


# Concepts and Measurement in Multimethod Research

Political Research Quarterly  
66(2) 280–291  
© 2011 University of Utah  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/1065912911427453  
prq.sagepub.com  


Ariel I. Ahram<sup>1</sup>

## Abstract

This article argues that concept misformation and conceptual stretching undermine efforts to combine qualitative and quantitative methods in multimethod research (MMR). Two related problems result from the mismatch of qualitatively and quantitatively construed concepts. Mechanism muddling occurs when differences in the connotation of qualitatively and quantitatively construed concepts embed different causal properties into conceptual definitions. Conceptual slippage occurs when qualitatively and quantitatively construed concepts use incompatible nominal, ordinal, or radial scales. Instead of gaining leverage from the synthesis of large- and small-*N* analysis, these problems can push MMR in two diametrically opposed directions, emphasizing one methodological facet at the cost of the other.

## Keywords

multi-method research, concepts, measurement

In the classic “Concept Misformation in Comparative Politics,” Sartori (1970) argues that in the course of seeking to describe the entire globe, core terms in the social scientific vocabulary, like democracy or development, are altered, diluted, or rendered vacuous. Instead of gaining analytical leverage through more powerful techniques of comparison, social science is likely only to stumble into a morass of terminology so general as to be vague in applicability and meaning. Sartori’s warning were often invoked during the so-called “paradigm wars” between qualitative and quantitative researchers in social sciences in the 1970s and 1980s (Bryman 2006; Teddlie and Tashakkori 2009, 14–16; Guba and Lincoln 2005). Fortunately, many of these now stale debates have given way to a more pluralistic vision of inter- and intramethodological dialogue. Interpretivism, which favors more reflexive and ideographic case studies, now stands alongside nomothetically oriented qualitative and quantitative work (Sil 2000; Yanow and Schwartz-Shea 2006). For those aspiring to develop generalizable theory, partnership between qualitative and quantitative methods is a new norm and multimethod research (MMR) an increasingly prominent technique.

This article argues, though, that in combining qualitative and quantitative methods, MMR often overlooks important questions about concepts and measurement. Different forms of concept misformation undermine MMR’s two-pronged effort to increase the comprehensiveness and validity of causal accounts in theory building and theory testing. Two related but distinct problems are

identified: The first problem is *mechanism muddling*. Differences between definitions used in qualitative and quantitative versions of concepts embed different causal properties into those definitions. This problem tends to arise as qualitative techniques are used to identify specific attributes of cases that are necessary for the enactment of specific mechanisms or causal pathways. These attributes are not, however, necessarily present in every case identified in the quantitatively construed domain. The result is that even though a consistent correlational pattern is identified, the mechanisms connecting antecedent with outcome remain underspecified and relate to only a subset of the larger domain of relevant cases. The second problem is *conceptual slippage*. This arises when different taxonomical schemas are used to organize what purports to be the same concept, typically as qualitative methods array cases in nominal categories while quantitative methods array them along ordinal or radial axes. This leads to analytical ambiguity, as some cases are categorized as equivalent in the qualitative sphere but different in the quantitative sphere and vice versa, making it extremely difficult to generate valid generalization. Instead of leading to the synthesis of qualitative and quantitative methods in MMR, attempts to mitigate these problems push MMR in two

<sup>1</sup>University of Oklahoma, Norman, OK, USA

## Corresponding Author:

Ariel I. Ahram, College of International Studies, University of Oklahoma, Hester Hall 129, Norman, OK 73019.  
Email: arielahram@ou.edu

opposing directions. On one hand, addressing conceptual slippage requires standardization of concepts across the quantitative and qualitative domain. On the other hand, addressing mechanism muddling requires greater conceptual flexibility to capture the variety of equifinite mechanisms leading to the same ultimate result.

To illustrate the hazards of concept misformation and suggest some solutions, the article examines three recent books that deploy MMR designs, Schultz's (2001) *Democracy and Coercive Diplomacy*, Lieberman's (2003) *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*, and Lange's (2009) *Lineages of Despotism and Development: British Colonialism and State Power*. These works break significant ground methodologically and substantively, and each has garnered considerable scholarly recognition. The goal is not to impugn them but rather to argue *a fortiori* that if they overlook problems stemming from conceptual misformation, then these challenges are likely common to the venture of MMR as a whole.

### Concept Misformation and the Commensurability Problem in MMR

Sartori raised the problem of concept misformation in the midst of a veritable revolution in social science methodology, with the widespread adoption of quantitative, statistical techniques and new commitment to positivism and the elucidation of universal laws of social science. Sartori warned that the idiom of quantitative, statistical metrics was often vacuous and thus ill suited as a means to development and to express such theories. Using the metaphor of the "ladder of abstraction," Sartori highlighted the interplay of two dimensions in concept formation: intension (connotation), the systematic and explicit definition of the characteristics of the concept, and extension (denotation), the range of cases that can be categorized as meeting the conceptual definition. Intension and extension are inversely related. On one hand, moving up the ladder of abstraction by increasing a concept's extension incorporates more cases and strips away some of its necessary intension or its specific conceptual attributes. On the other hand, moving down the ladder of abstraction by adding to a concept's intension means that it will describe fewer empirical cases because it adds to the criteria necessary to be judged as an instance of the linguistic term under consideration. Sartori used the terms *stretching* and *straining* to describe the process by which specific connotation is jettisoned in the course of extending denotation. When strained or stretched, conceptual terms cover more by saying less about the objects to which they refer (Sartori 1970, 1041).

Connotation and denotation are alternate sides of the same coin. While Collier and Goertz point out some important exceptions to Sartori's claim about the inverse relationship of connotation and denotation (most notably in the case of family-resemblance-type as opposed to necessary/sufficient-type concepts), the core of Sartori's contention about concepts remains intact.<sup>1</sup> An argument about whether or not to include a specific case in an analysis relates ultimately to the conceptual definition used to delineate criteria for inclusion under an abstract rubric. Disputes over selecting cases and establishing a referent domain necessarily imply ambiguity in the key terms of the concepts. Conversely, vagueness in the key terms implicates an imprecise delineating of referent cases (Collier and Mahoney 1996). This goes to Sartori's basic concern that concepts like democracy and political community might have their meanings distorted as they travel to cover new cases.

