set as near misses. The other eleven are instances where our coders could not identify any clear evidence of territorial control. Consultation with the source documents for the NonState Actor Dataset reveals no citations establishing territorial control for the rebel groups in question. We believe the codings in the territorial contender data set are valid, while accepting that there may be some occasional error in either our or the non-state actor (NSA) data sets. An 84 percent overlap is very high, nonetheless. Absent an argument specific to a subset, there is considerable evidence in favor of analysis of the entire category of territorial contenders over analysis of any one subset. (Lemke and Crabtree 2019, 19, emphasis is mine)

309. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.

Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

310. In the Two cultures book a key methodological point is causal asymmetry. Hafner-Burton’s (2014) review article does not talk about authoritarian regimes as a source of human rights abuses, while she talk about democracy as a cause of the respect of human rights. Discuss this causal asymmetry.

The central insight about state institutions is that, while fully developed democracy may be the best hope for the promotion and protection of human rights, efforts to transition to and institutionalize democracy often incite violence, and the alternatives to democracy are not all equally adverse for human rights. (Hafner-Burton 2014, 275)

311. Blainey in his oft-cited book (1973) argues for causal symmetry in understanding war and peace. Discuss. For an opposing view see Goertz et al. The puzzle of peace: the evolution of peace in the international system where the final chapter argues for asymmetry.

the causes of war and peace, logically, should dovetail into one another. A weak explanation of why Europe was at peace will lead to a weak explanation of why Europe was at war. A valid diagnosis of war will be reflected in a valid diagnosis of peace. (p. 3)

War and peace appear to share the same framework of causes. The same set of factors should appear in explanations of the outbreak of war; widening of war by the entry of new nations; outbreak of peace; surmounting crises during a period of peace; and, of course, the ending of peace. (p. 293)

Wars usually end when the fighting nations agree on their relative strength, and wars usually begin when fighting nations disagree on their relative strength. Agreement or disagreement is shaped by the same set of causal
factors. Thus each factor that is a prominent cause of war can at times be a prominent cause of peace. (p. 122)

312. Geddes (2003) emphasizes the importance of variation on the dependent variable. But, variation on the independent variable is also important. Analyze her test of Skocpol’s theory in terms of variation on the two key independent variables.

Answer:
She selects only cases of agrarian revolt for her analysis so there is no variation on this independent variable. What she is in fact testing is the interaction of agrarian revolt and state breakdown (which does have variation). Typically, statistical interaction terms include variation on both variables.

313. Recently statisticians have become much more concerned with problems of “unit homogeneity.” Here is Henry Brady defining the idea:

We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and $Z_A$ and $Z_B$ [Z is the treatment] are identical as well (footnote: By saying that $Z_A$ and $Z_B$ have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven't indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [Identicality of units and treatment or Unit Homogeneity]. That is, A and B are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on B is the same as the impact of $Z = 0$ and $Z = 1$ on A. (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:
Holland (1986) says “unit homogeneity” means that units are prepared carefully “so that they ‘look’ identical in all relevant aspects” (Holland 1986, 948). To be “identical” means that the concept must produce identical units, A and B in the Brady analysis. If a study looks at all “states” then the conceptualization of state means that there are no casually important differences in “stateness.” Treatments, $Z$ in the Brady analysis, also are assumed to be homogeneous. So if democracy is an dichotomous independent variable that means that variations in democracy have no causal impact on the dependent variable.

314. As discussed at various points in Two cultures the set theoretic approach to causation is asymmetric. A variable might be good at explaining not-Y but weak at explaining Y. Take the following quotes and (1) formulate the strong generalization in terms of a necessary condition, (2) explain how asymmetry solves Hensel’s constant independent variable problem to a large degree. Note that correlation between contiguity and militarized disputes and wars is considered one of the strongest in the international conflict literature:

5 percent of fatal disputes [aka MIDs] and 5 percent of full-scale wars between minor powers begin between adversaries that do not share a land or sea boundary. (Hensel 2012, 6)
In short, proximity increases the risk of armed conflict. Proximity between states is basically a constant; countries rarely gain or lose borders over time. As a result the proximity explanation faces difficulties in accounting for the outbreak of rare events such as conflict and war, since a dependent variable cannot account for fluctuations in a dependent variable. (Hensel 2012, 8).

Answer:
Contiguity is a necessary condition for minor power wars.
Asymmetry arises because non-contiguity gives a very good explanation for the lack of severe militarized conflict, but does a poor job of identifying when wars actually break out. Notice that non-war is a almost constant for most dyads. This is of course the converse of the fact that wars are rare events. So both are very constant variables.

315. Discuss the causal figures in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

316. Discuss Pahre’s claim that necessary conditions violate common statistical assumptions: “Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that ‘if two units have the same value of the key explanatory variable, the expected value of the dependent variable will be the same’ (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition x ← y; in other words, x may occur with or without y” (Pahre 2005, 131).

317. A key assumption in multidimensional concepts is that the various means of achieving a given aggregate level are equivalent (e.g., 2+3 and 4+1 are really the same). What is one criterion for defending equivalence? Bueno de Mesquita et al. (2005) provides a very good discussion of these issues.

Answer:
One common criterion for equivalence is similar causal effect in some key hypotheses. Bueno de Mesquita et al. (2005) illustrates this.

318. Homogeneity of the zero cases when the dependent variable is coded dichotomously is often very problematic because they include quite different outcomes all coded as zero. Kroenig (2009) is looking at nuclear states supplying sensitive technology to non-nuclear states. It is quite likely that different causal mechanisms are producing these different zero outcomes. The dependent variable is received sensitive nuclear assistance or not. Discuss how the “not” case may contain two (or perhaps more) quite different kinds of zero.

Answer:
The zero cases include those countries who never tried to get such technology and those states who requested it but were turned down. See Kroenig’s
Appendix C “Selected cases of nonassistance” for some examples, such as “Argentina to Iran (1992). Argentina denied an Iranian request for hot cells (Hibbs 1992).”

319. Sometimes concepts should not build in causal hypotheses. Evaluate the Cohen and Pavoncello (1987) critique of Schmitter’s concept of corporatism based on the idea that there are causal hypotheses built into definition.

320. Thomas Franck is well-known for his work (e.g., 1988, 1990, 1992) on the importance of “legitimacy” in international politics. Discuss the theoretical relationship between the concept of legitimacy and four factors related to it. Is the relationship best seen as (1) indicator, (2) ontological, (3) causal, and if causal in which direct do the causal arrows go? For example, Franck says “one may postulate four indicators: pedigree, determinacy, coherence, and adherence. . . . The content of these four indicators of legitimacy” (Franck 1992, 51).

Answer:
In the 1992 article, Franck uses the term “indicator.” However, much of the discussion could lead one to see these four factors as causes of legitimacy.

321. A central guideline in building concepts is the consideration of causal relationships involving secondary-level dimensions of a concept. The causal relationships that involve the various secondary-level dimensions open up the possibility that there is significant causal heterogeneity in concepts and hence that causal generalizations are likely to be problematic. Little appears to be making this argument:

Another way to state my position is to consider whether there are social kinds, analogous to natural kinds. A natural kind is a set of entities that share a common causal structure, and whose behavior can therefore be predicted on the basis of the laws that govern the behavior of such entities (Putnam, 1975b). . . . I deny that any social concepts serve to identify social kinds in the strong sense outlined above. Instead, social concepts function as ideal types or cluster concepts, permitting us to classify a range of diverse phenomena under a single concept. The notion of a cluster term captures many scientific concepts-terms that encompass a variety of phenomena that share some among a cluster of properties (Putnam 1975a, 50–54). An ideal-type concept is a complex description of a group of social phenomena that emphasizes some features and abstracts from others (Weber 1949). It is apparent that generalizations and predictions based on cluster concepts and ideal types demand a great deal of care. Since the entities that fall under such concepts do not share a homogeneous causal structure, we cannot infer that instances of the concept will behave in the typical way. . . . The metaphysical counterpart, then, to my view that there are no governing social regularities among social phenomena, is that there are no genuine social natural kinds. (Little 1993, 190–91).

Discuss Little’s claims about “social kinds” versus “natural kinds.” Also, explore the extent to which “cluster concepts” are the same as family resemblance concepts.
322. The concept of a welfare state is used as an ongoing example in the concept book. In particular, it is good because the use of expenditure data illustrates the completely fungibility and substitutability of the secondary-level dimensions. It could be that this assumption of substitutability masks causal heterogeneity. See Bonoli (2007) who argues that causal processes dealing with “new social risks” (e.g., women’s participation in the labor force and child care) are quite different from those dealing with classic welfare policies (e.g., unemployment insurance). How should you change the concept of the welfare state? Add an adjective?

323. A very difficult conceptual and methodological problem is dealing with norms, particularly norms that have the logical structure of rights, e.g., human rights. This problem has exercised me a great deal and in fact resulted in a book (Goertz International norms and decision making; a large part of that book was devoted precisely to analyzing the concept of a norm). (1) Discuss the concept of a right (contrasting it with prohibitions and obligations). (2) The positive and negative poles are interesting here in terms of behavior: do you focus on the positive exercise of rights (e.g., voting) or on the absence of the exercise of the right. This has theoretical consequences because in the former case you focus on what factors in addition to rights lead to the use of rights, while in the latter you focus on factors that eliminate the right itself in practice. (3) Rights are about possibility, relate this to the Possibility Principle of case selection and necessary conditions in general. Dowding and Hees (2003) is a very nice discussion of many of these issues. For measurement see also Rosenbaum (2000) and Carter (2000).

324. One serious issue with dichotomous data is the potential for serious heterogeneity in the zero or one category. This is particularly likely to be a problem in the zero category. The issue of heterogeneity arises almost by definition with nonordered, categorical data. For example, quantitative studies of international conflict that look at the outcome dispute variables that have anywhere from 3 to 9 categories. Often these can be reduced to settlement/compromise, win, draw. If the focus is, say, settlement or compromise then there is a tendency to code that as one and lump win and draw together as zero. Discuss the merits of this procedure. What would be alternatives?
   Answer:
   Lumping all these together in the zero seems particularly dubious when the causal mechanisms producing these zeros are likely to be quite different.
   Within the context of the conflict literature and the militarized disputes dataset, it might make sense to just include draw/stalemate as the contrasting case since this constitutes often 70+% of the zero cases (e.g., Goertz et al. 2005).
   This choice could well be brought into the research design and theory if the research were interested in making specific causal contrasts.

