

Oxford Handbooks Online

Overview Of Political Methodology: Post-Behavioral Movements and Trends

FREE

Henry E. Brady, David Collier, and Janet M. Box-Steffensmeier

The Oxford Handbook of Political Science

Edited by Robert E. Goodin

Print Publication Date: Jul 2011

Subject: Political Science, Political Methodology, Comparative Politics

Online Publication Date: Sep 2013

DOI: 10.1093/oxfordhb/9780199604456.013.0048

Abstract and Keywords

Political methodology offers techniques for clarifying the theoretical meaning of concepts such as revolution and for developing definitions of revolutions. It also provides descriptive indicators for comparing the scope of revolutionary change, and sample surveys for gauging the support for revolutions. It then presents an array of methods for making causal inferences that provide insights into the causes and consequences of revolutions. An overview of the book is given. Topics addressed include social theory and approaches to social science methodology; concepts and development measurement; causality and explanation in social research; experiments, quasi-experiments, and natural experiments; general methods of quantitative tools for causal and descriptive inference; quantitative tools for causal and descriptive inference; qualitative tools for causal inference; and organizations, institutions, and movements in the field of methodology. In general, the *Handbook* provides overviews of specific methodologies, but it also emphasizes three things: utility for understanding politics, pluralism of approaches, and cutting across boundaries. This volume discusses interpretive and constructivist methods, along with broader issues of situating alternative analytic tools in relation to an understanding of culture.

Keywords: political science, political methodology, causal inference, descriptive inference, social theory, revolution

1 Overview

With the ascendancy of “behavioralism” in political science during the mid-twentieth century, an emphasis upon careful conceptualization, precise measurement, and fastidious causal thinking emerged in the study of politics. These tasks have become the hallmarks of political methodology. Assessing the progress of methodology is therefore inextricably bound up with an evaluation of behavioralism.

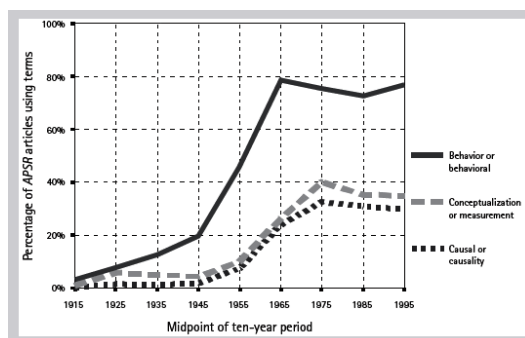


Fig. 48.1 Percentage of APSR articles using terms

(p. 1006) Behavioralism was a self-consciously “scientific” movement within political science that began in the 1930s and took off in the 1950s.¹ Now it is embedded in the warp and woof of the discipline. The movement pushed political scientists to be more specific and systematic about their methodology—about their concepts, their measures, and their causal arguments concerning the actions and behaviors of political actors. Figure 48.1 depicts the simultaneous development of behavioralism and political methodology by plotting the frequency of articles referring to these perspectives in the *American Political Science Review*, the flagship journal of the American political science profession, between 1910 and 2000. The solid line plots the terms “behavior” or “behavioral,” the dashed line the words “conceptualization” or “measurement,” and the dotted line the words “causal” or “causality.”² Behavioralism takes off around 1945 and from 1960 to the present about 80 percent of the articles include the word “behavior” or “behavioral.” Concern with causality and with conceptualization and measurement take off soon after 1945, and today about one-third of the articles in the *APSR* use the words “causal” or “causality” and about one-third use the words “conceptualization” or “measurement.” Since about 1970, roughly one-half of the articles use one or the other set of terms.

Behavioralism’s focus on human actions and its legacy of precision, care, and fastidiousness in studying them (that is, its emphasis upon methodology) is a good one, (p. 1007) but its weaknesses included a neglect of theory, an excessive reliance upon the tenets of logical empiricism, and sometimes an emphasis on investigating overt behavior to the exclusion of speech and language. Political methodology must be assessed in part in light of the weaknesses, as well as the strengths, of behavioralism. Political science has moved beyond behavioralism, but different observers see political science going in distinct directions. Some see it moving toward more theory (e.g. game theory) and empirical testing of theories—as with the Empirical Implications of Theoretical Models movement (Aldrich, Alt, and Lupia 2008); others associate it with the recent resurgence of a rigorous concern with qualitative and multimethod research (Collier and Elman 2008); and still others see it moving in the “post-positivist” direction of more narrative and interpretative work (Patterson and Monroe 1998; Yanow and Schwartz-Shea 2006).³ All these perspectives might see political methodology as too tightly bound to a “positivist” or at least nominalist, operationalist, and inductive approach to political science.

We agree that political science has moved beyond simple behavioralism, and we also believe that political methodology has done so, in ways that accommodate divergent perspectives on the road to post-behavioralism. We will argue that political methodology has made impressive strides by coming to grips with many of the criticisms of behavioralism, although recondite problems remain such as linking theory to methodology, developing methodologies for understanding speech acts and text, comprehending the role of case-studies and assessing causation for singular events, and developing methods for systematizing and improving qualitative methods. In this chapter, we begin by reviewing how methodologists now deal with conceptualization and measurement, then we discuss causal inference, and we end with a discussion of unresolved problems, the methodology organizations that are considering these problems, and future directions.

2 Conceptualization and Measurement

2.1 Introduction: Concepts and Measures

Philosophers, scientists, and political scientists have moved a long way from the Aristotelian notion that the best concepts designate “natural kinds” based upon “natural laws.” They have also moved beyond the classical philosophical notion (see Mill 1888, book I, ch. VIII)—developed further by the logical empiricists (Hempel 1952; Cohen and Nagel 1934)—that concepts can be defined by a set of necessary and sufficient conditions—by a set of attributes that the concept must have.

(p. 1008) To illustrate where we are now in our thinking about conceptualization, consider this example. In the social sciences, it was once thought that ethnic groups, nations, and races were natural kinds, but modern social science suggests that these concepts are constructed (Berger and Luckman 1966; Giddens 1986) and not natural. In their manifestation at any point in history, the coherence of ethnic groups may well rest on numerous coinciding cleavages, and the identities and solidarities associated with the groups may well be deeply embedded. But in their origin and evolution, one must consider carefully the role of social construction. At such times, individuals associated with the group may evoke a supposedly primordial characteristic such as language, descent, or territory, but a potential group usually has many possible such characteristics from which to choose and only a few

of them are selected—sometimes without a strict adherence to what an outsider might consider the reality of the presumed group (Anderson 1991; Brass 1991).