Sartori never suggested that qualitative and quantitative methods are fundamentally incommensurate at an epistemological level. On the contrary, he pointed out that they can share a considerable amount of common ground, including the aspiration for broader theories of social change (at least as long as they both held to a positivist epistemology). The problem Sartori touched on is one of concept commensurability. Changes in the criteria relevant to taxonomical categorization of cases rendered different meanings to specific terms or concepts, making their arguments incommensurate on a local scale. As Sankey (1991, 2000) elaborates, one of the key indicators of such a concept as incommensurability is that referents included in one categorization scheme are excluded from another. Though they might share a term, they are talking about different objects because meanings of specific conceptual vocabularies not only vary but also are impossible to translate meaningfully.<sup>2</sup>

A number of scholars have recently sought to follow Sartori's footsteps in offering a unified epistemological framework for qualitative and quantitative methodology. King, Keohane, and Verba (1994) set out to demonstrate a common toolkit for both methodologies. Collier and Brady (2004) reiterate the notion of shared standards even while they highlight the diverse practices. Techniques of "nested analysis"—sequential application of quantitative and qualitative methods on the same research puzzle—seem to offer the best of both methods and are gaining increasing notoriety in political science (Coppedge 2005; Lieberman 2005; Fearon and Laitin 2008).

Nested analysis offers two primary and related justifications for combining methods. The first is the idea that qualitative and quantitative methods offer complementary but distinctive forms of analysis that combine to offer a more comprehensive, multidimensional

causal account (Gerring, 2005, 2011; Mahoney 2008). As Mahoney and Goertz (2006, 231) observe,

an explanation of an outcome in one or a small number of cases leads one to wonder if the same factors are at work when a broader understanding of scope is adopted, stimulating a larger-N analysis in which the goal is less to explain particular cases and more to estimate average effects. Likewise, when the statistical results about the effects of causes are reported, it seems natural to ask if these results make sense in terms of the history of individual cases, one wishes to try to locate the effects in specific cases.

This distinction leads to a division of labor between different methodological components. The quantitative, statistical portion identifies regular macro-level correlative relationships. The qualitative, case study portion is used to infer what mechanistic processes connect antecedent (X) and outcome (Y) variables (Pawson 1995; Erzberger and Prein 1997; Bryman 2004). Importantly, this perspective does not see mechanisms as mere handmaidens of correlation or yet-to-be-discovered intervening variables but analytically distinct that entities that generate change directly (Waldner 2007, 154; Mahoney 2001).

The second justification for MMR is to leverage such a multidimensional account to increase validity. This is one of the oldest justifications for MMR. Put simply, a hypothesis that survives a series of tests with different methods is more valid than a hypothesis tested with the help of only a single method (Erzberger and Kelle 2003, 460; Denzin 1978; for a critique, see Reiss 2009). Failing the qualitative test prompts revision to the quantitative model and vice versa. Nesting or moving sequentially between qualitative and quantitative methods allows a check for spurious correlation and interrogate potential omitted variables qualitatively before integrating these new variables into the quantitative analysis for testing across a larger sample.

Although the techniques of nested analysis have been subject to methodological critique (Rohlfing 2007) and continued skepticism about the epistemological continuity between the two methods (Ahmed and Sil 2009; Chatterjee 2009), few have grappled with the question of conceptual commensurability *per se*. Coppedge (2009, 16), for one, cautions that “there is a flaw in any multi-method work that relies on mismatched concepts—which probably includes most multimethod research.” This is a significant predicament. If qualitative and quantitative methods are talking about different things, then it is impossible for them to be saying the same things (Kelle

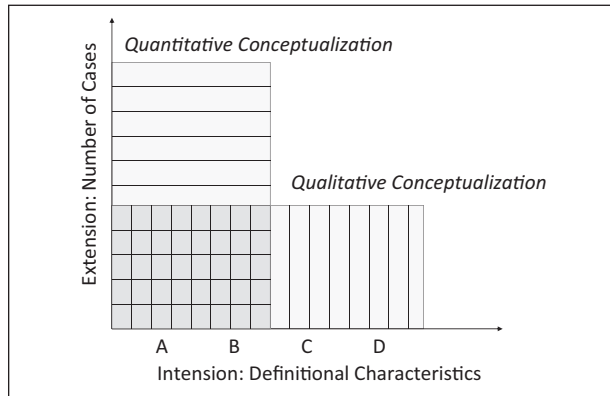
2001). Seemingly concurrent findings can be dismissed as mere felicitous coincidence. Still, the question remains, how can the cases of conceptual mismatches be diagnosed and addressed? The sections below describe two distinct forms of concept misformation in MMR and analyze the implications of each for goals of confirming and complementary causal accounts.

### From Conceptual Stretching to Mechanism Muddling

The possibility of conceptual stretching because of a mismatch between the concepts and variables used in the qualitative and quantitative setting has obvious pertinence for MMR. The concepts used in case-based, qualitative research tend to have complex or “thick” definitions developed iteratively through examination of a small number of cases. In contrast, the concepts used in variable-oriented, quantitative research tend to be “thin,” with relatively simple conceptual definition. Conceptual thickness and thinness translate inversely into narrowness or breadth at the level of extension. Because of their definitional intricacy and high intension, qualitative concepts are designed to apply to only a small number of cases (Coppedge 1999). In contrast, the simpler definitions and low intension involved in quantitative methods are amenable to incorporating a much wider universe of cases for interrogation via statistical analysis. Qualitative methods tend to work with concepts that are closer to the maximal or ideal type definition and thus have few cases that fall within their domain, while quantitative methods work with minimal definition and thus have more expansive domains (Gerring and Barresi 2003).

The concepts used in the qualitative and quantitative components of MMR may share a label or term, but they have different characteristics and refer to different categories of cases. This is depicted in Figure 1, where the horizontal axis lists four characteristics necessary for falling under the concepts rubric (A, B, C, and D) and the vertical axis lists the range of empirical cases that fall under it. In the figure, conceptual stretching is most obviously manifest at the level of extension: numerous cases that would be counted as “in” in a quantitative setting but “out” in a qualitative setting.