325. One of the key guidelines for constructing concepts involves thinking about causal relationships within concepts. There is a huge literature on the concept of “ethnicity.” Discuss Chandra and Wilkinson’s figure 1 (2008, 524) on there
concept of ethnicity. Which of the various factors listed does not really involve any causal relationships with outside variables?

Answer:

The “structure” factor is the only which really is only and exclusively about ethnicity. All the others involve others causal variables. For example, see the discussion of ECI measure and its application to India; ethnic group variable is an interaction between some measure of ethnic “structure” and colonial practices.

“At the broadest level, we can imagine the term ethnicity as encompassing two families of concepts – the structure of ethnic identities and the practice of ethnic identification. Ethnic structure refers to the distribution of descent-based attributes – and, therefore, the sets of nominal identities – that all individuals in a population possess, whether they identify with them or not. Ethnic practice refers to the act of using one or more identities embedded in this structure to guide behavior. In other words, it refers to the set of activated identities that individuals employ in any given context. The set of activated ethnic identities for any given country is typically a subset of the identities contained in the ethnic structure.” Chandra and Wilkinson (2008, 523)

326. State failure is the subject of much academic research as well as interest to policy organizations such as the World Bank and the Carnegie Endowment, not to mention the large state failure project. One of the important risks with complex concepts such as “failed state” (see Iqbal and Starr 2009 for a review) is that some of the secondary-level dimensions of the concept may also be viewed as causes or effects of the phenomenon. Examine the various concepts of state failure and determine which might really be seen as causes or effects of state failure rather than the concept itself.

Answer:

For example, as Iqbal and Starr (2009, 316) state: “Specifically, most existing measures or indices of state failure incorporate a number of factors that may, in fact, be determinants of state collapse – such as civil strife and poverty.”

327. Birch (2012) in a book on the causes of “Electoral malpractice” includes the Freedom House democracy measure as a lagged (five year) independent variable. Explain why she runs the risk of “conceptual tautology” with her dependent variable of electoral malpractice.

328. Recently statisticians have become much more concerned with problems of “unit homogeneity.” Here is Henry Brady defining the idea:

We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and $Z_A$ and $Z_B$ [Z is the treatment] are identical as well(footnote: By saying that $Z_A$ and $Z_B$ have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven’t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [Identicality of units and treatment or Unit Homogeneity]. That is, A and B are mirror images of one another so that
the impact of \( Z_A = 1 \) and \( Z = 0 \) on \( B \) is the same as the impact of \( Z = 0 \) and \( Z = 1 \) on \( A \). (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:

Holland (1986) says “unit homogeneity” means that units are prepared carefully “so that they ‘look’ identical in all relevant aspects” (Holland 1986, 948). To be “identical” means that the concept must produce identical units, \( A \) and \( B \) in the Brady analysis. If a study looks at all “states” then the conceptualization of state means that there are no casually important differences in “stateness.” Treatments, \( Z \) in the Brady analysis, also are assumed to be homogeneous. So if democracy is an dichotomous independent variable that means that variations in democracy have no causal impact on the dependent variable.

329. One potentially large issue is what might be called “conceptual endogeneity.” This is where a secondary-level dimension is part of the dependent variable and also considered as an independent variable. There has been an explosion of studies of elections in hybrid democratic regimes. Should these elections be considered a cause of democratization? Does this conflict with the fact that freer and more competitive elections will change the democracy score of the country by definition? For example, “Scholars also disagree about the role of elections in promoting democratic change. (p. 12) While a familiar indicator of democratic progress, we must remember, electoral turnover does not always support democratic development. (p. 13) These issues lead to a final and more fundamental concern about attributing too much democratic influence to elections” (Bunce and Wolchik 2011, 14).

330. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when \( X_1 \) is a subset of \( X_2 \). For example, in the debate about the territorial peace versus democratic peace, it turns out that the set of democratic dyads is a subset of the dyads at territorial peace: “This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies” (Owsiak and Vásquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in \( X_1 \) is taken up in \( X_2 \)? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if \( X_1 \) is a perfect subset of \( X_2 \) and one introduces an interaction term, \( X_1 \times X_2 \)?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as \( X_1 \) less than or equal to \( X_2 \) for all observation or vice versa).
Causal mechanisms

331. The multimethod book espoused the mantra “no causation without causal mechanism.” This obviously plays a key role in exceptional studies where there may be randomization treatments but where the result does not match any causal mechanism intuition. The question then is should the journal publish the research when there seems to be no plausible mechanism that would explain the results. Here’s a famous example in psychology. The Journal of Conflict Resolution had a similar controversial publication regarding the publication of an article “International Peace Project in the Middle East: The Effects of the Maharishi Technology of the Unified Field.” (Orme-Johnson et al. 1988). Russett (2017) describes in some detail the controversies and eventual publication of the article.

History may look back on 2011 as the year that changed psychology forever. It all began when the Journal of Personality and Social Psychology published an article called “Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect.” The paper, written by Daryl Bem of Cornell University, reported a series of experiments on psi or “precognition,” a supernatural phenomenon that supposedly enables people to see events in the future. Bem, himself a reputable psychologist, took an innovative approach to studying psi. Instead of using discredited parapsychological methods such as card tasks or dice tests, he selected a series of gold-standard psychological techniques and modified them in clever ways. . . . On this basis, Bem argued that the participants were able to benefit in the past from practice they had completed in the future. As you might expect, Bem’s results generated a flood of confusion and controversy. How could an event in the future possibly influence someone’s reaction time or memory in the past? If precognition truly did exist, in even a tiny minority of the population, how is it that casinos or stock markets turn profits? And how could such a bizarre conclusion find a home in a reputable scientific journal? (Chambers 2017, 2–3)


Representative political institutions may improve a government’s ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the
political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

333. The idea of a causal mechanism being a causal chain is very wide spread. Draw a causal mechanism figure for ethnic civil war according to Lindemann and Wimmer. For example, is each link necessary for the following one. How would you diagram the idea that different conditions are relevant in different stages?

Overall, this suggests that the conditions are not ordered in a historical sequence of causal connections. The historically prior condition for ethnic war does not causally generate the second condition for ethnic war. In other words, the logic here is not that of a necessary chain reaction. Rather, the uniform temporal ordering in which conditions become relevant is due to the logic of the escalation process itself, which makes certain conditions relevant at different points in the process, but does not ensure their emergence. The following stages of this escalation process could be hypothesized: mobilization of ethno-nationalists, their initial organization, subsequent radicalization and mass recruitment, and, finally, the emergence of armed organizations.

Different conditions become relevant during these subsequent stages. Further ‘grievance’-inducing conditions are relevant for mobilization. Limited state reach might be relevant for initial organization. Indiscriminate violence becomes relevant for subsequent radicalization and mass recruitment – without indiscriminate violence the process of escalation stops here. And finally, an external sanctuary is often relevant for the actual organization of an armed guerrilla front. (Lindemann and Wimmer 2018, 317)

334. Social constructivists frequently uses the term “permissive cause.” Carpenter (2007) discusses agenda-setting in terms of constraints and her causal mechanism figure, figure 1. p. 102. So the absence of permissive condition explains why children of rape is not on the agenda? What do the lines mean (i.e., not arrows). Do they just indicate temporal ordering?

335. The Keck and Sinkink (1998) boomerang model of social movements have been very influential. Think about the figure used to describe the causal mechanism, not an exhaustive list: (1) arrows versus lines what is the difference? (2) some arrows get “pressure” others do not, why? (3) is thickness of arrow an hypothesis about relative causal impact? (4) arrows between NGOs in State B, but not State A, why? (5) what do dotted lines (NGOs to blockage) mean? (6) are there arrows missing, e.g., from NGOs in State A to IGO?

336. Pevehouse is a main example in the discussion of causal mechanisms and multimethod research. One issue is that democratic IGO is relatively constant. For that mechanism to kick in it probably needs some triggering mechanism, e.g., crisis, coup attempt, shock. Add that to the Pevehouse figure to give a more complete mechanism; “What many studies of the liberalization process have in common is their treatment of the impetus for liberalization as an exogenous shock. These shocks may be political or economic in nature, but either can force elites to take some action to restore the legitimacy of their
regime. Disagreements then arise within the authoritarian bloc as to the prudent course of action. Some regimes may be able to weather the crisis given a variety of factors, ranging from the nature of the current autocratic regime, economic conditions, to the past performance of the regime. Other regimes may decide to liberalize in an attempt to restore legitimacy.” (Pevehouse 2005, 17)

337. Fearon and Laitin in their oft-cited work (2003) on civil war (in)famously include a variable on amount of mountainous terrain (see Collier and Hoeffler 2004). This is controversial in terms of causal inference because it essentially does not vary within countries. Influential causal inference scholars deny causal status to variables which cannot be manipulated. Discuss how one use large-N qualitative case studies to explore the causal mechanism linking mountainous terrain to civil war. Can this illustrate how a constant can be part of a causal mechanism?

338. Causal mechanisms and scope decisions can be closely related. Bush's main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.

339. Causal mechanisms and scope decisions can be closely related. Bush's main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy
abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.

A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

Consider the following path diagram. $X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y$ We are apt to consider $X_1$ to be the cause and causal factors $X_2 - X_4$ intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to $X_1$. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are ‘ultimate’ or ‘fundamental’.

Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.

The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find ‘structural’ variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between X and Y are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

The AMAR “rebellion” scale (Birnir et al. 2018) has as its lowest rebellion level “political banditry and/or sporadic terrorism.” Often zero on the scale (like in
Soifer (2012) explores the concept of critical juncture along with critical antecedents in the case of import substitution policies in Latin America.