What standard of conceptualization might be employed here? One option would be that the investigator fully accepted the idea of social construction, but also set an external standard for judging when this process of construction in fact had produced a group. For example, an ethnic group might be understood as really being a group when the following statements are true: There is common agreement among its members that they belong to the group, nonmembers see them as a group, and all of these actors behave as if this is so.

Further, the perspective of social constructivism should not be taken to suggest that just anything is possible, even though some researchers (Tajfel 1970) have concluded that when experimenters bring people together in a laboratory, even the most minimal pretext can serve as the basis for the emergence of ideas about group differences, and hence for the idea that distinct groups have formed. Most discussions of this “minimal group” paradigm ignore the fact that the experimenter brought the people together in the first place. Thus, the creation of a group requires not only a pretext for imagining group differences, but also something powerful enough to bring people together in the first place to get the word out about their supposed differences from other groups. This may, for example, require that elites (sometimes an intelligentsia) deliberately formulate an ethnic identity out of the available histories and characteristics of a people (Hroch 1986). As Anderson puts it, “imagined” communities may thereby be created, involving a creative use of the facts.

We see here a crucial link to issues of measurement and data. For the methodologist, this discussion shows why census, survey, and other data with social categories or concepts must be approached with great caution. As is discussed in much research on racial and ethnic politics in the USA (Lee 1993; Nobles 2000; Lee 2007) the questions on race and ethnicity on the US Census have constantly undergone change. There have been periods when citizens were asked whether they were colored or white, whether colored, quadroons, or octaroons, and whether Negro or White. Most recently, citizens could identify themselves as belonging to multiple races. Similarly, the Soviet Census (Hirsch 2005) defined and redefined nationalities and peoples in ways that fit the needs of the regime. Hence, any notion that census data are based on “hard facts” that can readily be treated as the basis for a rigorous social science must (p. 1009) be examined with great care. A student of race and ethnicity must think carefully about measuring the concept in multiple ways (Brady and Kaplan 2000; 2009) and about being sensitive to the history and narratives of the people and the regime.

This example also shows that social science concepts have an additional feature that distinguishes them from concepts in the natural sciences. Because social science concepts routinely reflect the thinking of those being studied, they must take into account those thoughts. This, in turn, calls for criteria to establish the credibility of the concepts being formed. For example, the existence of an ethnic identity requires that people in a group believe that they have the identity, and it requires that others believe that the members of the group have that identity, and it requires that the members of the group believe that others believe that they have that identity, and so forth. Thus, any analysis of a situation must take into account not only what each person thinks about themselves, but what each person thinks about each other person, and what each person thinks each other person thinks about them, and so on. In game theory, the notion that everyone has the same knowledge about everyone is called “common knowledge” (Geanakopols 1992), and studying departures from this common knowledge assumption is a vigorous field of investigation in game theory. The study of these departures links game theory, constructivism, and some ethnographic approaches. Consider, for example, the difficulties that could ensue in an interaction if a man thinks that another person (named Pat) is a woman and he thinks that Pat believes that she is a woman and he believes that Pat thinks that he thinks of her in this way, but suppose Pat does not think of (herself?) as a woman and Pat thinks that people should not think of (her?) in this way at all. Just such confusions are at the root of some of the work of constructivists, symbolic interactionists (Blumer 1969), and the ethnographic work of Erving Goffman (1959) and Harold Garfinkel (1967). Thus, constructivism and game theory are coming together in a common concern with common understandings and misunderstandings.

These examples of race, ethnicity, gender, and identity—while perhaps representing unusually difficult analytic challenges—remind us that in many domains of research, concepts and analytic categories are hard to pin down and refer to human behavior of great complexity. The challenges of conceptualization and measurement well merit our attention as methodologists, and the remainder of this section explores them.

2.2 Three Approaches to Conceptualization and Measurement

This section explores three approaches to conceptualization and measurement which we call the classic syntactic statistical approach, the semantic-pragmatic approach, and the formal modeling approach.

2.2.1 The Classical Syntactic-Statistical Approach

The classical approach to concept formation relied heavily on the tenets of logical empiricism (Hempel 1952) which distinguished between theoretical terms (e.g. (p. 1010) concepts) connected to one another through theoretical axioms and logic (thus forming a theory) and observational terms or empirical statements about the world which were connected to, and provided meaning for, concepts through rules of correspondence. Empirical statements, in turn, were meaningful only if they were verifiable—capable of being assessed as true or false.

To employ, for example, the concept of “utility” would require the truth of both some theoretical and empirical statements. At the theoretical level, the preference relation P must satisfy the crucial transitive axiom that “if alternatives A , B , and C are such that APB (A preferred to B) and BPC , then we also have APC .” With this axiom, it can be proved that an “ordinal utility function” exists with meaningful properties. One set of corresponding empirical statements is that when presented with choices A and B , the person chooses A , when presented with B and C , the person chooses B , and when presented with A and C , the person chooses A .

These statements could be verified by simply presenting a person with each pair of choices and observing whether or not each statement were true. If the empirical statements are true, then we can use the theoretical concept of “utility function” to describe the behavior of this person. This simple description hides a spate of difficulties including whether or not axioms and logic are sufficient to describe a theory, how empirical verification actually occurs, and how correspondence rules (between theoretical concepts and empirical observations) actually work. We will not get into those thickets here. Instead, we will briefly review how some early social science methodology relied upon a simplistic version of these tenets which often ignored the need for theory and the importance of meaning by inductively developing concepts from observational statements.

Standard tools of those forming concepts from quantitative data are reliability analysis (Lord and Novick 1968), validity analysis (Campbell and Fiske 1959), and scaling and factor analysis (Jackman 2008). In the 1950s and 1960s, many social science researchers started with a set of survey items or measures of the characteristics of nations (which were certainly empirically true statements) and used factor analysis or related techniques to find their underlying “dimensions” which were treated as theoretical concepts. The primary emphasis was upon discovering the structure of the data, and meaning fell out from the structures that were found.