This form of conceptual stretching has severe implications for the ability to offer complementary accounts of correlation and mechanism. In some cases, the difference between qualitative and quantitative conceptualization may be so stark that the area of overlap is minimal or nonexistent. The terms themselves may be mere homonyms. The qualitative and quantitative definitions of the concept are so disparate that they refer to entirely different cases possessing entirely different attributes. Short of

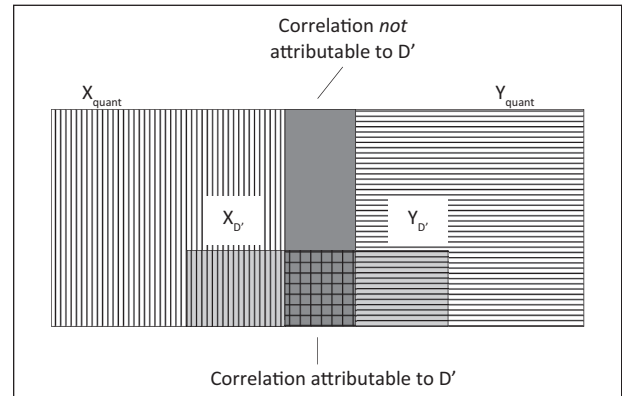


**Figure 1.** Overlay of qualitative and quantitative conceptualizations.

such an egregious malpractice, though, it is still important to disambiguate between qualitative and quantitative definitions of what purports to be the same concept. In fact, Gerring (2001, 65-85) goes so far as to suggest that we consider treating them as two distinctive concepts altogether. If they were expressed as a set, we could imagine qualitative and quantitatively construed concepts referring to two separate domains,  $X_{qual}$  and  $X_{quant}$ , respectively. These domains might overlap to a greater or lesser extent, but they are not congruent.

The situation becomes even more complex when we consider the difference in the kinds of techniques of concept formation commonly practiced in qualitative and quantitative methods. As Goertz notes, though both qualitative and quantitative approaches to conceptualization can embed specific etiological properties into the definition of the concept itself, this is much more common in qualitative case study analysis. Identifying these properties is crucial to explicating the mechanisms by which causes produce their effects in particular instances (Goertz and Mazur 2008, 34; Goertz 2006, 64). In a narrower (quantitative) conceptualization, however, these key attributes may not be included in the connotation, meaning many cases counted as “in” in the quantitative case domain do not share this attribute. A qualitative definition of a concept specifies four attributes: A, B, C, and D'. Attribute D' is vital because it is the locus of causal potential. The quantitative definition, however, includes only attributes A and B.

Incongruence between conceptual definitions poses a significant problem for the division of labor in MMR. The quantitative portion identifies a regular correlational pattern between antecedent  $X_{quant}$  and the outcome, Y. Consideration of the specific causal mechanisms that are responsible for this correlation is left to the qualitative case study. If, however, the qualitative method defines its key concepts to include a



**Figure 2.** Mechanism and correlation in a multimethod explanation.

causal property (D') absent in some of the cases in the quantitative domain, then any claims to have identified both a correlational and causal pattern are bound to be underspecified. The qualitative component can make a claim about the mechanisms involved only in the domain of  $X_{D'}$ , not the larger  $X_{quant}$ . Stated differently, a significant portion of the  $R^2$  observed in the statistical portion must have been the result of mechanisms *other* than the ones identified in the case study. It is even possible that the observed correlation is for the most part spurious and the only real causal relationship exists for the small subset of cases contained in  $X_{D'}$ . This situation is shown graphically in Figure 2, where the covariation between  $X_{quant}$  and  $Y_{quant}$  is much larger than the covariation between  $X_{D'}$  and  $Y_{D'}$ .

A good example of the problems of conceptual stretching and mechanism muddling comes from Lieberman's (2003) *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. The puzzle Lieberman addresses is why states adopt more progressive redistributive tax regimes. The key variable, Lieberman argues, is a state's national political community (NPC). NPC is a novel concept, which Lieberman defines as the official, state-sponsored definition of the nation specified in the constitution and other founding documents. NPC constrains the ability of political entrepreneurs to form coalitions to demand redistribution. As Lieberman describes the mechanism,

The specification of group rights in the form of official state documents and policies provides a stronger set of incentives for political entrepreneurs to make claims based on such identities. . . . Federalism, for example, tends to give important political salience to regional identities, and official racial exclusion tends to give much more salience to racial identities. (Lieberman 2003, 14)

When societies are divided by significant racial and regional heterogeneity, an officially sanctioned racial exclusion or federalist structure channels collective action into certain racial or regionalist forms.

Lieberman traces a surprising but persuasive historical account of this process using a paired comparison of Brazil and South Africa, two states with significant racial and regional disparities but different legal definitions of NPC. Brazil's 1891 federalist constitution privileged claims based on regional equity but was explicitly inclusive on racial grounds. Faced with a state that seemed intent on equalizing racial disparities, Brazil's white economic elite successfully worked to block demands for greater redistribution. In contrast, South Africa's 1909 Constitution specified white racial supremacy while denying recognition to regional differences. By establishing whites as a formal legal category, this cornerstone legal document encouraged the white economic elite to cooperate with the state in establishing a social safety net system that raised the living standards of their poorer coethnics and increased the solidarity of the ruling white minority. After apartheid's downfall, this redistributive regime was opened to all races, turning a tool of racial exclusion into one of socioeconomic equalization.

Following the case studies, Lieberman deploys large-*N* regression analyses to test whether the similar legal definitions of NPC along racial and regional dimensions have the same effect in other cases. Examining constitutions and other legal documents from other cases, he converts the data into a series of dummy variables that he enters on the left-hand side of the equation. The statistical results show a correlation consistent with the small-*N* study. When constitutions and other founding documents enshrine federalism, states tend to have less capacity for redistribution; when they enshrine racial supremacy, states have higher redistributive capacity.

But NPC, the crucial independent variable, is stretched severely in this effort to shift from a thick to a thin conceptualization, muddling the claim to have linked correlation with underlying causal mechanisms. The problems begin with the attempt to convert ontological categories of NPC into discrete dummy variables. The qualitatively derived elaboration of NPC proves too narrow to incorporate the majority of empirically relevant cases. As shown in Tables 1 and 2, South Africa and Brazil are representative of only six of the sixty-nine cases (8.5 percent) included in the quantitatively construed population. Only twenty-one countries (30.4 percent) in the quantitatively construed population share South Africa and Brazil's racial fragmentation. In the population, forty-three cases (62.3 percent) have neither relevant racial or regional cleavage. They essentially are outside the initial definition of NPC used in the qualitative narrative. In the quantitative analysis, Lieberman incorporates these cases

**Table 1.** Initial Typology of National Political Community by Type in Qualitative Analysis.

	Federal	Unitary
Race exclusionary		South Africa
Race inclusionary	Brazil	

Source: Lieberman (2003, 79).