The distinct feature of a historical juncture with the potential to be critical is the loosening of the constraints of structure to allow for agency or contingency to shape divergence from the past, or divergence across cases. . . . we must distinguish between two types of causal conditions at work during the critical juncture: the permissive conditions that represent the easing of the constraints of structure and make change possible and the productive conditions that, in the presence of the permissive conditions, produce the outcome or range of outcomes that are then reproduced after the permissive conditions disappear and the juncture comes to a close. The two types of conditions are nearly always framed as separately necessary and jointly sufficient for divergence to occur. (Soifer 2012, 1573)

They [Slater and Simmons 2010] define a critical antecedent as “factors or conditions preceding a critical juncture that combine in a causal sequence with factors operating during that juncture to produce a divergent outcome” (p. 889) . . . . They write that the critical antecedent “does not produce its causal effect by causing the independent variable to emerge. It does so by helping to determine the differential causal effect of the independent variable across cases when the critical juncture exogenously comes about” (p. 891). Here we see that the critical antecedent is, in their formulation, unrelated to the permissive condition, which (for them) emerges exogenously. Permissive conditions are also distinct from critical antecedents in that although the former mark the temporal bounds of the critical juncture, the latter are operant before the juncture emerges. On the other hand, critical antecedents are connected to the productive condition. (Soifer 2012, 1576)

Draw a causal mechanism figure for the table below.

<table>
<thead>
<tr>
<th>Critical antecedent</th>
<th>Permissive condition</th>
<th>Productive condition</th>
<th>Outcome</th>
<th>End of critical juncture</th>
<th>Mechanisms of reproduction</th>
<th>Consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strength of middle class and labor as of 1929</td>
<td>Collapse of world trade and economic challenges of Great Depression and World War II</td>
<td>Economic ideas of ECLA and more general rise of economic nationalism</td>
<td>Inward-looking industrialization implemented to varying degrees in Latin America</td>
<td>Recovery of world trade by 1950, and especially after the Korean War</td>
<td>New political coalitions among bureaucrats, domestic elites, and organized labor</td>
</tr>
</tbody>
</table>
Slater and Simmons (2010) provide quite a few causal mechanism figures, such as

Figure 2. McAdam on King Cotton and Black insurgency
Figure 3. Grzymała-Busse on Communist Party regeneration
Figure 4. Family law and gender inequality in North Africa.
Figure 5. The Colliers on labour incorporation in Latin America
Figure 6. Luebbert on interwar European regime outcomes
Figure 7. Kalyvas on Christian Democratic parties in Western Europe
Figure 8. Regime outcomes in Central America [Yashar, Mahoney]
Figure 9. Ethnic violence in India [Varshney, Wilkenson, Kohli]
Figure 10. Racial politics in Brazil and South Africa [Marx, Lieberman]

Discuss these causal mechanism figures. For example, Do the arrows all mean the same think? Are some of the ellipses necessary conditions? What are the causal connections between text with boxes in front of them and the ellipses?

For example, here is the McAdams figure.

**Figure 2. McAdam on King Cotton and Black insurgency**

Mansfield and Pevehouse (2000) give a variety of mechanisms that could link preferential trade agreements to reduce military conflict. Do a causal mechanism graph of their theory (like those in chapter 2 of *Multimethod research, causal mechanisms, and case studies*, note that the core hypothesis is an interaction term between trade and PTA membership. A similar exercise can be done with the causal mechanisms leading from common IGO membership to less international conflict: “Indeed, a close look at various international organizations indicates they may serve any of six functions: coercing norm breakers; mediating among conflicting parties; reducing uncertainty by conveying information; problem-solving, including expanding states’ conception
of their self-interest to be more inclusive and long-term; socialization and shaping norms; and generating narratives of mutual identification. (Russett et al. 1998, 444–45). Discuss whether these “functional” arrows in the causal mechanism diagram should be interpreted as causal.

345. Discuss the causal figures in *Multimethod research, causal mechanisms, and case studies*, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

346. Grzymała-Busse (2007) is a core example in the causal mechanism chapter of *Multimethod research, causal mechanisms, and case studies: an integrated approach*. In my figure 2.4 of her theory “Robust competition” is constituted by three mechanisms (1) moderation, (2) anticipation, and (3) cooptation. Discuss how on might aggregate or combine these causal mechanisms (e.g., OR, AND, addition). Mikkelsen (2017) uses Grzymała-Busse (2007) a core example in his discussion of fuzzy logic case studies. Compare and discuss his version of the mechanism of competition, e.g., Figure 2 and how he uses fuzzy logic to combine the three mechanisms with the discussion in chapter 3 of *Multimethod research, causal mechanisms, and case studies: an integrated approach* of the same mechanism.

347. Different causal mechanisms and control variables. Sometimes it is not clear whether something is a separate causal mechanism (e.g., separate X) or whether it is part of some larger causal mechanism, i.e., an $M_i$ factor. For example, Simmons and Elkins (2004) in a widely cited article discuss various mechanisms of policy “diffusion.” (I use the scare quotes because it is not clear that they are all diffusion mechanisms, not other kinds of mechanisms, e.g., market competition.) They contrast three theories about why states adopt liberal economic policies. The first is via market competition, if competitor nations are adopting liberal policies and the state adopts to compete in markets. A second mechanism is global norms about economic policies. A third is learning from networks, successful states, or culturally similar states. There are independent variables in their statistical model representing each of these mechanisms. So should these independent variables be considered separate mechanisms, e.g., separate $X_i$ or $M_i$ within a larger theory of diffusion?

348. The $\rightarrow$ in causal mechanism figures is often ambiguous. Sometimes it can mean a noncausal relationship. What would be some situations where this is the case?

Answer: One set of situations is where the arrow is ontological or definitional in nature. Another is game trees or other decisionmaking figures where the arrows indicate different choice options.
Constraint mechanisms

349. Social constructivists frequently use the term “permissive cause.” Typical of social constructivists. Think Carpenter (2007) discusses agenda-setting in terms of constraints and her causal mechanism figure, figure 1. p. 102. So the absence of permissive condition explains why children of rape is not on the agenda? What do the lines mean (i.e., not arrows). Do they just indicate temporal ordering?

350. Discuss the Rodrik (2007) figure below. Note that everything is conceptualized as a constraint (see chapters 2 and 4 of Multimethod research, causal mechanisms, and case studies: an integrated approach). Which case studies would be useful? Which ones does he do? Discuss the nature of the arrows – e.g., are they causal – and the two-level structure of the causal mechanism. Could the arrows be casually going up instead of down?

Moving down the branches of the decision tree is tantamount to discarding candidates for the most binding constraint on growth. The overarching lesson from our theoretical analysis is that it is this constraint, once identified, that deserves the most attention from policymakers. (Rodrik 2007, 66)

Fig. 2.1. Growth diagnostics

Process tracing, within case causal inference, counterfactuals

351. It is not uncommon to confuse necessary condition counterfactuals with sufficient condition ones. See Fenoaltea (1973) for a beautiful discussion of this in
the context of economic history. Fischer's classic book (1970) also deals with this. Discuss Fisher's examples.

One common form of the reductive fallacy is the confusion of necessary with sufficient cause – the confusion of a causal component without which an effect will not occur, with all the other causal components which are required to make it occur. This sort of error appears in causal explanations which are constructed like a single chain and stretched taut across a vast chasm of complexity. The classic example is the legendary battle that was lost for the want of a horseshoe nail; for the the want of a nail the shoe was lost, for the want of a shoe the horse was lost, for the want of a horse the rider was lost, for the want of a rider the message was lost, for the want of the message the regiment was lost, and for the want of a regiment the battle was lost. (Fischer 1970, 172; he then goes on to use the classic example of the battle of Antietam for which the accident loss of confederate orders to the North was a key factor in the North's victory)

A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

Consider the following path diagram. $X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y$ We are apt to consider $X_1$ to be the cause and causal factors $X_2 - X_4$ intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to $X_1$. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are ‘ultimate’ or ‘fundamental’.

Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.

The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find ‘structural’ variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between
X and Y are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

353. There is significant debate about path dependency, historical contingency and the like. Explain why the “for example counterfactual” is different than the historical contingency definition preceding it.

By historical contingency we mean chance-influenced events with substantial long-term effects, that is, events which clearly take history down a different path than it otherwise would have followed. . . . For example, had Empress Elizabeth died a month later, Prussia might have lost the Seven Years’ War. The chance event caused a different path to be taken. (Bendor and Shapiro 2019, 131)

354. Discuss the validity and methodology of a core counterfactual from Simmons’s very influential book on human rights:

No available secondary source account argues that Japan would have passed legislation to address some of the most egregious aspects of discrimination against women in employment had not the question of Japan’s position on the CEDAW come up. Clearly, this was a case in which an externally negotiated agreement changed the country’s legislative agenda, placing issues of women’s equality much higher on the list of legislative priorities than would have existed had Japan not been faced with the issue of what to do about the CEDAW. To be sure, women’s equality had its domestic supporters, but they were clearly in a minority and somewhat isolated politically and bureaucratically. They were hardly a match for Japan’s powerful business interests that wanted to maintain the status quo. Second, the existence of the CEDAW increased the size of the coalition that was to support the EEOL. The Ministry of Foreign Affairs, for example, would not normally weigh in on domestic legislation on equality in women’s employment, but they did in this case because they wanted Japan on board with this major UN initiative. (Simmons 2009, 240)

355. Discuss the “light-switch” analogy as a way to think about process tracing and equifinality. The basic analogy is that the light-switch is on and the light is on (essentially the (1,1) cases). Discuss process tracing as following electricity through various junction boxes to the light bulb. Equifinality lies in the fact that there might be multiple switches that can possibly control the same light. Discuss what would be the analogy in case study work of “turning on or off” the switch (within cases or between cases). What about counterfactuals? (Thanks to Mike Desch for suggesting this analogy)

356. In the literature on within-case causal inference researchers often talk about hoop tests and smoking gun tests in terms of “certainty” and “uniqueness”: “The first dimension is called certitude or certainty and captures how likely it is to confirm a specific observable implication in process tracing. The second dimension is called uniqueness and asks whether an observable implication can be derived from a single or multiple hypotheses” (Rohlfing 2014 table 1).
Rohlfing goes on to claim: “A hoop test is characterized by high certainty and no uniqueness.” (Rohlfing 2014, 612) Describe a situation where a hoop test has a high degree of uniqueness.