For example, in a series of papers and books, R. J. Rummel (1963; 1966a; 1966b; 1972) and Raymond Tanter (1966) measured the “dimensions” of nations. Their method was to factor analyze numerous items (empirical statements) coded from various sources about a large number of countries. The resulting dimensions were then treated as theoretical concepts. Sometimes this led to interesting results as in the distinction (Tanter 1966) between “turmoil” (strikes, riots, demonstrations, governmental crises) and “internal war” (guerrilla war, purges, revolts, and domestic deaths from violence) within countries, but sometimes it led to odd and indefensible dimensions (Doran, Pendley, and Atunes 1973). A defender of logical empiricism might note that the failing here is a lack of underlying theory, but a critic might note that the deeper problem is the inattention to meaning and an emphasis upon structure and syntax instead of semantics. This approach was repeated in other areas (p. 1011) of political science, with more success, in the scaling of legislative votes (MacRae 1970) and the scaling of attitude items (Stouffer et al. 1950), partly because more attention was paid to the meaning of the resulting concepts. Our take is that these methods can be very useful, but they require substantial attention to both meaning and theory.

2.2.2 The Semantic-Pragmatic Approach

This perspective on concepts and measurement systematically explores the diverse meanings of the concepts employed by social scientists, and seeks to develop productive ways—in light of a pragmatic concern with the research tasks at hand—of understanding, coordinating, adapting, and sometimes sharply modifying these meanings. A central idea is that, especially in light of the high degree of technification of much contemporary work on measurement in political science, it is invaluable to maintain a sustained focus on the sometimes contrasting meanings of the concepts used by researchers. This focus is routinely enriched by careful attention to context,

which is valuable in its own right, and—to the extent that the scholar is concerned with systematic measurement—can contribute decisively to meeting more adequately the conventional standards of measurement validity.

The semantic-pragmatic perspective has sometimes been seen as the Sartori–Collier tradition of concept analysis (Bevir and Kedar 2008; Goertz 2006, 1, 69). It has been called a tradition of “qualitative concept formation” (Bevir and Kedar 2008, 503, 509). Yet this designation overlooks the interest of many relevant scholars—including Sartori—in linking careful work with concepts not only to qualitative, but also to quantitative, measurement. In addition to the studies by Sartori and Collier discussed below, important work in this tradition includes books and numerous articles by Gerring (e.g. 2001) and Goertz (e.g. 2006), Elman’s (2005) analysis of typologies, efforts by Kotowski (1984) and Kurtz (2000) to untangle concepts through a systematic accounting of contrasting meanings, and studies that seek to address dilemmas of causal inference through more careful treatment of concepts (Levitsky 1998; Paxton 2000; Kurtz 2000).

In this tradition, the pragmatic goals that shape choices about concepts are sometimes contradictory. For example, these goals may include a concern with achieving analytic generality, an objective that routinely must be traded off against the priorities of adapting concepts to different spatial and historical contexts and/or working with actor-defined rather than observer-defined sources of meaning. Other central objectives include dealing in a coherent and productive way with the normative valence of concepts; seeking good procedures for linking concepts with observations about the world; and finding appropriate standards for sorting out the interplay between disputes about concepts and alternative choices about observation and measurement.

The wider history of concept analysis in political science certainly extends back to classical political theory, and many of the concerns addressed in the semantic-pragmatic approach have their roots in the work of major figures as different as Weber and Wittgenstein. However, viewing the semantic-pragmatic approach as a component of contemporary political science methodology, we see it as deriving from four analytic currents. The first and most important is the period of intense (p. 1012) methodological innovation in the late 1960s and early 1970s in the field of comparative and international studies. For present purposes, this innovation is seen most crucially in the work of Sartori (1970; see also 1975; 1984; 1991) and in the contributions of many other scholars, for example, Przeworski and Teune (1970) and Verba (1967; 1971a). The second current is work in sociology by scholars such as Barton (1955) and Barton and Lazarsfeld (1969; see also McKinney 1966 and Tiryakian 1968), who focused on typologies, classification, and the interplay between theoretical and inductive analysis in concept-formation. A third line of influence can be dated back to the work of the British political scientist W. B. Gallie (1956a; 1956b; see also Collier, Hidalgo, and Maciuceanu 2006), who has stimulated a long trajectory of reflection and writing on conceptual structure, conceptual confusion, and the normative content of concepts. Finally, the fields of linguistics and cognitive science (Lakoff 1987; for recent syntheses see Taylor 2003; Cruse 2004) have contributed important insights into conceptual hierarchies, framing, and questions of the well-boundedness (or otherwise) of concepts.

Several features of the semantic-pragmatic approach should be underscored. First, at the same time that it is rigorous and systematic, it is committed to recognizing diversity of conceptual meanings, both as a potential source of analytic richness to be preserved, and as a source of potential confusion that can lead to analytic disaster. A recurring concern is with choices about when the standardization of meaning is productive, and when it gives up too much.

Second, we find a focus on the specificity of context, involving the challenges of addressing contrasts in the national, subnational, and temporal/historical settings under analysis. Sartori (1970) played a key role in inaugurating the discussion of this challenge with his arguments about how scholars can adapt their concepts to distinct contexts—thereby avoiding conceptual stretching—through movement on the ladder of abstraction. Numerous subsequent discussions have explored tools for the valid conceptualization and measurement of political phenomena across different settings (a brief overview is found in Adcock and Collier 2001, 534–36).

Sartori’s concern with the ladder of abstraction has been extended in discussions of two types of conceptual hierarchies: the kind hierarchy (basically the same as Sartori’s ladder) and the part–whole hierarchy (Collier and Levitsky 2008). An important aspect of movement on these hierarchies is the creation of subtypes (the concept with adjectives—see Collier and Levitsky 1997; Goertz 2006). With regard to kind hierarchies, Goertz has worked with the idea from cognitive linguistics of the “basic level,” which is then a point of departure for his argument

about developing two-level theories.

A third focus is on the normative valence of concepts, i.e. how they convey ethical judgements (Gallie 1956a). One need not spend much time reasoning about such standard political science topics as democracy, justice, equality, and the rule of law—as opposed to totalitarianism, war, genocide, torture, and rape—to recognize that each of these concepts has a fundamental evaluative component. An important goal that animates political science as a discipline is indeed to understand and explain the successes and disasters of human politics. Finding dependent variables that are (p. 1013) deemed humanly important is thus routinely seen as a high priority. This is certainly the case even in areas of the discipline where advanced quantitative or formal tools are employed.