**Table 2.** Distribution of Cases in Quantitative Analysis.

		Federal	Unitary
Fragmented	Race exclusionary	3	4
	Race inclusionary	2	10
Not fragmented		7	43

Adapted from Lieberman (2003, 242)

by inventing a new category ("non-fragmented"), where racial divisions are not present. In so doing, however, Lieberman diminishes his ability to account for the mechanisms that connect antecedent to outcome in these cases. Based on historical investigation of Brazil and South Africa, the qualitatively derived version of NPC makes no mention of what constrains collective action in the absence of significant racial and regional cleavages. Overall, the study is mute about the disposition of the majority of the case it claims to explain.

While Lieberman's resort to the residual category of "non-fragmented" to describe the NPC of the majority of this cases might be blamed on the relative novelty of the concept itself, the same problems of conceptual stretching and mechanism muddling also manifest in projects using relatively well-known and well-defined variables as well. In *Democracy and Coercive Diplomacy*, for instance, Schultz (2001) examines the impact of democracy on crisis management and conflict resolution in light of democratic peace theory. The study begins by using POLITY scores to define a domain of some fifty-eight cases in which democracies faced challenges of deterrence and coding whether or not the opposition stood with the government during the crisis and whether the deterrence was a success. Through chi-squared and probit analyses, he demonstrates that success or failure is closely correlated with whether or not the democratic opposition sides with the government during a crisis.

Based on these statistical findings, Schultz uses case studies of the British in the Fashoda Crisis, the British in the Boer Wars, the British and French in the Rhineland Crisis, the British and French in the Suez Crisis, and the British during Rhodesian Independence to trace the way opposition behavior tips the government's hand and

reveals information to foreign rivals about the level of determination in the crisis. He concludes that

whereas nondemocratic governments have substantial leeway to bluff and probe, democratic governments are less willing to make threats they do not intend to carry out. Underlying this probabilistic prediction is a specific causal mechanism: democratic governments face domestic competitors who have an incentive to oppose the use of force when political and military conditions are unfavorable. (Schultz 2001, 233)

When the opposition supports the government, it lends credence to the national leadership's claims of resolve in a crisis and chastens potential challengers. When the opposition stands apart from the government, it weakens the government's credibility in a crisis.

Similar to the muddling mechanisms in Lieberman's conceptualization of NPC, Schultz's definition of democracy becomes connotatively ambiguous as it travels between small and large-*N*. Transparency in foreign policy decision making—a key characteristic that the qualitative case studies identify as necessary to trigger the signaling mechanisms—is not a component of the POLITY-based quantitative definition of a democracy. Great Britain's parliamentary system is renowned for vigorous and public debate on all manner of policy decisions. But Great Britain was a party in only eighteen of the fifty-eight cases (31 percent) of democratic deterrence in the quantitative sample. A number of other factors can affect how a democratic opposition can meaningfully signal to rival states during crisis, including, to name just a few, democracies headed by presidents rather than parliaments, different institutions of civil–military interactions, and different constellations of opposition forces at the time of crisis. In the United States, for instance, the concentration of foreign policy decision making in a relatively opaque and autonomous executive branch excludes the opposition from foreign policy debates and renders moot most signals that emanate from the legislative opposition. Democracies may be even *less* transparent in their foreign policy deliberations than authoritarian regimes.<sup>3</sup> Even if Schultz's claim about the mechanism is correct for Great Britain, it could not account for many of the other instances of crisis resolution within his initial quantitatively delimited domain.

Whether moving up the ladder of abstraction from qualitatively to a quantitatively construed concept or vice versa, MMR must contend with conceptual mismatch that leads to the muddling of mechanisms. Mechanism muddling damages the claim that nested analysis yields a more comprehensive, multidimensional causal account. This occurs because not all the cases within a large-*N*

population share the properties necessary to activate a particular causal pathway specified in the qualitative domain. Mechanism muddling leads causal theories to be underspecified and imbalanced.

## Conceptual Slippage

The disparities between qualitative and quantitative approaches to concepts are not limited to connotation and denotation. Rather, they extend to the use of different schemas for measuring and categorizing cases. Qualitative measurement necessarily involves categorizing cases using specific nominal and ordinal criteria that often combine multiple dimensions or characteristics. Membership in each category is absolute, and each case is equivalent. By contrast, quantitative measurement involves scoring cases so that every case is related to the other on a basic interval or ratio scale. Cases are measured along a single dimension (Mahoney 2000; Goertz 2008). Because differing schemas imply a different set of relationships among referent cases, conceptual slippage can be seen as a form of measurement error where cases are inaccurately or inconsistently deemed equivalent (or disparate) because of the discrepancies in taxonomical systems (Franzosi 2004, 281; Jacoby 1999; Zerubavel 1996). Of course, this can occur in single-method studies, sometimes as a result of sloppiness or the limitations of natural language (Bryman 1988, 127). Claiming simultaneously that two cases are equivalent members of a particular set and that one is “more” of that set than the other introduces significant ambiguity to the term and the measurement scale (Mahoney 2003). Such a combination of nominal and ordinal schemas requires considerable conceptual elaboration (Goertz 2006, 82).

Slippage is a particularly acute problem for MMR because of the inherent disparity between qualitative and quantitative taxonomical schemas. Consider as examples the concepts of development and regime type. Quantitative research can make ready use of per capita GDP as a continuous variable to measure development. Since the variable is expressed in interval format, quantitative research can tell that the United Kingdom, Sweden, Denmark, and the United Arab Emirates are more or less equivalent (around \$35,000 per capita GDP) and are five times more developed than Ukraine (\$7,000). Qualitative researchers, in contrast, cannot make use of these kinds of fine-tuned distinctions. Instead, they rely on a categorical set that treats certain degrees of difference as less relevant than others. Thus, Sweden, Denmark, and the United Kingdom might be grouped as “developed” economies, Ukraine as moderately developed. The United Arab Emirates, despite a high per capita GDP, might be grouped with the likes of Kuwait, Saudi Arabia, and

Venezuela in a special category of “highly developed rentier states.”<sup>4</sup>