Answer:
As the hoop test gets closer to being a sufficient condition, its uniqueness increases, i.e., when the hoop is very small.

357. Use the democratic peace to propose a specific counterfactual so that World War I does not occur.

358. Discuss the following claim by one of the most prominent historians today. Does it invoke a necessary or sufficient condition claim? Do a counterfactual analysis of it.

The argument that British intervention in the war was made inevitable by the violation of Belgian neutrality has been repeated by historians ever since [Lloyd George]. (Ferguson 1999a: 231)

359. The confusion of necessary with sufficient conditions also arises when defining what “cause” means. Explain how this works in Wendt’s analysis of cause:

In saying that ‘X causes Y’ we assume that: (1) X and Y exist independent of each other, (2) X precedes Y temporally, and (3) but for X, Y would not have occurred. . . . The logical empiricist model of causal explanation, usually called the deductive-nomological model or D-N model, is rooted in David Hume’s seminal discussion of causality. Hume argued that when we see putative causes followed by effects, i.e., when we have met conditions (1) and (2), all we can be certain about is that they stand in relations of constant conjunction. The actual mechanism by which X causes Y is not observable (and thus uncertain), and appeal to it is therefore epistemically illegitimate. Even if there is necessity in nature, we cannot know it. How then to satisfy the third, counterfactual condition for causality, which implies necessity? (Wendt 1999: 79)

360. Quite a few different terms are often used to distinguish between different kinds of causes, e.g., “contributory cause,” “remote cause,” “intervening/mediating cause.” Use Russett’s very interesting analysis of World War I (1962) – in particular Fay’s account – to discuss what these kinds of cause mean. See also Fischer’s classic book (1970) which talks about eight kinds of cause: (1) all antecedents, (2) regularistic antecedents, (3) controllable antecedents, (4) rational and/or motivational antecedents, (5) abnormal antecedents, (6) structural antecedents, (7) contingent-series antecedents, (8) precipitant antecedents.

361. Brooks and Wohlforth argue for a probabilistic interpretation of their causal claims in their response to English. For the philosophically inclined, there is a major debate in philosophy about the “sense” that probably statements have for individual cases. Many argue that probability statements only make sense in over repeated events. This is called the “frequentist” position, e.g., “Our probability theory has nothing to do with questions such as: ‘is there
a probability of Germany being at some time in the future involved in a war with Liberia.’ . . . probability when it refers to a single event has no meaning at all for us” (von Mises 1957, 9, 11).

362. Van Evera’s well-known article (1984) on World War I argues that the “cult of the offensive” was an important cause of the war. Find three key counterfactuals in that article that imply necessary condition hypotheses. One of the three should involve the “window of opportunity” idea.

363. What is the logical form of the hypothesis that U.S. military pressure would put such a strain on the Soviet Union “that the Soviet leadership would have little choice but to make substantial concessions on arms control” (Gaddis 1989, 13)?

364. Many historians feel that events like World War I were “over-determined.” Explain why this would generally lead them to discount necessary condition counterfactuals, e.g., “The fact that so many plausible explanations for the outbreak of the war [World War I] have been advanced over the years indicates on the one hand that it was massively overdetermined, and on the other that no effort to analyze the causal factors involved can ever fully succeed” (Schroeder 1972, 320).

365. Define the contrapositive: (1) starting with necessary conditions, (2) starting with sufficient conditions, (3) with multiple necessary conditions, (4) with multiple sufficient conditions.

366. The Schroeder chapter in the Goertz and Levy anthology is a brilliant example of how one can use the contrapositive to make a very provocative claim. (1) Explain what he did. (2) Do the same thing with a causal claim about the end of the Cold War.

367. The following quote summarizes the argument of Sagan (1986) on the causes of World War I:

Sir Edward Grey’s failure to present a clear and credible threat of British intervention early in the July crisis and the specific preemptive aspects of Germany’s offensive war plans caused by the slow Russian mobilization and the Liege bottleneck are linked together as an immediate cause of the First World War. (Sagan 1986: 168; this is key thesis of whole article).

In the conclusions Sagan proposes the following counterfactual which is potentially logically inconsistent with the basic theory given above. What is that potential logical inconsistency? Can you give an interpretation which makes everything consistent?

If Grey had given a clear warning earlier, if the Czar had further delayed Russian mobilization against Austria and then Germany, and if the German offensive war plans had not been able to depend upon a preemptive coup de main against Liege and the decisive battle in France before Russian mobilization was completed in the East, then it is possible, just possible, that Bethmann-Hollweg would have had the time and the courage necessary to apply sufficient pressure on Vienna to accept the “Halt in
Belgrade.” And if this had occurred, 1914 might today appear as only another one of a series of Balkan crises that almost led to a world war. (Sagan 1986, 169).

368. A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

Consider the following path diagram. \( X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y \) We are apt to consider \( X_1 \) to be the cause and causal factors \( X_2 - X_4 \) intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g., if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to \( X_1 \). Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are ‘ultimate’ or ‘fundamental’.

Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.

The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find ‘structural’ variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between \( X \) and \( Y \) are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

369. The following is a well-known nursery verse that describes King Richard III’s fall from power. Analyze the temporal distance causal claims embedded in it.

For want of a nail the shoe was lost,  
For want of a shoe the horse was lost,
For want of a horse the rider was lost,
For want of a rider the battle was lost,
For want of a battle the kingdom was lost,
And all for the want of a horseshoe nail.

Answer:
Fischer’s classic book on historical explanation (1970) discusses this example. See Lowe (1980) for a more technical philosophical analysis that focuses on counterfactuals and causal explanations.

370. United States law gives particular importance to the person or action that was the last one which could have potentially avoided a bad event (e.g., car accident). What does this mean about the relative causal importance of recent or distant events.

371. How can you use the powder keg model to combine the arguments of English and Brooks/Wohlforth debate about the end of the Cold War?

Answer:
Basically the Soviet decline is the powder keg which needs the spark of new ideas and leadership. This idea is very common in the social constructivist literature, e.g., “fire [powder keg] metaphor serves for the types of interactions that we describe throughout this book. There must be an idea, advocates to spread the idea, and a public ready to receive it” (Keck and Sikkink 1998: 46).

372. A key issue in the concept of a turning point revolves around whether a turning point should be defined in terms of points in history where things changed or should also include “potential” turning points but where things did not change. For example, Herrmann and Lebow do not include potential turn points in their definition: “We define a turning point in terms of two properties. First, it must be a change of significant magnitude, not an incremental adjustment but a substantial departure from previous practice. Second, it must be a change that would be difficult to undo” (Herrmann and Lebow 2004). Discuss the following example from A.J.P. Taylor regarding a turning point in German history where Germany did not turn:

1848 was the decisive year of German, and so of European, history: it recapitulated Germany’s past and anticipated Germany’s future. Echoes of the Holy Roman Empire merged into a prelude of the Nazi ‘New Order’; the doctrines of Rousseau and the doctrines of Marx, the shade of Luther and the shadow of Hitler, jostled each other in bewildering succession. Never has there been a revolution so inspired by a limitless faith in the power of ideas; never has a revolution so discredited the power of ideas in its result. The success of the revolution discredited conservative ideas; the failure of the revolution discredited liberal ideas. After it, nothing remained but the idea of Force, and this idea stood at the helm of Germany history from then on. For the first time since 1521, the German people stepped on to the centre of the German stage only to miss their cues once more. German history reached its turning-point and failed to turn. This was the fateful essence of 1848. (Taylor 1945: 68; for a different perspective see Blackbourn and Eley 1984)
Use the idea of a window of opportunity to define all potential turning points. What is the role of necessary versus sufficient conditions?

Many historians feel that the “Pleiku” analysis of World War I is basically correct (i.e., that there were many Balkan crises and that one of them would have sparked a world war). Some similar arguments have been made with regard to the impact of the Versailles Treaty on the occurrence of World War II, that is made World War II very likely. What would be one counterfactual argument that strongly contests such views? Are there parallels to the end of the Cold War debate?

Answer:
One counterfactual would stress the absolutely central role of Hitler. For example, Mueller has in several places used this idea. For example,

Thus it is simply not true that the seeds of another great war were planted at the peace conference of 1919. In order to bring about another major war it was necessary for Germany, first, to desire to expand into non-German areas, second, to be willing to risk and threaten military action in order to get these areas, and third, to be willing to pursue war when these desires were opposed by other major countries. It seems to me that none of these propositions – particularly the last two – were very popular in Hitler’s Germany, that almost no one accepted all of them, and that only Hitler, it appears, combined a fanatical acceptance of them with a maniacally determined and effective capacity to carry them out. (Mueller 1991, 22, see also Mueller 1989; Byman and Pollack 2001, 118)

Can you make counterfactual arguments about the role of key individual decision-makers in the World War I case? End of Cold War?

Answer:
Obviously this is a question about which historians have debated for years.

Byman and Pollack sees German and Austrian leadership in these terms, “Absent the Iron Chancellor, it is hard to imagine a defeated Austria aligning with Prussia after the humiliations of Sadowa and Königgrätz. Similarly, it is equally hard to imagine a leader other than Wilhelm II repeatedly antagonizing Britain for so little purpose” (2001: 134).

Brown argues for the essential role of Gorbachev by arguing against the sufficiency of alternative explanations: “None of these interpretations [(1) learning, (2) power, and (3) domestic political pressure] on its own provides an adequate explanation of the changing Soviet political agenda in the second half of the 1980s, although the first and the third – or a combination of the two – have often been seen as sufficient by their proponents (p. 13) . . . It is tempting to see Gorbachev as the handmaiden of history or the embodiment of social forces which, if Gorbachev rather than Dmitry Ustinov had died in December 1984, would have brought forth an alternative leader in the mid-1980s whose policies would have been broadly the same as Gorbachev’s producing similar results. This, however, is a temptation that should be resisted, for it has got little but a ‘retrospective determinism’ to commend it” (Brown 1996, 317).
Here are some claims that constitute core ideas of prominent work (both by political scientists and historians) on World War I (I do not include those defended in various chapters of the Goertz and Levy anthology). In all cases these are central claims made by the author(s). The quotes below are obviously dramatic simplifications of complex arguments but they do capture the core of the scholar’s argument and are useful for classroom discussion. Some are expressed in necessary condition counterfactual terms. Others are not; for these generate some necessary condition counterfactuals. More generally, analyze the causal explanations in them. These can be used as the basis for class discussion or additional reading.