The point here is not that the normative evaluations entailed in particular concepts will gain universal agreement. There are “just wars,” wars that should be avoided, and debates about which is which. Given these debates, scholars should deal frankly with the normative content of concepts. Classic substantive studies—for example Dahl’s *Polyarchy* (1971, esp. ch. 2) and O’Donnell and Schmitter’s *Transitions from Authoritarian Rule* (1986, esp. ch. 1)—are admirable for many reasons, including their frank attention to normative issues. These issues must be a central focus in the treatment of concepts, and a key ongoing issue is the sharply contrasting degrees of optimism and pessimism about whether normative disputes undermine coherent work with political science concepts (Gallie 1956a).

A further question is whether this approach focused on normative issues—and indeed the semantic-pragmatic approach more broadly—is exclusively concerned with concepts used by scholars. Whereas Gallie (1956a, 183) focused on concepts employed in the academy, subsequent writing in the tradition of Gallie provides a useful reminder that careful work with political science concepts must be alert to the embeddedness of their meanings in a larger context of politics and public discourse (Freeden 1994, 141). And indeed, as one looks at the dramatic evolution over the decades in the concepts used by political scientists, as well as the larger forces that motivate this evolution, it is easy to see that this wider perspective is essential. An excellent example is the evolving meaning of “neoliberalism,” which Boas and Gans-Morse (2009) analyze, drawing on Gallie’s framework.

A fourth and final issue is the interplay between qualitative and quantitative approaches. Many scholars in the semantic-pragmatic tradition are indeed focused on qualitative methodology. For example, work on typologies—such as that by Elman (2005) and Collier, LaPorte, and Seawright (2008)—explores how careful work with concepts can yield well-crafted categorical variables. Yet many analysts are likewise concerned with quantitative measurement, and Sartori’s famous injunction (1970, 1038) that “concept formation stands prior to quantification” does not at all mean that he opposes quantitative work in political science (or mathematical formalization, for that matter). Indeed, he states that ultimately, building on appropriate work with concepts, “the more we measure, the better” (1975, 296). Adcock and Collier (2001) discuss how the complexity and ambiguities of work with concepts can be linked to careful measurement. Among many points, they advocate maintaining a clear separation between dealing with and potentially resolving disputes about conceptual meaning, as opposed to questions of measurement validity. The first must be dealt with before the second can be appropriately addressed. Various approaches have been adopted in building this bridge. Coppedge and Reinicke (1990), as well as Munck and Verkuilen (2002)—among many other studies—link careful work with concepts and the development of indicators. Elkins (2000) uses quantitative tools to address the question, sometimes identified with the qualitative tradition (see Collier and Adcock 1999), (p. 1014) of whether it is most productive to view democracy versus nondemocracy as a dichotomy, or in terms of gradations.

To conclude, the semantic-pragmatic approach is not narrowly a tradition of qualitative concept analysis. It is a multimethod tradition in which many scholars are strongly concerned with building bridges to quantitative measurement. However, we find a distinctive focus, which might be thought of as more typical of the qualitative tradition, on sustained attention to concepts, careful reasoning with categorical variables, and attention to crafting concepts (and measures) to accommodate distinctive features of context. During a period in which “multimethod” sometimes becomes a slogan in our discipline that pushes scholars to use multiple tools poorly—rather than one tool with skill—we feel that this semantic-pragmatic approach is an area of multimethod work that points scholarly attention in a productive direction.

2.2.3 The Formal Modeling Approach

The simple example of the “utility model” described earlier shows that models can help to clarify, even perhaps

define, concepts. Choice-theoretic and game theory modeling have certainly done this for many political science concepts. Consider, for example, the notions of “the general will” and “democracy.” One exegesis of Rousseau’s idea of the general will is that it is the expression of the public’s most preferred policy as determined by democratic voting procedures such as majority rule.⁴ But modern social choice theory (Riker 1982) suggests that things are not so simple. Because of “cycling” (Arrow 1951), democratic voting procedures may not lead to any one best choice that could be considered the “most preferred alternative.” Social choice theory, therefore, certainly rules out one interpretation of “the general will” as the “most preferred alternative.”

Or take the concept of “cleavages,” which starts to appear in political science articles around the 1890s and is used casually to mean differences in politically relevant sociodemographic characteristics,⁵ in politically relevant issue positions,⁶ or in political groups themselves.⁷ Initial elaborations of the concept (Rice 1928) were purely statistical and used differences in voting support for party candidates (p. 1015) based upon geographic location or type of economic activity to examine cleavages.⁸ Rice even noted that “[t]heoretically the electorate might be divided up in well-nigh infinite variety of ways,” but he says nothing about how some ways might become paramount for politics.⁹ His approach is baldly nominal, operational, and empirical, with no explanation of which cleavages are important or where they come from.

Work by Peter Odegard in 1935 developed a simple model in which cleavages were the result of conflicting “group pressures” based upon interests “which are compounds of economic, historical, traditional, and even ethnic influences” (Odegard 1935, 69). E. E. Schattschneider (1952, 19) extended this model by noting “that an indefinite multiplication of conflicts is extremely unlikely even in a free political system” because “conflicts interfere with each other, because the lines of cleavage never or rarely coincide.” In 1960 he expanded on these ideas in a book where he presented simple two-dimensional circles in which each member of the electorate had some position in the circle representing his or her interests. Lines ran across these circles indicating “cleavages” between groups such as Democrats or Republicans of the 1870s who supported different policies for “Civil War Reconstruction” (the North–South cleavage) or the Populists and the Bourbons of the 1890s who supported different agricultural and monetary policies. Schattschneider’s work represents the first pictorial description of a cleavage in a multidimensional issue space, but he did not explain the origins of these issue cleavages nor did he explain how the line of cleavage formed, except that it involved political groups such as political parties or movements.