On the other hand, democracy is a concept difficult to quantify because of its multidimensionality and manifest types and subtypes. Case-based researchers usually apply a proliferation of such classifications and typologies (Collier and Levitsky 1997). To gain greater specificity, qualitative researchers move down the ladder of abstract, adding features that differentiate cases from one another. The category of democracy is further specified by differentiating parliamentary and presidential-type democracies. Some conceptualizations of regime type are tailored to specific regions.<sup>5</sup> On discovering a case too dissimilar to fit under the existing definitional rubric, the qualitative researcher will move up the ladder of abstraction to create a higher level category that subsumes the existing lower level. Thus, democracy becomes a type of political regime alongside authoritarianism and totalitarianism. These are not placed on a continuum, but each shares a variety of characteristics in common with each other (Linz 2000). By comparison, quantitative analysis relies on numerical datasets, such as Freedom House and POLITY, which treat democracy as a continuum. At best, these datasets reveal that two countries are equally democratic (or undemocratic). On quantitative scales, Iran is equivalent to Swaziland since both received Freedom House ratings of 6 for 2008. Even though the values are expressed in an interval format, there is no substantive meaning to the distance between intervals, such that the exact relationship between Iran (6) and North Korea (7) remains ambiguous (Munck and Verkuilen 2002; Hanson and Kopstein 2005).

When combining continuum and categorical scaling in MMR, significant conceptual slippage can occur. A good example comes from Lange’s (2009) *Lineages of Despotism and Development*. Lange hypothesizes that the difference between direct and indirect rule under the British empire lay the ground work for a path-dependent process of social, economic, and political development among postcolonial states. States that were directly governed by Britain inherited superior institutions that contributed over time to superior attainment, while states that endured more indirect forms of control inherited less capable institutions and saw the empowerment of local strongmen that retarded development over time. Lange describes the differences between the ideal types of indirect and direct colonial rule. Indirect rule involves collaboration between a dominant colonial center and several regionally based indigenous institutions. In contrast, direct rule entails the construction of a complete system of colonial domination in which both local and central institutions are well integrated and governed by the same authority and organizational principles. Lange (2009, 29) notes, however, that in practice the distinction between

**Table 3.** Percentage of Cases in Customary Courts.

Colony	% of customary court cases	Colony	% of customary court cases
Australia	0	Gambia	37
Bahamas	0	South Africa	39
Barbados	0	Zimbabwe	40
Belize	0	Botswana	43
Canada	0	India	49
Cyprus	0	Swaziland	49
Egypt	0	Bangladesh	50
Guyana	0	Lesotho	50
Hong Kong	0	Pakistan	50
Jamaica	0	Solomon Islands	52
Mauritius	0	Fiji	55
New Zealand	0	Kenya	59
Singapore	0	Zambia	60
Sri Lanka	0	Ghana	65
Trinidad	0	Sudan	73
United States	0	Tanzania	75
Brunei	1	Uganda	80
Malaysia	6	Sierra Leone	81
Myanmar	16	Malawi	82
		Nigeria	93

Source: Lange (2009, 48).

direct and indirect rule is “more of a spectrum than a strict dichotomy.” On one hand, there were only a handful of cases of completely integrated direct colonial rule. On the other hand, indirect rule never relied entirely on native collaboration. Instead, Lange reports, there were many cases of hybrid forms, such as Malaysia and India, “combining colonial and indigenous institutions in different ways and to different extents.”

In his initial quantitative approach, Lange operationalizes the form of control by measuring the percentage of court cases in each colony handled by indigenous courts based on customary law as opposed to colonial courts administered by the British. The larger the role played by customary courts, the less direct the form of colonialism. Using colonial historical records, Lange covers thirty-nine cases (see Table 3). When added as an independent variable to a statistical equation of economic and social development, the results are consistent with the hypothesis: countries that were subject to more direct forms of colonization have markedly better performance than those with less direct colonization in terms of per capita GDP, levels of democracy, average school attainment, and infant mortality.

Following the techniques of nesting analysis, Lange proceeds to select four cases for further in-depth interrogation: Mauritius, Botswana, Guyana, and Sierra Leone.

**Table 4.** Rubric for Qualitative Comparison.

	Outcome (Y): Positive development	Outcome (Y): Negative development
Antecedent condition (X): Direct rule	Mauritius (0 percent cases in customary courts)	Guyana (0 percent cases in customary courts)
Antecedent condition (X): Indirect rule	Botswana (43 percent of cases in customary courts)	Sierra Leone (81 percent of cases in customary courts)

Source: Adapted from Lange (2009, 15).

He selects these cases with an eye to applying the logic of both most similar and most different designs, as shown in Table 4. The Mauritius–Botswana and Guyana–Sierra Leone comparisons are used to highlight determinants of similar outcomes. The Mauritius–Guyana and Botswana–Sierra Leone comparisons are used to highlight determinants of dissimilar outcomes.

Yet this qualitative structure betrays a change in the conceptualization of the key independent variable. Rather than being a matter of degrees between indirect and direct rule, as the concept had initially been treated in quantitative analysis, in the qualitative portion colonial rule is treated as a simple bivariate. This switch implies some troubling logical claims: In this qualitative setting, Botswana is considered equivalent to Sierra Leone as an examples of indirect rule. In the quantitative setting, however, Botswana's metric of indirect rule based on the prevalence of indigenous courts is only about half that of Sierra Leone. Moreover, the relationship between Botswana (43 percent) and cases such as Malaysia (6 percent) and India (49 percent), which Lange had previously characterized as hybrid, is also ambiguous. The categories and scales by which Lange arrays his cases are internally inconsistent within the study. The units of analysis are not homogenous. At various points Lange has treated his key independent variable as a matter of degrees (as used in the quantitative component), a trichotomy (direct, hybrid, or indirect rule), or a dichotomy (as direct or indirect rule). How the cases relate to one another analytically is unclear. Consequently, efforts to generalize from any specific cases to wider categories of cases becomes extremely problematic.

Similar to the effect of conceptual stretching, slippage leads to ambiguity in denotation of cases. Specifically, it affects the ways cases are related to one another and are located within a specific coding schema. This violates the core assumption of unit homogeneity and equivalence (Gerring 2007, 52). The problem is particularly

dire, then, for MMR's goal of mutual confirmation of results. If the cases selected for analysis via large-*N* and small-*N* are not equivalent, then the coincidental congruence of findings in these two metrics is analytically inconsequential.