Goemans:
I have to put the blame for four years of war and over ten million deaths squarely on the shoulders of Germany’s regime. If Germany had not been a semirepressive, moderately exclusionary regime, would the war have ended sooner? Such counterfactual claims are notoriously hard to evaluate. . . . An argument can nevertheless be made that a nonrepressive, nonexclusionary Germany would have ended the war before 1918, with 1917 a likely termination year. (Goemans 2000, 314–15, in conclusions)

Ferguson:
Rather than joining the Allied war effect, he [Ferguson] said, Britain should have maintained its neutrality and allowed the Germans to win a limited Continental war against the French and the Germans. In that event, he postulated, Germany whose war aims in 1914 were relatively modest, would have respected the territorial integrity of Belgium, France, and Holland and settled for a German-led European federation. Had Britain “stood aside” he continued, it is likely that the century would have been spared the Bolshevik Revolution, the Second World War, and perhaps even the Holocaust. (Boynton 1999, 43; New Yorker piece on Ferguson’s book)

Geiss:
The events of July and early August 1914 cannot be properly understood without a knowledge of the historical background provided by the preceding decades of Imperialism. On the other hand, that background alone is not sufficient to explain the outbreak of the First World War. Two general historical factors proved to be decisive, and both were fused by a third to produce the explosion known as the First World War. Imperialism, with Wilhelmine Weltpolitik as its specifically German version, provided the general framework and the basic tensions; the principle of national self-determination constituted, with its revolutionary potential, a permanent but latent threat to the old dynastic empires and built-up tensions in south-east Europe. The determination of the German Empire – then the most powerful conservative force in the world after Tsarist Russia – to uphold the conservative and monarchic principles by any means against the rising flood of democracy, plus its Weltpolitik made war inevitable. (Geiss 1984, 46; original is the introduction of his book July 1914)

Taylor:
But there was only one decision which turned the little Balkan conflict between
Austria-Hungary and Serbia into a European war. That was the German decision to start general mobilization on 31st July, and that was in turn decisive because of the academic ingenuity with which Schlieffen, now in his grave, had attempted to solve the problem of a two-front war. (Taylor 1969, 101)

Maier:
The irreversible momentum toward general war in 1914 is usually seen as a result of three factors: the hopeless, long-term instability of the Habsburg empire, the rigid structure of opposing alliances, and the ineluctable pull of military preparations. (Maier 1988, 822)

Williamson:
Of all the central actors in 1914, Conrad alone could have – by saying no to Berchtold or expressing hesitation to Franz Joseph or accepting some modified “Halt in Belgrade” – brought the crisis to a more peaceful conclusion. (Williamson 1988, 815–16)

Van Evera:
The consequences of the cult of the offensive are illuminated by imagining the politics of 1914 had European leaders recognized the actual power of the defense. . . . Thus the logic that led Germany to provoke the 1914 crisis would have been undermined, and the chain reaction by which the war spread outward from the Balkans would have been very improbable. In all likelihood, the Austro-Serbian conflict would have been a minor and soon-forgotten disturbance on the periphery of European politics. (Van Evera 1984, 105)

Snyder:
These war plans and the offensive doctrines behind them were in themselves an important and perhaps decisive cause of the war. (Snyder 1984, 108) In short, the European militaries cannot be blamed for the belligerent diplomacy that set the ball rolling towards World War I. Once the process began, however, their penchant for offense and their quickness to view war as inevitable created a slide towards war that the diplomats did not foresee. (Snyder 1984, 138–39)

Trachtenberg:
If in 1914 everyone understood the system and knew, for example, that a Russian or German general mobilization would lead to war, and if, in addition, the political authorities were free agents – that is, if their hands were not being forced by military imperatives, or by pressure from the generals – then the existence of the system of interlocking mobilization plans could hardly be said in itself to have been a “cause” of war because, once it was set off, the time for negotiation was cut short. But if the working of the system was understood in advance, a decision for general mobilization was a decision for war; statesmen would be opting for war with their eyes open. To argue that the system was, in such a case, a “cause” of war makes about as much sense as saying that any military operation which marked the effective beginning of hostilities . . . was a real “cause” of an armed conflict (Trachtenberg 1990–91, 122)

Sagan:
If Grey had given a clear warning earlier, if the Czar had further delayed Russian mobilization against Austria and then Germany, and if the German offensive war plans had not been able to depend upon a preemptive coup
*de main* against Liege and the decisive battle in France before Russian mobilization was completed in the East, *then* it is possible, just possible, that Bethmann-Hollweg would have had the time and the courage necessary to apply sufficient pressure on Vienna to accept the “Halt in Belgrade.” And if this had occurred, 1914 might today appear as only another one of series of Balkan crises that almost led to a world war. (Sagan 1986, 169)

Riker:
The process of confusion concerning the cause of α [World War I], which, being ambiguous, cannot properly be said to have a cause, starts with the assertion that the cause of B [the Austro-Serbian war] is the assassination at Sarajevo, event D. It can perhaps be demonstrated by valid arguments—or at least by arguments likely to be accepted by many historians—that the assassination of Franz Ferdinand is a sufficient condition of the Austro-Serbian war; and (somewhat dubiously, however) it can even be argued that it is was a necessary condition. (Riker 1957, 63–64)

### Generalization, external validity, Large-N Qualitative Analysis (LNQA), scope, etc.

377. The notion of “conceptual replication” is wide-spread in psychology. In fact, conceptual replication is in fact theoretical generalization. Discuss the merits of this notion. Chambers provides an example of someone who thinks it is a bad idea because it is not actual replication of the original experiment.

> Although used widely in psychology, the term conceptual replication does not feature in the scientific method of other disciplines. In fact, the term itself is misleading because conceptual replications don’t actually replicate previous experiments; they instead assume (rather than test) the truth of the findings revealed in those experiments, infer the underlying cause, and then seek converging evidence for that inferred cause using an entirely different experimental procedure. Viewed within the framework of the HD scientific method, this process can be thought of as extrapolating from a body of findings to refine theory and generate new hypotheses (see figure 3.1). As important as this step is, it depends first and foremost on the reliability of the underlying evidence base. To rely solely on extrapolation at the expense of direct replication is to build a house on sand. (Chambers 2017, 49)

378. Sechser and Fuhrmann 2017 is an excellent example of large-N qualitative analysis. Discuss their use of case studies first the (1,1) cases then the set of cases that potentially falsify their theory:

> The previous chapter analyzed clear failures of nuclear coercion. We showed that in six serious nuclear crises, and three less severe cases, countries failed to coerce their adversaries despite engaging in nuclear brinkmanship. In this chapter, we turn our attention to cases in which nuclear blackmail seemingly worked. These crises are apparent outliers for nuclear skepticism theory and fertile ground for evidence in support of the nuclear coercionist school. Indeed, the leaders involved in these
episodes – as well as many scholars who have studied them – asserted that nuclear weapons were useful for military blackmail. . . . We critically assess ten crises that are widely regarded as successful instances of nuclear coercion, focusing our attention on the seven most serious cases. None of these crises provides unequivocal evidence that nuclear coercion works. In each instance, at least one of three factors mitigates the conclusion that a nuclear threat resulted in a coercive victory. First, factors other than nuclear weapons often played a significant role in states’ decisions to back down. Second, on close inspection, some crisis outcomes were not truly “victories” for the coercer. Third, when nuclear weapons have helped countries in crises, they have aided in deterrence rather than coercion. (Sechser and Fuhrmann 2017, 172–73)

379. A key issue in many experiments is extrapolation to the population of interest. Discuss Lupton’s discussion of these issues:

Furthermore, our theories about how people think should be broadly applicable across individuals from a variety of backgrounds. Researchers can also employ sample populations that mirror the demographics of policymakers in certain ways. The samples I use skew male, as do world leaders more generally, and are rather highly educated (with almost 90 percent of participants in my surveys having some form of a college education—compared with 80 percent of world leaders). The majority of participants in my study are also interested in international affairs and international events. While I do not claim that my sample populations perfectly represent the demographics of policymakers, the participants in my study are not as uneducated or uninterested in politics as one might initially think. Moreover, I randomly assign participants to different treatment groups, which helps address the issue of external validity. (Lupton 2019, 52)

380. Concept generalization across groups is a core methodological issue. Discuss the sense of “meaning” in the abstract and how it relates to generalization of a concept across groups.

Measurement invariance assesses the (psychometric) equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups or across repeated measurements . . . . Appropriate and proper comparison of a construct between groups or across times, therefore, depends first on ensuring equivalence of meaning of the construct. The untoward consequences of measurement noninvariance can be readily illustrated in the study of depression in men and women. Suppose frequency of crying, weight gain, and feelings of hopelessness are indicative of the severity of depression in women, but only feelings of hopelessness are indicative of the severity of depression in men. If the three indicators are combined into a scale to compare depression in women and men, mean differences on the scale may mislead because crying and weight gain have little relation to depression in men. In this example, men may score lower than women on the depression scale because they cry less and gain less weight. (Putnick and Bornstein 2016, 71)

381. Discuss case selection in Stasavage’s influential book Public debt and the birth of the democratic state France and Great Britain, 1688–1789. Clearly he is using
the very popular paired comparison, in terms the causal mechanism book a (0,0) case compared to a (1,1) case. There are lots of nondemocracies in this. Why choose France? What about choosing more (1,1) cases to generalize?