Independently of Schattschneider’s work, Duncan Black (1948) and Anthony Downs (1957) introduced a one-dimensional “left–right” issue space in which legislators or voters voted for alternative motions (Black) or parties (Downs) located in this same space. Black’s contribution was to show that if individual preferences along this one dimension satisfied the property of “single-peakedness”¹⁰ then in a pairwise consideration of motions at most only one would get a simple majority over all others and this motion would be the one preferred by the median voter—the voter with his or her ideal point at the median. Downs added two more features: the notion that voters would compare distances and vote for the alternative nearest them (which automatically satisfies single-peakedness) and the notion that just two parties would compete in the space by taking up locations in it and obtaining votes depending upon their distance from the voters. In this model, the line of cleavage between those who vote for one party versus those who vote for the other runs perpendicular to (p. 1016) the midpoint of the line drawn between the locations of the two parties in the issue space. The model stipulates very clearly that a cleavage is defined by both individual preferences and the locations of parties.¹¹ Downs’s model was extended to multiple dimensions by Davis, Hinich, and Ordeshook (1970) who generalized his result in a model where cleavage lines continue to be perpendicular to the midpoint of the line drawn between the locations of the two parties.¹²

These models make it clear that a cleavage depends both on the distribution of individual preferences and on the political actors such as parties who help to define them by their position-taking. This elaboration of the concept of political cleavage has implications for their genesis and their alteration. If major historical events (Reformation and Counter-Reformation, French and Democratic Revolutions, Industrial Revolution, or the Russian and Communist Revolutions) cause changes in preferences as in the classic Lipset and Rokkan (1967) paper on “Cleavage structures, party systems, and voter alignments,” then lines of cleavage might change but *only* if there are political entrepreneurs who exploit them. Even without any change in preferences, politicians might maneuver and change the lines of cleavage by modifying their positions or by emphasizing one cleavage over another (Riker 1986; Johnston et al. 1992, ch. 3). In this case, formal theory helps to clarify the concept of cleavage by showing that it is the result of the interaction between position-taking by political elites and the preferences of mass publics. More

generally, formal theory has helped to clarify many concepts such as utility, deterrence, arms races, party identification, collective action, and many more.

2.3 Data Collection and Measurement

The data available to social scientists have increased dramatically in the past sixty years, partly as a result of the behavioral revolution which emphasized the collection of data of all sorts including quantitative and interview data. It is true that prior to 1950 some political scientists analyzed aggregate election data (e.g. Gosnell and Gill 1935; Key 1949), legislative votes (Rice 1928; Beyle 1931; Brimhall and Otis 1948; Gage and Shimberg 1949), judicial votes (Pritchett 1941; 1945), elite biographical data (Lasswell and Sereno 1937), in-depth interviews (Weeks 1930), or quantitative media data (Lasswell 1941), but there was not much of this kind of research. For example, articles about multiple in-depth interviews do not even appear in the *APSR* until the decade of the 1930s when there were eight articles (six of which used interviews and two of which merely advocated their use). There was no increase in the 1940s with only seven articles (of which four were advocacy), but the number increased (p. 1017) dramatically in the 1950s with twenty-three articles and the 1960s with forty articles (of which only one was advocacy). We see similar patterns with surveys and legislative voting.¹³ We cannot review every mode of data collection here, but we will review a few areas where major advances have been made.

2.3.1 Surveys

Scientific surveys first appeared in the 1930s, but they did not take off in political science until two major election studies in 1948—a panel study in Erie County, Ohio by Lazarsfeld and his colleagues at Columbia University (reported in *Voting* 1954) and another national post-election study at Michigan by Campbell and Kahn at the University of Michigan (reported in *The People Elect a President* 1952).¹⁴ Since that time, surveys have been used to study political culture (Almond and Verba 1963), political participation (Verba and Nie 1972), political socialization (Jennings and Niemi 1974; 1981), political parties (Eldersveld 1964), and many other topics. There are many national election studies,¹⁵ single-point-in-time cross-national studies, and ongoing cross national studies such as the World Values Surveys, the International Social Survey Programme, the Comparative Study of Electoral Systems, the Pew Global Attitudes Survey, Gallup's Voice of the People, and many others.¹⁶

Surveys provided political scientists with the first opportunities to collect microdata on people's attitudes, beliefs, and behaviors, and they continue to be an important method of data collection. Moreover, they have become much more useful because there are better instruments, better designs, greater comparability over time, and greater comparability over space.

We now know a great deal about how to write better questions (Bradburn, Sudman, and Wansink 2004; Tourangeau, Kips, and Rasinski 2000). We also know how to ask questions about political information, political attitudes (liberalism-conservatism, political issues and agendas, racial attitudes, patriotism, trust, etc.), political and economic values (tolerance, democratic values, civil liberties, economic equality, capitalist values, the Protestant ethic), identity (partisanship, ethnic identity, racial identity), political emotions, candidate and group traits, the use of the news media, and political participation and other behaviors (p. 1018) Robinson, Shaver, and Wrightsman 1993; Abdelal et al. 2009; Marcus and MacKuen 1993; Price and Zaller 1993). We take seriously issues of reliability and validity, and we have better techniques for analyzing batteries of questions (Jackman, in Brady and Collier 2008). We have better techniques for making interpersonal comparisons across places and times using "vignettes" (King et al. 2003). We know a great deal about the problem of social desirability bias and other mechanisms affecting responses and memory, and we have learned a lot about interpreting and improving responses on racial attitudes (Berinsky 2002; Krysan and Couper 2003), religious attendance (Presser and Stinson 1998), voting (Beilli et al. 1999), and welfare spell lengths (Luks and Brady 2003). We have innovative methods for getting at racism such as the "list experiment" (Kuklinski, Cobb, and Gilens 1997) and the Implicit Association Test and its close relatives (Greenwald, McGhee, and Schwartz 1998; Fazio and Olson 2003; but see Arkes and Tetlock 2004). Still, it is frustrating that we cannot reliably (and perhaps not even validly) identify the true extent of racism, the role of emotions, and the extent of social desirability bias and other mechanisms which distort response.

We also have much better designs for surveys along three distinct dimensions. First, experiments embedded within surveys (Sniderman and Grob 1996) have been used to probe general population attitudes and opinions in novel and interesting ways. Second, panel studies in which the same people are interviewed repeatedly and rolling

cross-sections in which true daily random samples are repeated day after day over a long period of time such as a political campaign have been used to study change over time (Brady and Johnston 2006). Researchers now combine these approaches into rolling cross-sectional panel designs. Third, researchers now combine experiments, panels, rolling cross-sections, and other designs across the modes of in-person, telephone, Internet, and self-administered mail questionnaires (Johnston, in Brady and Collier 2008). By repeating questions across surveys, across countries, and over time and by building up collections of data, we are able to make interesting descriptive comparisons (e.g. with respect to partisan identification, ethnic and racial attitudes, or concern for economic issues across countries and over time) and we can often make some causal inferences by using differences across countries or over time (or changes in differences) to pin down (technically, “to identify”) causes.