### *Addressing Concept Misformation: Standardization versus Flexibility?*

The most intuitive way to address the problem of concept slippage is to push toward greater conceptual standardization. Coppedge (1999, 2009) expresses hope that eventually the scholarly community might converge on definitions of key terms that are thick enough for use in case studies but thin enough to be applied on a large scale in statistical analyses:

As we climb the ladder of abstraction, we must leave behind the attributes that are irrelevant and take with us all the attributes that matter for the theory at hand. Unfortunately, knowing which attributes matter is hard. It requires round after round of theorizing and systematic testing. (Coppedge 2009, 16)

Similarly, Lieberman (2005, 436) claims that nested analysis MMR can contribute to the progress of concept formation by systematically and iteratively juxtaposing the connotation and denotation of qualitative and quantitative concepts, thereby forcing scholars to contend head on with the potential mismatch.

These measures face significant practical and theoretical obstacles, however. Dunning (2007) recommends that researchers familiarize themselves with variable features and underlying conceptual definitions in a data set by coding at least a handful of cases manually to recognize and rectify divergence between qualitative and quantitative conceptualizations. Even before this labor-intensive inductive process begins, though, a more or less exhaustive schema of conceptual subcategories and antonyms must be developed deductively to avoid shunting problematic cases into meaningless residual categories (Bailey 1994, 17-34; Goertz 2006, 32-33).

Lange's work, for instance, could have benefitted from a more thorough examination of the initial statistical data to develop a single, unified schema to measure the type of colonial control in both qualitative and quantitative domains. Instead of using incommensurate scalar and dichotomous variables, he might have offered a consistent conceptualization of colonization (such as the direct, hybrid, or indirect trichotomy) that establishes clear, conceptually defensible cut points. He would have to offer a substantive explanation for whether Malaysia (6 percent cases in customary courts) should be considered closer



to directly colonized Guyana (0 percent) or indirectly colonized Myanmar (16 percent). Indicatively, some of the most promising efforts in this regard come through multi-scholar collaborative projects, such as the Research Network on Gender Politics and the State (McBride and Mazur 2010) and the World Bank's project on the causes of civil wars (Sambanis 2004a, 2004b).

Standardizing concepts, however, does not address the question of mechanism muddling, which is essentially a problem of equifinality and multiple pathways leading to the same outcome. While statistical analysis often overlooks the phenomenon, it is a critical component in effective qualitative research (Mahoney and Goertz 2006; Bennett and Elman 2006; Braumoeller 2003). In the cases of Lieberman and Schultz discussed above, the problem is not that the mechanisms are specified inaccurately but that they were logically insufficient to account for the regularity across all the cases in the set because of the concept of misformation. In other words, a significant portion of the correlation observed in the statistical portion must have been from mechanisms *other* than the ones identified in the case studies.

Considerable conceptual elaboration is necessary to enable the search for potentially manifest but equifinite mechanistic pathways. Conceptual definitions must be made more flexible to capture the variety of the referent cases and reflect a range of embedded causal properties. Just as was the case for conceptual standardization, concept formation must be wary of creating residual categories that have little analytical purchase. Rather than trying to simplify categorization schemas, though, the likely outcome is a proliferation of subtypes and the creation of an even more intricate conceptual hierarchy.<sup>6</sup>

Returning to the example of Lieberman, while the factors of race and regionalism identified in the case studies of South Africa and Brazil may have had a significant impact in defining NPC and in turn spurring the creation of more or less robust welfare states, these mechanisms cannot account for the majority of the cases where formally specified regions and racial cleavages are both absent. A more flexible conceptual format would suggest other forms of NPC that could have a similar impact. For instance, the experience of war and mass mobilization could redefine boundaries of citizenship and national belonging in much the same way that Lieberman claims racial or regional cleavages do. In fact, a number of studies show that when NPC is defined by war, citizens often demand redistributive fiscal and tax policies (Centeno 2003; Skocpol 1995). This third NPC subtype could be examined qualitatively to identify another mechanism that accounts for the emergence of progressive (or regressive) fiscal and tax policies in a wider span of relevant cases.

Similarly, in the example of Schultz, recognizing diversity within democratic states rather than relying on Great Britain as a prototype would prompt further qualitative research to explain peaceful conflict resolution when transparent foreign policy decision making is absent. In such opaque democracies, it is possible that a perception of shared liberal values causes them to adopt a more conciliatory attitude toward democratic rivals (Hayes 2009). In MMR's aim for comprehensive causal accounts, these alternative mechanisms constitute not rival hypotheses but complementary efforts that offer a more inclusive delineation of the pathways that undergird correlative relationships. To recognize these potential mechanisms, however, concepts must be built that cover the range of empirical variation within the relevant cases.

Pursuing the goals of concept standardization and flexibility pulls the MMR practitioner in two different directions. On one hand, standardization leads MMR closer to mimicking the techniques of quantitative analysis. Standardization, after all, ensures there is unit homogeneity between qualitative and quantitative components and that there is no measurement error or discrepancies between the two. On the other hand, flexibility fosters the recognition of multiple causal pathways by proliferating an array of conceptual subtypes. This is more in line with the techniques of contextualized comparisons of seemingly disparate cases commonly deployed in qualitative analysis (Locke and Thelen 1995). These steps, though, make it harder to fit cases into a statistical format because of the numerous forms and permutations of case categorizations. Reversing the logic of "thickening" thin concepts, some concepts are already too thick to be easily converted into statistical format. Ultimately, MMR practitioners are likely to be drawn toward whichever pole they deem adds greater analytical value, sacrificing the other component of the study.

## Conclusion

No conceptual definition is written in stone. Any meaningful comparative method depends ultimately on pushing concepts to new and unfamiliar historical, geographic, or theoretical terrain. In this manner, unfounded (and often unstated) assumptions about a particular concept's applicable scope are challenged empirically and the concept's essential components are uncovered and made more explicit (C. Chen and Sil 2007). Indeed, among the promises of MMR is that it can integrate qualitatively derived, region-specific knowledge with larger cross-national, quantitative analysis in a new form of comparative area studies (Ahram 2010).