Representative political institutions may improve a government's ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

382. Large-N qualitative analysis can become particularly attractive when there is an interaction term and events that are quite rare for both parts of the interaction. Analyze how this might be possible following hypothesis:

Hypothesis: Salient territorial threats make mass killing more likely to occur only when leaders espouse exclusionary ideologies. Otherwise, territorial threats have no effect on mass killing. (Hong and Kim 2019, 355)

383. Discuss Alter’s rationale for her generalization case studies: “Although there are many European cases one could examine, only three of the eighteen case studies in this book focus on European legal institutions because I want to show that the new terrain of international law exists beyond Europe. Because I prefer less likely cases, I focus on human rights courts from Latin America and Africa, and cases where ICs with economic subject matter jurisdiction end up speaking to human rights issues, rather than a case study involving the European Court of Human Rights.” (Alter 2014, 25)

384. Shannon 2009 is a case study of the US invasion of Panama in 1989. If one were to do a large-N qualitative analysis starting from this article what would be the basic case selection criteria using figure 2. Should you consider case selection further down the tree? Why is that important?

Answer:

The top of the decision tree is those leaders who are motivated to violate the norm. These are the cases where we can see the norm potentially in action.

385. King and Zeng (2007) have presented a large-N argument about counterfactuals. Basically the article deals with the perils of extrapolation and a technique
for deciding when a counterfactual is an extrapolation outside the data. Basically, counterfactuals within the data are more reliable than extrapolations outside it: a bit of conventional wisdom. One conclusion from this methodology is that one should collect more data to increase the region covered by the data; a conclusion not too surprising from a statistical methods point of view. Use the Possibility Principle cube to think about these issues.

(1) It is quite possible for counterfactuals within the boundaries of the data to be quite distant from actual data points. Illustrate this in the cube.

(2) Use the cube to show how counterfactuals outside the data might be closer to actual data points than points in the data cube.

(3) Use the democracy and war literature to illustrate how interpolation might be more problematic than extrapolation.

(4) Applying the Possibility Principle principle usually means reducing the size of the sample space, while King and Zeng implicitly argue for increasing it. However, both argue that causal or counterfactuals are best made with contrasting cases that are “close” to each other. Discuss.

Answer:

For (1) and (2) put points at the corners, and a few outside the corners and leave the middle of the cube empty.

The democratic peace literature is paralleled by a small “autocratic” peace literature. But there is a fair amount of evidence that hybrid or transitional regimes are more war prone. Hence interpolation of causal effects may not work in this case. Also, extrapolation to higher quality democracy is probably going to be more valid than interpolation to gray zone democracies.

For (4) the key issue hidden by the King and Zeng article is that one wants to make contrasts with cases that are nearby. Data in the convex hull may be more distant than data outside it. The Possibility Principle says that increasing the convex hull may be counter-productive if the contrasting cases are too far away.

386. Fearon and Laitin in their oft-cited work (2003) on civil war (in)famously include a variable on amount of mountainous terrain (see Collier and Hoeffler 2004). This is controversial in terms of causal inference because it essentially does not vary within countries. Influential causal inference scholars deny causal status to variables which cannot be manipulated. Discuss how one use large-N qualitative case studies to explore the causal mechanism linking mountainous terrain to civil war. Can this illustrate how a constant can be part of a causal mechanism?

387. Large-N qualitative testing works when there are few cases in the (1,1) cell. One generic situation where this can occur (illustrated by the Mansfield and Synder example in Goertz 2017) is an interaction term where already one or both of the terms is relatively rare. For international conflict scholars this can occur when joint democracy is interacted with virtually any other variable. Discuss this, for example, in Ghatak et al. 2017 who interact joint democracy with territorial conflict.
388. Fearon and Laitin in their oft-cited work (2003) on civil war (in)famously include a variable on amount of mountainous terrain (see Collier and Hoeffler 2004). This is controversial in terms of causal inference because it essentially does not vary within countries. Influential causal inference scholars deny causal status to variables which cannot be manipulated. Discuss how one use large-N qualitative case studies to explore the causal mechanism linking mountainous terrain to civil war. Can this illustrate how a constant can be part of a causal mechanism?

389. The medium-N paradigm means in this context exploring cases in border areas of the (1,1) zone as secondary case studies. For the democratic peace this would be borderline democracies or borderline peace (in the joint democracy is sufficient for peace version of the democratic peace). An interesting exercise to see what the cases scholars have used when critiquing of the democratic peace, e.g., Elman (1997) or Rosato (2003).

390. Large-N qualitative testing. Analyze this example of large-N qualitative testing.

In the replicated version of Schultz’s data set, there are 147 democratic victories—that is, unreciprocated MIDs initiated by democracies. Of these episodes, Table 2 lists the twenty-five most influential cases as measured by the ddfbeta statistic, which reports the change in the coefficient of interest when an individual case is excluded from the regression. These twenty-five cases—roughly 2 percent of the overall data set—exert the greatest downward impact on the democratic initiator variable, in effect doing the most work to confirm the predictions of the democratic credibility hypothesis. The empirical analysis in Democracy and Coercive Diplomacy relied on the MID data set because, according to the study, the data set is composed of “cases in which states used threats of force, regardless of how prominent or how severe the ensuing crisis eventually became.” If true, the data set would indeed be useful for testing the democratic credibility hypothesis, because the theory aims to explain the success and failure of militarized threats. However, this characterization of the MID data set is largely incorrect. Our research indicates that none of the twenty-five most influential democratic victories in our replication of Schultz’s analysis represents an actual threat made by a democracy. In fact, as Table 2 reports, eight of the twenty-five cases appear to involve no militarized dispute at all. The remaining cases entail unilateral raids, skirmishes, or border violations (twelve cases!), troop movements or exercises without a coercive demand (two cases), and encounters with fishing boats or other civilian vessels (three cases!). Because these cases do not involve coercive threats, they do not belong in an empirical test of democratic credibility theory. (Downes and Sechser 2012, 468)

391. Large-N qualitative testing and multimethod research more generally involves looking for evidence of the causal mechanism in cases. Discuss Simmons and Danner (2011, 235) where the number of (1,1) cases is manageable, but they cannot find much qualitative evidence. Does that invalidate the statistical analysis?

392. Copeland in his analysis of causal importance of economic interdependence and future expectations looks all cases of major power crisis and war: “In this
book, I look at the onset of essentially all the significant great power crises and wars from 1790 to 1991” (2015, 76). In the context of large-N qualitative testing, discuss the advantages and disadvantages of this Y-centric approach.

**QCA, set theory, logic, Boolean algebra, fuzzy logic**

NOTE: For calibration exercises see the semantic transformation section above. See also the section “Two-level theories” for many examples involving QCA, INUS, and SUIN.

393. Grzymala-Busse (2007) is a core example in the causal mechanism chapter of *Multimethod research, causal mechanisms, and case studies: an integrated approach*. In my figure 2.4 of her theory “Robust competition” is constituted by three mechanisms (1) moderation, (2) anticipation, and (3) cooptation. Discuss how on might aggregate or combine these causal mechanisms (e.g., OR, AND, addition). Mikkelsen (2017) uses Grzymała-Busse (2007) a core example in his discussion of fuzzy logic case studies. Compare and discuss his version of the mechanism of competition, e.g., Figure 2 and how he uses fuzzy logic to combine the three mechanisms with the discussion in chapter 3 of *Multimethod research, causal mechanisms, and case studies: an integrated approach* of the same mechanism.

394. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when $X_1$ is a subset of $X_2$. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace: “This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies” (Owsiak and Vasquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in $X_1$ is taken up in $X_2$? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if $X_1$ is a perfect subset of $X_2$ and one introduces an interaction term, $X_1 \times X_2$?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as $X_1$ less than or equal to $X_2$ for all observation or vice versa).

395. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.
Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

396. Give a fuzzy logic interpretation of the relationship between the two democracy dimensions in Coppedge’s (2012) analysis.

![Graph showing distribution of countries on two dimensions of democracy](image)

Figure 2.3 Distribution of Countries on Two Dimensions of Democracy, 2000. Both indicators were estimated by exploratory Principal Components Analysis of fourteen democracy indicators. The units of measurement are standard deviations. The estimation procedure and justifications for interpreting these as indicators of Dahl's two dimensions of polyarchy are in Coppedge, Alvarez, and Maldonado (2008).

Answer:
The Coppedge figure would be interpreted in a set theoretic fashion that inclusiveness is necessary for contestation.

397. Sometimes the concept may have no necessary conditions in terms of its definition and structure but there can be set theoretic relationships empirically between dimensions. For example, this can happen when one dimension is empirically a subset of another. Discuss the implications of this in terms of the polity data set and Bueno de Mesquita et al.’s discussion of it.

Answer:
Bueno de Mesquita et al. (2005) provides an example. “For instance, only 26 percent of the times when countries have open executive recruitment have
they also institutionalized multiparty competition. But, if they have multiparty competition, then 97 percent of the time they also have open executive recruitment. Open recruitment appears to be necessary but not sufficient for multiparty competition. . . . Clearly, it is very hard to have multiparty competition without the other dimensions, but it is not difficult to achieve thresholds on other dimensions without having multiparty competition.” (pp. 453-55)

398. Since people in general and social scientists in particular are not taught mathematical logic is not surprising that they have issues regarding questions of logic. Below Chambers describes an extremely famous and classic experiment – that while he discusses it in terms of confirmation bias – is really a question about logic. Give the two correct ways to test this hypotheses. Why is one much less common practice than the other?

Since then, many studies have explored the basis of confirmation bias in a range of laboratory controlled situations. Perhaps the most famous of these is the ingenious Selection Task, which was also developed by Wason in 1968. The Selection Task works like this. Suppose I were to show you four cards on a table, labeled D, B, 3, and 7 (see figure 1.1). I tell you that if the card shows a letter on one side then it will have a number on the other side, and I provide you with a more specific rule (hypothesis) that may be true or false: “If there is a D on one side of any card, then there is a 3 on its other side.” Finally, I ask you to tell me which cards you would need to turn over in order to determine whether this rule is true or false. Leaving an informative card unturned or turning over an uninformative card (i.e., one that doesn’t test the rule) would be considered an incorrect response. Before reading further, take a moment and ask yourself, which cards would you choose and which would you avoid? (Chambers 2017, 5)

Answer:

Here is Chambers:

If you chose D and avoided B then you’re in good company. Both responses are correct and are made by the majority of participants. Selecting D seeks to test the rule by confirming it, whereas avoiding B is correct because the flip side would be uninformative regardless of the outcome. Did you choose 3? Wason found that most participants did, even though 3 should be avoided. This is because if the flip side isn’t a D, we learn nothing – the rule states that cards with D on one side are paired a 3 on the other, not that D is the only letter to be paired with a 3 (drawing such a conclusion would be a logical fallacy known as “affirming the consequent”). And even if the flip side is a D then the outcome would be consistent with the rule but wouldn’t confirm it, for exactly the same reason. Finally, did you choose 7 or avoid it? Interestingly, Wason found that few participants selected 7, even though doing so is correct – in fact, it is just as correct as selecting D. If the flip side to 7 were discovered to be a D then the rule would be categorically disproven – a logical test of what’s known as the “contrapositive.” And herein lies the key result: the fact that most participants correctly select D but fail to select 7 provides evidence that people seek to test rules or hypotheses by confirming them rather than by falsifying them.
It is not surprising that few people thought of the contrapositive test given its nonintuitive nature. The sufficient condition test is straightforward.