2.3.2 Legislative and Judicial Voting Data

Another area where there have been great advances is in the collection and scaling of voting data which has moved considerably beyond the classic books by Rice (1928) and MacRae (1970). Keith Poole (in Brady and Collier 2008) and his collaborator Harold Rosenthal have created the definitive version of American congressional roll-call data which are available (and constantly updated) on the VoteView¹⁷ website, and they have produced the classic interpretation of these data (Poole and Rosenthal 1997) using their widely used “Nominat” method for analyzing legislative votes (p. 1019) (Poole and Rosenthal 1985).¹⁸ Data from around the world are archived and disseminated on the VoteWorld website,¹⁹ and the VoteWorld project aims to create “open source” software standards for formatting and archiving roll-call data-sets. Judicial data, including votes on every Supreme Court case back to 1953, are available at the “Oyez” website (<www.oyez.org>). The site also includes Martin–Quinn measures of judicial ideology (Martin and Quinn 2007) based upon a dynamic Bayesian ideal-point model.

2.3.3 Events and Media Data

Since the pioneering work of Gosnell and Rice in the 1920s, political scientists have used aggregate voting data, survey data, and legislative and judicial voting data. “Content analysis” of the media began in the 1930s and 1940s with the work of Harold Lasswell (1941), but the study of political events has been a more recent innovation. At least five types of events are now regularly studied in political science: political campaign events, agenda-setting actions such as hearings or the introduction of legislation, contentious (protest) events, civil wars and civil strife, and international interactions and transactions. The recording of events can be traced back to the efforts forty years ago to create cross-national data books (Banks and Textor 1963; Russett et al. 1964; Banks 1971; Taylor and Hudson 1972). These compendia contained aggregate data on countries over time on variables such as “domestic group violence deaths” (Russett et al. 1964) or “coups d’etat, assassinations, general strikes, or riots” (Banks 1971). In the past forty years, much more refined data have been produced. Some important publications in the development of specific types of events data are Feierabend and Feierabend (1966) for internal conflict, Gurr (1968; 1970a) for protest and civil strife, Azar (1970) for international interactions, Ben-Dak and Azar (1972) with a symposium on Arab–Israeli conflict using mostly events data, Singer and Small (1972) on war and alliances, Tilly, Tilly, and Tilly (1975) for protest (“contentious politics”), Bartels (1987) on primaries as campaign events, Allsop and Weisberg (1988) on campaign events and party identification, Holbrook (1996) and Shaw (1999) on campaign events, and Baumgartner, Jones, and MacLeod (1998) on agenda-setting actions.

Events data are most useful when many different features of the events are coded: For example, it is helpful to know the date (and length) of the protest, its size, its location, the characteristics of its participants, its targets, and any violence associated with it. Or it is helpful to know all participants in an international event (such as a war), its date, its duration, the sizes and economic characteristics of the participants, whether they have other linkages with one another, and the outcome of the interaction. Knowing these things requires having good sources, and events have typically been coded from newspapers, yearbooks, police records, and historical records, which can be problematic. From the very beginning (Azar et al. 1972) (p. 1020) and continuing up to the present day (Schrodt and Gerner 1994; Woolley 2000; Oliver and Maney 2000), there has been a concern with the representativeness and accuracy of these sources. But with good event data, sophisticated event history methods can be used (Golub 2008), and important works can be completed such as Beissinger’s (2002) analysis of how protest events contributed to the fall of the Soviet Union, Freeman and Goldstein’s (1990) analysis of reciprocity in the international system, and Shaw’s (2006) analysis of American political campaigns.

2.3.4 Other Measurement Methods

In addition to the improvement of older methods of data collection, some fascinating new methods are available. Game theorists have added to our stock of measurement techniques with simple games that tell us about people's willingness to free-ride, their trust in one another, and many other features of interactive situations (Camerer 2003). Two new methods inform us about the neurological bases of attitudes: fMRI techniques provide real-time brain scans that indicate what parts of the brain are activated when subjects are presented with stimuli (Lieberman, Schreiber, and Ochsner 2003; Phelps and Thomas 2003). ERP or event-related potential methods (Morris et al. 2003) provide another, somewhat less difficult to obtain, measure of brain activity.

2.3.5 Textual data

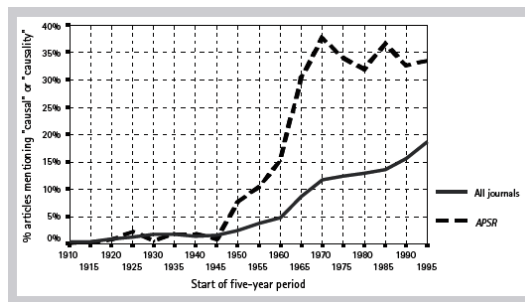


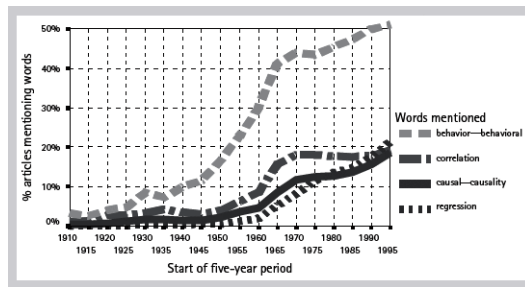
Fig. 48.2 Growth of causal thinking

One of the most exciting frontiers in political science research at the present time is developing ways to code and use the enormous bodies of textual data that are now available on the web and through other sources. Human coding of text into categories through “content analysis” has been used to analyze textual data since the 1930s, but the very large bodies of data that are now available and the increasing capabilities of computers make it both feasible and necessary to automate this process (King and Lowe 2003). The possibilities are extraordinary. We have large collections of newspaper texts and books (including biographies) that have been scanned. We also have judicial opinions and briefs, congressional debate and testimony, political advertisements in text and video form, nightly news in text and video, party platforms, and, of course, the entire contents of the web. The challenge is to find ways to code these data and to link them with political surveys, judicial or legislative votes, or events. To take just one example, as part of its Social, Political, and Economic Event Database project, the Cline Center at the University of Illinois has assembled in a machine readable form over 30 million reports dating from the Second World War from the *New York Times*, the *Wall Street Journal*, the BBC's Summary of World Broadcasts, and the Foreign Broadcast Information Service reports, and it intends to use computers to code these data into over 200 event categories. The results will provide researchers with a worldwide, sixty-year record of events in the world. (p. 1021)

3 Causal Thinking

Although not all of modern political science is about causation, Figure 48.2 shows that between 1990 and 1999, about one-third of the articles in the *American Political Science Review* included the words “causal” or “causality,” and 17 percent of the political science journal articles in JSTOR for this period mentioned them. Moreover, the mentions of these terms grew rapidly from less than 2 percent of the JSTOR articles from 1910 to 1950 to an increasing proportion from 1950 onwards, with the *APSR* leading the way.