Still, Sartori's warnings about conceptual misformation stand as a reminder that transparency and consistency in refining concepts are also critical. Otherwise, the search

**Table 5.** Impact and Possible Solutions to the Problems of Concept Misformation in Multimethod Research (MMR).

	Multidimensionality	Mutual validation
Described objective	MMR produces more comprehensive causal accounts covering both correlation and causal pathways	Findings in separate qualitative and quantitative investigations confirm each other
Potential problem	Mechanism muddling: $X_{\text{qual}}$ and $X_{\text{quant}}$ imply the presence of different causal properties	Conceptual slippage: $X_{\text{qual}}$ and $X_{\text{quant}}$ use different classification schemas
Impact	Mechanistic account remains underspecified relative to the broader correlations	Generalization from individual cases to categories of cases is impossible
Ways to address the problem	Develop flexible concepts that support multiple causal pathways and mechanisms undergirding a single correlation	Standardize concepts and classification schemas using inductive and deductive concept formation

for generality in application is liable to yield concepts so distorted as to be meaningless. This challenge is significantly magnified for MMR and thus requires more proactive efforts to head off problems associated with concept misformation. As summarized in Table 5, different forms of concept misformation undermine different aspects of MMR's claim to improve social science research overall. The goal of producing multidimensional accounts of both correlative patterns and mechanistic causal pathways is damaged by mechanism muddling. If different causal properties are implicit in different conceptual definitions, then any causal pathway that is identified through qualitative inquiry is bound to be underspecified relative to the broad correlation identified in the quantitative setting. The goal of mutual validation between qualitative and quantitative methods is damaged by conceptual slippage. If qualitative and quantitative methods work with different schemas to categorize cases, then the relationship between cases is inconsistent, making it impossible to generalize about groups of cases.

There is no easy fix to these twin problems, and addressing one may well exacerbate the other. On one hand, concept standardization confronts slippage. It ensures that units are homogeneous, a key assumption in quantitative analysis. On the other hand, greater conceptual flexibility is often necessary to capture multiple mechanistic pathways in qualitative analysis. Ultimately, MMR confronts the same dilemma as single method qualitative and quantitative studies, forced to choose between depth and breadth of findings.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### Notes

1. See Goertz (2008), Collier and Mahon (1993), and Adcock and Collier (2001).
2. Also see X. Chen (1990).
3. On different subtypes of democracy and foreign policy, see Peterson (1996), Elman (2000), and Palmer, London, and Regan (2004).
4. For an extended discussion of categorization schemes, see Ragin (2000). For an example of a perfunctory attempt to justify excluding oil-producing states from the study of development, see Przeworski et al. (2000).
5. See Bratton and Van de Walle (1994) on African regime types.
6. For a good example of such an elaborate conceptualization, see Møller and Skaaning (2010).

### References

- Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95:529-46.
- Ahmed, Amel, and Rudra Sil. 2009. "Is Multi-Method Research Really Better?" *Qualitative and Multi-Method Research* 7:2-6.
- Ahram, Ariel. 2010. "The Theory and Method of Comparative Area Studies." *Qualitative Research* 11:69-90.
- Bailey, K. D. 1994. *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, CA: SAGE.
- Bennett, Andrew, and Colin Elman. 2006. "Complex Causal Relations and Case Study Methods: The Example of Path Dependence." *Political Analysis* 14:250-67.
- Bratton, Michael, and Nicolas Van De Walle. 1994. "Neopatrimonial Regimes and Political Transitions in Africa." *World Politics* 46:453-89.
- Braumoeller, Bear. 2003. "Causal Complexity and the Study of Politics." *Political Analysis* 11:209-33.
- Bryman, Alan. 1988. *Quantity and Quality in Social Research*. London: Routledge.

- Bryman, Alan. 2004. *Social Research Methods*. New York: Oxford University Press.
- Bryman, Alan. 2006. "Paradigm Peace and the Implications for Quality." *International Journal of Social Research Methodology* 9:111-26.
- Centeno, Miguel. 2003. *Blood and Debt: War and the Nation-State in Latin America*. University Park: Pennsylvania State University Press.
- Chatterjee, A. 2009. "Ontology, Epistemology, and Multi-Methods." *Qualitative and Multi-Method Research* 7 (2): 11-14.
- Chen, Cheng, and Rudra Sil. 2007. "Stretching Post-Communism: Diversity, Context, and Comparative Historical Analysis." *Post-Soviet Affairs* 23:275-301.
- Chen, Xiang. 1990. "Local Incommensurability and Communicability." *Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1:67-76.
- Collier, David, and Henry Brady, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Roman & Littlefield.
- Collier, David and James Mahon. 1993. "Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Analysis." *American Political Science Review*, 87: 845-855.
- Collier, David, and James Mahoney. 1996. "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49:56-91.
- Collier, David, and Steven Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49:430-51.
- Coppedge, Michael. 1999. "Thickening Thin Concepts and Theories: Combining Large N and Small in Comparative Politics." *Comparative Politics* 31:465-76.
- Coppedge, Michael. 2005. "Explaining Democratic Deterioration in Venezuela Through Nested Induction" In *The Third Wave of Democratization in Latin America*, edited by Francis Hagopian and Scott Mainwaring, 289-318. New York: Cambridge University Press.
- Coppedge, Michael. 2009. "Speedbumps on the Road to Multimethod Consensus." *Qualitative and Multi-Method Research* 7 (2): 15-17.
- Denzin, Norman. 1978. *The Research Act: A Theoretical Introduction to Sociological Methods*. New York: McGraw-Hill.
- Dunning, Thad. 2007. "The Role of Iteration in Multi-Method Research." *Qualitative Methods* 5:22-24.
- Elman, Miriam. 2000. "Unpacking Democracy: Presidentialism, Parliamentarism, and Theories of Democratic Peace." *Security Studies* 9 (4): 91-126.
- Erzberger, Christian, and Udo Kelle. 2003. "Handbook of Mixed Methods in Social and Behavioral Research." In *Handbook of Mixed Methods in Social and Behavioral Research*, edited by A. Tashakkori and C. Teddlie, 457-90. Thousand Oaks, CA: SAGE.
- Erzberger, Christian, and Gerald Prein. 1997. "Triangulation: Validity and Empirically-Based Hypothesis Construction." *Quality and Quantity* 31:141-54.
- Fearon, James, and David Laitin. 2008. "Integrating Qualitative and Quantitative Methods." In *Oxford Handbook of Political Methodology*, edited by Janet Box-Steffensmeier, Henry Brady, and David Collier, 756-78. New York: Oxford University Press.
- Franzosi, Robert. 2004. *From Words to Numbers: Narrative, Data, and Social Science*. New York: Cambridge University Press.
- Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. New York: Cambridge University Press.
- Gerring, John. 2005. "Causation: A Unified Framework for the Social Sciences." *Journal of Theoretical Politics* 17:163-98.
- Gerring, John. 2007. *Case Study Research: Principles and Practices*. New York: Cambridge University Press.
- Gerring, John. 2011. "How Good Is Good Enough? A Multi-dimensional, Best-Possible Standard for Research Design." *Political Research Quarterly* 64:625-36.
- Gerring, John and Paul Baressi. 2003. "Putting Ordinary Language to Work: A Min-Max Strategy of Concept Formation in the Social Sciences." *Journal of Theoretical Politics*, 15:201-232.
- Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.
- Goertz, Gary. 2008. "Concepts, Theories, and Numbers: A Checklist for Constructing, Evaluating, and Using Concepts or Quantitative Measures." In *Oxford Handbook of Political Methodology*, edited by Janet Box-Steffensmeier, Henry Brady, and David Collier, 97-118. New York: Oxford University Press.
- Goertz, Gary, and Amy Mazur. 2008. "Mapping Gender and Politics Concepts: Ten Guidelines." In *Politics, Gender and Concepts*, edited by Gary Goertz and Amy Mazur, 14-46. New York: Cambridge University Press.
- Guba, Egon, and Yvonna Lincoln. 2005. "Paradigmatic Controversies, Contradictions, and Emerging Confluences." In *The Sage Handbook of Qualitative Research*, edited by Norman K. Denzin and Yvonna S. Lincoln, 191-215. Thousand Oaks, CA: SAGE.
- Hanson, Stephen, and Jeffrey Kopstein. 2005. "Regime Type and Diffusion in Comparative Politics." *Canadian Journal of Political Science* 38:66-99.
- Hayes, Jarrod. 2009. "Identity and Securitization in the Democratic Peace: The United States and the Divergence of Response to India and Iran's Nuclear Programs." *International Studies Quarterly* 58:977-99.
- Jacoby, William. 1999. "Levels of Measurement and Political Research: An Optimistic View." *American Journal of Political Science* 43:271-301.
- Kelle, Udo. 2001. "Sociological Explanations between Micro and Macro and the Integration of Qualitative and Quantitative Methods." *Qualitative Social Research* 2:1-19.
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University press