399. One often sees scatterplots like the Tilly one in the study of the dimensions of democracy. What would be the QCA interpretation of these? What might that suggest about aggregation?

Answer:
Empirically the triangular scatterplot makes it look like civil liberties are sufficient for political rights. More substantively, they place a floor under the level of political rights.

![Civil liberties and political rights](image)

**Figure 7.1:** Freedom House ratings of European countries on political rights and civil liberties, 2001. *Source:* Compiled from Freedom House 2002.

400. Joshi and Wallensteen (2018) give five dimensions [(1) security provisions, (2) governance, (3) economic reconstruction, (4) Reconciliation and transitional justice, (5) Civil society] for quality peace: “This volume explores five dimensions that both theoretically and empirically necessary for quality peace in a post-acord society” (p. 5). Often scholars give a list of necessary conditions without stating if they are jointly sufficient. Would it be reasonable to also assume that they are jointly sufficient?

401. Discuss the claim below that it is better to look at the superset than the subset.
For the sixty-two sovereign states included in the territorial contenders data set, Cunningham et al. identify ninety-seven rebel groups that controlled territory. The territorial contenders data set includes eighty-one of them (83.5 percent). Of the remaining sixteen rebels identified by Cunningham et al. as territorial, five are included among the ancillary materials of the territorial contender data set as near misses. The other eleven are instances where our coders could not identify any clear evidence of territorial control. Consultation with the source documents for the NonState Actor Dataset reveals no citations establishing territorial control for the rebel groups in question. We believe the codings in the territorial contender data set are valid, while accepting that there may be some occasional error in either our or the non-state actor (NSA) data sets. An 84 percent overlap is very high, nonetheless. Absent an argument specific to a subset, there is considerable evidence in favor of analysis of the entire category of territorial contenders over analysis of any one subset. (Lemke and Crabtree 2019, 19, emphasis is mine)

402. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.

Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

403. Discuss the very strong Boolean claim that they have found a necessary and sufficient condition: “For the 500 Ush denomination game, we find qualitative support for the claim that players exhibit in-group bias if and only if they are observed, although the results do not reach significance at conventional levels (row 3; columns 3 and 4). In the 100 Ush denomination game, however, we find strong evidence that players discriminate in favor of co-ethnics if and only if they can be seen to be doing so (row 3; columns 1 and 2). Taken together, these findings offer strong support for this strategy selection mechanism as an important source of the variation we observe in public goods provision across ethnically homogeneous and heterogeneous settings.” (Habyarimana et al. 2007, 721, emphasis is the authors’)

404. Fuzzy-set analysis is often problematic when there are many observations near the origin (i.e., observations whose values on all variables approach zero). When observations are near the origin it is extremely easy for them to pass tests of sufficiency since $X$ is likely to be less than $Y$. How does the Possibility Principle aid fuzzy-set methodology by dealing with these cases?

Answer:

This geometric interpretation of the Possibility Principle offers a solution to an important problem that arises in fuzzy-set analysis. The problem involves what to do with cases that are near the origin (i.e., cases near the (0,0,0) corner). As Ragin notes (2000, 250–51), when testing whether variables are causally
sufficient for an outcome, observations with a zero for all the independent variables will always satisfy causal sufficiency and thus artificially inflate the number of cases where the theory works (this dilemma is the Raven Paradox). See also the exchange between Osherson and Smith (1981) and Zadeh (1982). The former is a classic in the cognitive psychology literature on categorization and Zadeh is the inventor of fuzzy-set logic.

405. Explain how the cube (Mahoney and Goertz 2004) provides a visual summary of disconfirming observations for necessary and/or sufficient condition hypothesis.

Answer:

With a necessary condition hypothesis, theory disconfirmation occurs when the necessary condition is absent, but the outcome of interest is present. Thus, disconfirming cases are located in the top portion of the cube where only one necessary condition is present (i.e., the (1,0,1) and (0,1,1) corners).

With a sufficient condition hypothesis, theory disconfirmation occurs when the sufficient condition is present, but the outcome of interest is absent. Thus, disconfirming cases are located in the bottom portion of the cube where the two jointly sufficient causes are present (i.e., the (1,1,0) corner).

406. What would be the set of appropriate negative cases for the following Boolean theory?

\[ Y = A + B + C + D \] (3)

Does this contradict the Possibility Principle? Why does this example not fit the general kind of problem that typifies case selection in small-N studies? Hint: Which combinations of A, B, C, and D predict that Y does not occur? (Thanks to Charles Ragin for this exercise.)

Answer:

This is a question dealing with the logic of the Possibility Principle. The problem here is that if we followed the Possibility Principle for selecting negative cases, we would not have any negative cases at all since the Possibility Principle excludes all cases where A–D are absent (i.e., not-A AND not-B AND not-C AND not-D, this is a hint in the question). But this is exactly the set of cases we need to test the Boolean theory \( Y = A + B + C + D \). This situation arises because A–D are individually sufficient for Y. However, models like \( Y = A + B + C + D \) are extremely rare in the qualitative-comparative literature.


Answer:

Negotiated settlement is sufficient condition for no genocide. Military victory is necessary condition for genocide.

408. In the literature on within-case causal inference researchers often talk about hoop tests and smoking gun tests in terms of “certainty” and “uniqueness”: 
“The first dimension is called certitude or certainty and captures how likely it is to confirm a specific observable implication in process tracing. The second dimension is called uniqueness and asks whether an observable implication can be derived from a single or multiple hypotheses” (Rohlfing 2014 table 1). Rohlfing goes on to claim: “A hoop test is characterized by high certainty and no uniqueness.” (Rohlfing 2014, 612) Describe a situation where a hoop test has a high degree of uniqueness.

Answer:
As the hoop test gets closer to being a sufficient condition, its uniqueness increases, i.e., when the hoop is very small.

409. Should one assume an interaction term or just a straight additive model, or a Boolean model for the data in the table below? Table 4 is a stripped-down version of their table 2 (p. 88) which gives the two core independent variables: (1) federal versus centralized state and (2) multiparty system versus two-party system. I use their primary dependent variable “macrocorporatism” (the other one they discuss is “sector coordination”). There is no statistical analysis in this article; the empirical analysis is carried out via four case studies, one from each cell in table 4.

<table>
<thead>
<tr>
<th>Number of Parties</th>
<th>Level of State Centralization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Centralized</td>
</tr>
<tr>
<td>Multiparty</td>
<td>Denmark</td>
</tr>
<tr>
<td></td>
<td>macrocorporatism=.72</td>
</tr>
<tr>
<td>Two party</td>
<td>Britain</td>
</tr>
<tr>
<td></td>
<td>macrocorporatism=−.14</td>
</tr>
</tbody>
</table>

Source: Martin and Swank 2011, table 2.

Answer:
As in almost all two-variable models – i.e., $X_1 \text{ AND } X_2 \rightarrow Y$ – the $(1,1,1)$ cell is clear as is the $(0,0,0)$ cell. The critical feature that will influence the decision between a QCA interpretation against an additive linear one is what happens with the cases where one independent variable is present and the other is absent, i.e., $(1,0)$ or $(0,1)$.

In set theoretic models the value of the $(1,0)$ and $(0,1)$ cells would be close to that of $(0,0)$ because of the necessary condition relationship. The additive model suggests that when one independent variable is present and the other absent we should see intermediate results, somewhere between the extreme values on the dependent variable. Hence, a signal that the additive relationship is at work is when the dependent variable is significantly greater than the $(0,0)$ but yet significantly less than the $(1,1)$ dependent variable value. This is in fact what we find in table 4. The two off-diagonal cells have values in between those in the on-diagonal cells.
It is pretty clear that a set theoretic interpretation of these data would not be appropriate because of the intermediate values of the dependent variable when only one independent variable is present. However, it is very difficult to make any strong conclusions regarding whether there should be an interaction term.

410. Discuss Pahre’s claim that necessary conditions violate common statistical assumptions: “Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that ‘if two units have the same value of the key explanatory variable, the expected value of the dependent variable will be the same’ (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition x ← y; in other words, x may occur with or without y” (Pahre 2005, 131).

411. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when \( X_1 \) is a subset of \( X_2 \). For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace: “This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies” (Owsiak and Vasquez 2016, 339). What is the causal interpretation of the subset variable when all of the variation in \( X_1 \) is taken up in \( X_2 \)? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if \( X_1 \) is a perfect subset of \( X_2 \) and one introduces an interaction term, \( X_1 \ast X_2 \)?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as \( X_1 \) less than or equal to \( X_2 \) for all observation or vice versa).

412. Walter’s oft-cited book on the peaceful settlement of civil wars makes much use of necessary or sufficient conditions both in hypotheses and the data analysis. Discuss how she mixes the two. For example, “three sections follow. The first section tests whether third-party security guarantees and power-sharing pacts are necessary in combination to bring a peaceful solution to war. The data reveal that civil wars are significantly more likely to end in a successfully implemented settlement if both types of guarantees are present” (Walter 2003, 92).