In our introductory chapter to the *Handbook of Political Methodology*, we explored the roots of this dramatic increase in mentions of “causal” or “causality.” Our discussion of the rise of “causal thinking” was, in many ways, simply a “toy example,” meant to show the difficulties of *explaining* anything—even something as prosaic as the rise of causal thinking within political science. In demonstrating these difficulties, we hoped that the example illustrated the problems of making defensible causal claims.



[Click to view larger](#)

Fig. 48.3 Growth of mentions of words related to causal thinking in political science

We explored three possible causes of the increased emphasis on causality in political science. Two potential causes involved the availability of new tools such as correlation (Pearson 1909) or regression analysis (Pearson 1896; Yule 1907) that might have fostered causal thinking because they made it easier for scholars to determine causality. A third potential cause, a commitment to behavioralism, might have placed more emphasis within the discipline upon making causal inferences. The growth in the mention of words representing each of these possible explanations (“correlation,” “regression,” and “behavioralism”) is plotted in Figure 48.3, which demonstrates that they grew in tandem with causal thinking as depicted in Figure 48.2. There may be other plausible explanations for the growth of causal thinking, but these three provide (p. 1022) us with a trio of interesting possibilities. Indeed, these categories of explanation—new inventions and new values—crop up again and again in social science. We concluded, with some trepidation given the incompleteness of our analysis, that values and inventions both helped to explain the rise of “causal thinking” in political science. The behavioral movement furthered “scientific values” like causal thinking, while regression analysis (but not correlation) provided an invention that seemingly provided political scientists with estimates of causal effects with minimal fuss and bother.

Despite this long-held presumption that regression could uncover causal relationships, the experiences of the last thirty years have undermined the belief that regression is the philosopher’s stone that can turn base observational studies into gold-standard experimental studies. Doubts have grown about inferences made from simple regressions, and increasingly sophisticated methods have been developed to improve regression analysis and causal inference. Researchers now know that most regression equations simply provide a multivariate summary of the data—at best a *descriptive inference*—not a sure-fire *causal inference* about them (King, Keohane, and Verba 1994) because the conditions for justifying a causal interpretation of regression coefficients are not met. Although establishing the Humean conditions of constant conjunction and temporal precedence with regression-like methods often takes pride of place when people use these methods, we now know that they seldom deliver a reliable causal inference. Rather regressions are often more usefully thought of as ways to describe complex data-sets by estimating parameters that summarize important things about the data. For example, Auto-Regressive Integrated Moving Average (ARIMA) models can quickly tell us a lot about a time series through the standard “p,d,q” parameters, which are the order of the autoregression (p), the level of differencing (d) required for stationarity, and the order of the moving average (p. 1023) component (q). And a graph of a hazard rate over time derived from an events history model reveals at a glance important facts about the ending of wars or the dissolution of coalition governments. Descriptive inference is often underrated in the social sciences (although survey methodologists proudly focus on this problem), but even more worrisome is the tendency for social scientists to mistakenly assume that a descriptive inference is a valid causal inference. Most regression analyses in the social sciences are probably useful descriptions of the relationships among various variables, but usually they cannot properly be used for causal inferences because they omit variables, fail to deal with selection bias and endogeneity, and lack theoretical grounding. In short, they typically do not establish causal relationships (Freedman 1997).

3.1 Nature of Causality

Why does regression typically fail to provide valid causal inferences? The basic problem is that establishing a causal relationship requires much more than estimating conditional expectations using regression. Brady (2008; and this volume) presents an overview of the challenges of making causal inferences by describing four perspectives on them. The *neo-Humean regularity approach* focuses on “lawlike” constant conjunction and temporal antecedence, and many statistical methods—preeminently regression analysis—are designed to provide

just the kind of information to satisfy the requirements of the Humean model. Regression analysis can be used to determine whether a dependent variable is still correlated (“constantly conjoined”) with an independent variable when other plausible causes of the dependent variable are held constant by being included in the regression, and time series regressions can look for temporal antecedence by regressing a dependent variable on lagged independent variables. But regression can fail to make valid inferences because researchers typically cannot consider all the plausible alternative explanations that can confound their analysis.

The *counterfactual approach* to causation asks what would have happened had a putative cause not occurred in the most similar possible world without the cause. It requires either finding a similar situation in which the cause is not present or imagining what such a situation would be like. But researchers seldom have a perfect counterfactual situation.

The *manipulation approach* asks what happens when we actively manipulate the cause—does it lead to the putative effect? This approach is what characterizes most laboratory experiments in the natural sciences. Some factors (temperature, pressure, chemicals, DNA, and so forth) are manipulated, and the scientist observes the results. Typically there is no “control” (or “counterfactual”) condition because it is thought to be obvious that the manipulation of the factors produces the result. But manipulations of some factors (e.g. party identification, race, or gender) is hard, if not impossible, in the social sciences, and the multitude of confounding factors that can cause a result makes it hard to determine the impact of a manipulation.

(p. 1024) Finally, the *mechanism and capacities approach* asks what detailed steps lead from the cause to the effect and what underlying theoretical perspective might explain the causal relationship. When done well, this approach helps to pin down the nature of the cause or treatment, it clarifies the intervening or contextual variables needed for the result to occur, and it helps to generalize to other situations. But social science researchers often have both too many loosely worked-out explanations and too few that are adequately worked out.

These four approaches provide a framework for understanding two successive developments in the study of causality in political science that have changed our understanding of it. First, there has been the realization that regression approaches that (seemingly) satisfy the Humean conditions for causation may suffer from problems of omitted variables, endogeneity, and selection bias. Second, there has been the recognition that researchers must focus on causes that can be manipulated; they must be able to make counterfactual statements about what would occur without the putative cause; and they must be able to describe an explicit causal mechanism (Brady 2008; and this volume). These developments have been accompanied by a growing emphasis on experimentation as an excellent way to deal with many (although not all) of these concerns. The change in our understanding of how to make causal inferences is captured by considering Figure 48.4, which displays the number of *APSR* articles in each decade that mentioned some of these things—endogeneity, selection bias, counterfactuals, and experiments. (During this period, there were roughly 500 *APSR* articles each decade.)