- Lange, Matthew. 2009. *Lineages of Despotism and Development: British Colonialism and State Power*. Chicago: University of Chicago Press.
- Lieberman, Evan. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. New York: Cambridge University Press.
- Lieberman, Evan. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99:435-52.
- Linz, Juan. 2000. *Totalitarian and Authoritarian Regimes*. Boulder, CO: Lynne Rienner.
- Locke, Richard M., and Kathleen Thelen. 1995. "Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics." *Politics & Society* 23:337-67.
- Mahoney, James. 2000. "Strategies of Causal Inference in Small-N Analysis." *Sociological Methods and Research* 28:387-424.
- Mahoney, James. 2001. "Beyond Correlational Analysis: Recent Innovations in Theory and Method." *Sociological Forum* 16:575-93.
- Mahoney, James. 2003. "Strategies of Causal Assessment in Comparative Historical Analysis." In *Comparative Historical Analysis in the Social Science*, edited by James Mahoney and Dietrich Rueschemeyer, 337-73. New York: Cambridge University Press.
- Mahoney, James. 2008. "Toward a Unified Theory of Causality." *Comparative Political Studies* 41:412-36.
- Mahoney, James, and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:227-49.
- McBride, Dorothy, and Amy Mazur. 2010. "Integrating Two Cultures in Mixed-Methods Research: A Tale of the State Feminist Project." *Qualitative and Multi-Method Research* 8 (1): 35-39.
- Møller, Jørgen, and Svend-Erik Skaaning. 2010. "Beyond the Radial Delusion: Conceptualizing and Measuring Democracy and Non-Democracy." *International Political Science Review* 21:261-83.
- Munck, Gerardo, and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35:5-34.
- Palmer, Glen, Tamar London, and Patrick Regan. 2004. "What's Stopping You? The Sources of Political Constraint on International Conflict Behavior in Parliamentary Democracies." *International Interactions* 30:1-24.
- Pawson, Ray. 1995. "Quality and Quantity, Agency and Structure, Mechanism and Context, Dons and Cons." *Bulletin de méthodologie sociologique* 47:5-48.
- Peterson, Susan. 1996. *Crisis Bargaining and the State: The Domestic Politics of International Crisis*. Ann Arbor: University of Michigan Press.
- Przeworski, Adam, Michael Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. New York: Cambridge University Press.
- Ragin, Charles. 2000. *Fuzzy Set Social Science*. Chicago: University of Chicago Press.
- Reiss, Julian. 2009. "Causation in the Social Sciences: Evidence, Inference, and Purpose." *Philosophy of the Social Sciences* 39:20-40.
- Rohlfing, Ingo. 2007. "What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative Research." *Comparative Political Studies* 41:1492-1514.
- Sambanis, Nicholas. 2004a. "Using Case Studies to Expand Economic Models of Civil War." *Perspectives on Politics* 2:260-79.
- Sambanis, Nicholas. 2004b. "What Is a Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48:814-58.
- Sankey, Howard. 1991. "Incommensurability, Translation and Understanding." *Philosophical Quarterly* 41:414-26.
- Sankey, Howard. 2000. "Kuhn's Ontological Relativism." *Science and Education* 9:59-75.
- Sartori, Giovanni. 1970. "Concept Misformation in Comparative Politics." *American Political Science Review* 64: 1033-53.
- Schultz, Kenneth. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- Sil, Rudra. 2000. "The Division of Labor in Social Science Research: Unified Methodology or 'Organic Solidarity.'" *Polity* 32:499-531.
- Skocpol, Theda. 1995. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, MA: Belknap.
- Teddlie, Charles, and Abbas Tashakkori. 2009. *Foundations of Mixed Methods Research*. Thousand Oaks, CA: SAGE.
- Waldner, David. 2007. "Transforming Inference into Explanations: Lessons from the Study of Mass Extinction." In *Theory and Evidence in Comparative Politics and International Relations*, edited by Richard Lebow and Mark Lichbach, 145-77. New York: Palgrave.
- Yanow, Dvora, and Peregrine Schwartz-Shea. 2006. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Armonk, NY: M.E. Sharpe.
- Zerubavel, Eviatar. 1996. "Lumping and Splitting: Notes on Social Classification." *Sociological Forum* 11:421-33.