413. Discuss the “light-switch” analogy as a way to think about process tracing and equifinality. The basic analogy is that the light-switch is on and the light is on (essentially the (1,1) cases). Discuss process tracing as following electricity through various junction boxes to the light bulb. Equifinality lies in the fact
that there might be multiple switches that can possibly control the same light. Discuss what would be the analogy in case study work of “turning on or off” the switch (within cases or between cases). What about counterfactuals? (Thanks to Mike Desch for suggesting this analogy)

414. Copeland (2015) has at the core of his book an interaction; positive levels of economic interdependence (often trade) interacted with expectations about future economic conditions. Discuss the situation where one of the interaction terms is basically a subset of the other. For example, what case studies we be appropriate? Is the superset variable (positive economic independence) more like a scope condition?

To determine whether the liberal prediction or realist prediction will prevail, we must introduce an additional causal variable—namely, a state’s expectations of the future trade and investment environment. When a dependent state has positive expectations about this future environment, it is more likely to see all the benefits of continuing the current peace and all the opportunity costs of turning to war. Economic interdependence would then be a force for peace. Yet if a dependent state has negative expectations about the future economic environment—seeing itself being cut off from access to foreign trade and investment, or believing that other states will soon cut it off—then the realist logic will kick in. Such a state will tend to believe that without access to the vital raw materials, investments, and export markets needed for its economic health, its economy will start to fall relative to other less vulnerable actors. If this economic decline is anticipated to be severe, the leaders of the dependent state will begin to view war as the rational lesser of two evils—that is, as better than allowing their state to fall to a point where rising states can attack it later or coerce it into submission. (Copeland 2015, 2; basic hypothesis of the book)

415. Levy and Thompson (2005) discuss the perhaps most classic hypotheses in international security: “threats of hegemony generate great-power balancing coalitions.” They define “balancing coalitions” as an forming an alliance. The sufficient condition hypothesis: “If threat of hegemony then great powers form balancing coalitions.” Make the data in the table below such that they would pass the QCA criterion for sufficient condition hypotheses. Then change the data so that the data (1) are a trivial sufficient condition hypothesis, (2) nontrivial sufficient condition hypothesis, and (3) nontrivial but with the relationship in the wrong direction.
Table 2
CAPABILITY CONCENTRATION AND ALLIANCE RESPONSE AGAINST THE LEADING POWER, 1495–1999

<table>
<thead>
<tr>
<th>Weak relative position (less than 33%)</th>
<th>Stronger relative position (33% or greater)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No alliance response</td>
<td>151 (.702)</td>
</tr>
<tr>
<td>Alliance response</td>
<td>64 (.298)</td>
</tr>
<tr>
<td>N = 445</td>
<td>215 (1.000)</td>
</tr>
<tr>
<td></td>
<td>230 (1.000)</td>
</tr>
</tbody>
</table>

\( \chi^2 = 28.42; p = .000 \)

Note: The years in which the state exceeded 33 percent of capability shares are indicated in parentheses in the “leading land power” column.

Answer:
A percentage of 80 percent or higher in the alliance formation column, would pass the QCA bar which is usually around 70–80 percent. For a trivial sufficient condition, make the percentage in the left-column the same as in the right. For nontrivial making it significantly lower, e.g., 50 percent. For nontrivial in the wrong direction make the percentage in the left-hand column around 95 percent.

416. Clarke gives an excellent discussion of the logic and methodology, in particular the common logical error of affirming the antecedent. Explain how the Avoid Overdetermination Guideline tries to deal with this by eliminating “other mortal creatures.”

The invalid argument is a logical fallacy known as “affirming the consequent.” Consider the following invalid deduction: If x is human, then x is mortal [therefore] If x is mortal then x is human. Obviously, not all mortal creatures are humans. Despite the fact that x is mortal, we cannot conclude that x is human. . . . The problem with affirming the consequent is that there are numerous conditions that imply that a thing is mortal, and being human is simply one of them. In the example given above, for instance, x could be a cat, a dog, or even a dandelion, and observing that “x is mortal” does not give us warrant to conclude that any particular one is true. (Clarke 2007, 888)

Answer:
The Avoid Overdetermination Guideline selects cases where alternative “mortal” variables are zero. If one had a complete list of mortal beings and applied the Avoid Overdetermination Guideline (Z = 0) then the only possible case selection options left are humans.

417. One reason why qualitative methods (based on logic and set theory) are different from statistical methods is because most statistics use linear algebra while set theory is based on Boolean algebra. One difference between Boolean and linear algebra is that Boolean algebra requires a second distributive law: 
\[ a + (b * c) = (a + b) * (a + c). \]  
Explain why this does not hold for linear

Answer:

The second distributive law can be neatly visualized for sets with two Venn diagrams of three overlapping circles each. The law says . The left hand-side of the equation marks the whole of “” plus the intersection of “” and “,” and the right hand-side will mark exactly the same region.

Since students in the social sciences are not taught basic mathematical logic it is easy to make errors. Explain the logical error in the following descriptions of the democratic peace:

Immanuel Kant’s argument that democratic institutions . . . are a necessary condition for peace has been empirically substantiated. (Risse-Kappen 1996, 366)

The recent flurry of studies of the theory of the “democratic peace” follows upon Kant’s argument that a necessary condition for peace between states is constitutional republics. (Holsti 1996, 180)

Answer:

The authors have confused a necessary condition with a sufficient condition. The democratic peace says that joint democracy is sufficient for peace.

It is not too hard to make mistakes when manipulating necessary and sufficient conditions. Schweller (1992) manages with the hypotheses below to be both redundant and contradictory. Explain.

1. A power transition involving a declining democratic leader is both a necessary and sufficient condition for the absence of preventive war. 1a. When a declining democratic leader confronts a rising democratic challenger, accommodation results. 1b. When a declining democratic leader confronts a rising nondemocratic challenger, the leader tries to forma a defensive alliance system to counterbalance the threat. 2. A power transition involving a declining nondemocratic state is a necessary but not sufficient condition for a preventive war, regardless of the regime type of the challenger. (Schweller 1992, 248–49)

Answer:

Hypothesis 1 stipulates that, given a power transition and a declining leader, (a) the leader’s democracy is necessary for the absence of war and (b) the leader’s democracy is sufficient for the absence of war. Hypothesis 2 stipulates that the leader’s nondemocracy is necessary for war. Since there are only two kinds of leaders, democratic and nondemocratic, and since necessary and sufficient conditions are fungible via the contrapositive operation, stating that the leader’s democracy is sufficient for the absence of war is logically equivalent to stating that the leader’s nondemocracy is necessary for the presence of war: if \[ P(Y|X) = 0, \text{ then } P(\neg Y|\neg X) = 1. \] Therefore, the assertion in Hypothesis 2 that nondemocracy is necessary for war is already implied by Hypothesis 1. On the other hand, Hypothesis 2 specifically states that, given a
power transition and a declining leader, the leader’s nondemocracy is not sufficient for preventive war. Translating this statement into necessary-condition terms yields the statement, “The leader’s democracy is not necessary for the absence of preventive war.” Hypothesis 1 makes the claim that democracy is necessary for the absence of war. Hypothesis 2, therefore, merely restates one half of Hypothesis 1 and contradicts the other.

Express Przeworski et al.’s (2000) main conclusion regarding GDP/capita and democracy using the logic of necessary and sufficient conditions with (1) two independent variables, (2) as a sufficient condition, (3) with time coefficients, and (4) dependent variable involves authoritarianism.

Answer:
Minimum level of GNP/capita (time $t$) AND Democracy (time $t$) are jointly sufficient for no transition to authoritarianism (from time $t$ onwards).

Notice how this orients theory toward to question of transitions to nondemocracy instead of transitions to democracy. What theory is there about transitions to authoritarianism?

Researchers often use set theoretic or necessary (or sufficient) condition language. For example, “all sociologists are Democrats” means sociologists are a subset of all democrats; which also means being a Democrat is a necessary condition for being a sociologist. Such language can be common but it is often not recognized as such. Take Lipset’s 1993 American Sociological Association Presidential address and find all examples of the use of logic (e.g., necessary condition) or set theory in hypotheses or discussions of empirical results (e.g., all sociologists are Democrats). Express these descriptive statements in terms of hypotheses about necessary or sufficient conditions.

Answer:
There are quite a few. For example, (1) “every country with a population of at least 1 million that has emerged from colonial rule and has had a continuous democratic experience is a former British colony” (Lipset 1993, 5, citing Weiner 1987), (2) “there has been no case of political democracy that has not been a market economy” (Lipset 1993, 5, citing Berger 1992). Others can easily be found.

Here is how the logic works using a descriptive statement from Dahl (1971): “all highest-level [developed] countries are polyarchies” (Cited by Diamond 1992, 97).

All highly developed (Wealth) countries are polyarchies (Democracy). (1)
All $W$ are $D$. (2)
$D$ is necessary for $W$. (3)
$\neg D$ is sufficient for $\neg W$ (“$\neg$” means “not”). (4)

Thus (3) or (4) are equivalent logical statements the express Dahl’s empirical findings.

The “democratic peace” idea is that democracies never fight wars with each other. This can be expressed as a either a necessary condition hypothesis or a sufficient condition hypothesis. Give both hypotheses.
It is not uncommon to confuse necessary condition counterfactuals with sufficient condition ones. See Fenoaltea (1973) for a beautiful discussion of this in the context of economic history. Fischer’s classic book (1970) also deals with this. Discuss Fisher’s examples.

One common form of the reductive fallacy is the confusion of necessary with sufficient cause – the confusion of a causal component without which an effect will not occur, with all the other causal components which are required to make it occur. This sort of error appears in causal explanations which are constructed like a single chain and stretched taut across a vast chasm of complexity. The classic example is the legendary battle that was lost for the want of a horseshoe nail; for the want of a nail the shoe was lost, for the want of a shoe the horse was lost, for the want of a horse the rider was lost, for the want of a rider the message was lost, for the want of the message the regiment was lost, and for the want of a regiment the battle was lost. (Fischer 1970, 172; he then goes on to use the classic example of the battle of Antietam for which the accident loss of confederate orders to the North was a key factor in the North’s victory)
References


EU Social Protection Committee. 2012. Social protection performance monitor (sppm) methodological report by the indicators sub-group of the social protection committee.


169