Before 1960, researchers adduced causal inferences from regressions with very little self-consciousness about their limitations. There are essentially no mentions of any of these concerns (omitted variables, endogeneity, selection bias, manipulations, counterfactuals, or mechanisms) before 1960. At that point, a nearly exponential growth begins in the use of the terms “endogenous” or “endogeneity.” This growth corresponds with the realization that statistical regression models with endogenous regressors are incorrectly estimated by standard (OLS) regression methods. Many of the mentions of endogeneity in these articles are associated with the spread of structural equation modeling methods with their “systems of equations” in which variables appear on both the left- and right-hand sides of equations—thus making it impossible to accept the standard regression assumption of no correlation between the right-hand-side variables and the error term. The growth in the use of the terms “endogenous” or “endogeneity” increases dramatically in 1970 and 1980; and by 1990, 101 articles mention the term and in the five-year period from 2000 to 2004, 75 articles mention it.²⁰

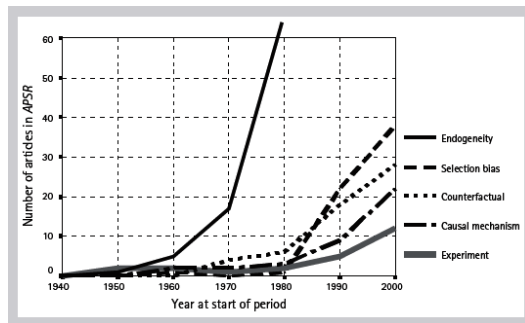


Fig. 48.4. Number of articles in APSR dealing with specific causal issues

However, regression methods were “saved” for a long time by the presumption that instrumental variables could be found (typically within the structural equation model itself) that would provide a statistical fix to the problem of endogeneity (Jackson 2008). It was only when researchers began to realize that instrumental variables (p. 1025) are hard to justify (that is, statistical identification is hard to come by), and that a supposed “treatment” in a simple regression equation (such as exposure to an advertisement in an equation estimating candidate preference) might be assigned nonrandomly through a “selection” process, that worries arose that regression methods, including structural equation modeling, might not solve the causal inference problem (Achen 1986).

Concern with the problem of selection bias,²¹ for example, rises dramatically in the 1990s (see Figure 48.4), which suggests that researchers began to realize that even without the problem of simultaneity, regression models might be incorrectly specified and incorrectly estimated without instruments or some other method of identification. An increase in mentions of the words “counterfactual” and “mechanism” in the 1990s suggests an even more detailed understanding of the problems with causal inference, and the increasing number of experiments described in the APSR suggests a modest move toward solving causal inference problems using that method. But there are practically no mentions of the words “manipulation” in the APSR or in other journals, suggesting that political scientists have not yet made sufficient use of this approach to causal inference.

In addition to an awakening to the complexity of causal inference involving a single putative cause and effect, political science has also become more attuned to the complexity of causal inference involving many causal factors. Since the beginning of the use of statistical methods within political science, there has been an (p. 1026) understanding that causes are often probabilistic, and not deterministic, but there has been surprisingly little discussion of what this means. Perhaps the most important theoretical question is how a deterministic world might nevertheless lead to a probabilistic social science.²² One answer is that we observe only part of a complex web of causation. Brady (2008; and this volume), for example, discusses the INUS model of causation which gets beyond simple necessary or sufficient conditions for an effect by arguing that often there are different sufficient pathways (but no pathway is strictly necessary) to causation—each pathway consisting of an *insufficient* but *nonredundant* part of an *unnecessary* but *sufficient* (INUS) condition for the effect. For example, revolutions may occur because of an interaction between state breakdown and peasant revolts (Skocpol 1979) or they may occur because of an interaction between indecisive secular repressive regimes and growing religious fundamentalism (Arjomand 1986). Hence, none of these four conditions (state breakdown, peasant revolts, indecisive repressive regimes, or religious fundamentalism) is either necessary or sufficient for a revolution, but there are two pairs of sufficient conditions. In most situations, we would expect to have multiple causal paths, and Brady shows that if we only observe some of the factors which affect outcomes, then the world will appear probabilistic. Sekhon (2004) concludes from this that the use of Mill’s methods and other deterministic approaches is inherently flawed. We would not go so far. In fact, the work on necessary and sufficient conditions (e.g. Goertz and Starr 2002; Goertz and Levy 2007), conjunctural causation (Collier and Collier 1991; Pierson 2000), and Qualitative Comparative Analysis (Ragin 1987) has made political scientists more aware of the complexity of social phenomena (Achen 2002) and the need for exploring multiple pathways and complex interactions.

3.2 Statistical Methods for Establishing Causality: Quantitative Tools for Causal and Descriptive Inference

As noted earlier, regression analysis, much more than correlation analysis, provides a seductive technology for exploring causality. Its inherent asymmetry with a dependent variable that is a function of a number of independent

variables lends itself to discussions of causes (independent variables) and effect (dependent variable) whereas correlation (even partial correlation) analysis is essentially symmetrical. The path analysis generalizations of regression make it even more attractive because of the widespread use of “path diagrams” with directed pathways that look just like causal arrows between variables. And social scientists and statisticians provide even more credibility for the method by their proofs that under certain conditions regression coefficients will be an unbiased estimate of the impact of the independent variables on the dependent variables (Simon 1954; Blalock 1964). (p. 1027)

Table 48.1. Results of regressing whether “causal thinking” was mentioned on mentions of potential explanatory factors for 1970–9 (all political science journal articles in JSTOR)

Independent variables	Regression coefficient (standard error)	
	One	Two
Behavior	.122 (.006)**	.110 (.006)**
Regression	.169 (.010)**	.061 (.021)*
Correlation	.157 (.008)**	.150 (.015)**
Behavior X regression		.135 (.022)**
Behavior X correlation		.004 (.017)
Regression X correlation		.027 (.021)
Constant	.022 (.008)**	.028 (.004)**
R ² /N	.149/ 12,305	.152/ 12,305

Notes: (***) Significant at .001 level;

(**) significant at .01 level;

* significant at .05 level.

Finally, regression analysis provides the striking capacity to predict that if there is a one-unit change in some independent variable, then there will be a change in the dependent variable equal to the value of the independent variable’s regression coefficient. In short, regression analysis seems to deliver a great deal whereas correlation analysis appears to deliver much less, so that it seems likely that regression analysis contributed much more to the emphasis on causal thinking than correlations.

We can illustrate these points and test our theories about the rise of causal thinking with some data from JSTOR. The classic regression approach to causality suggests estimating a simple regression equation such as the following for cross-sectional data on all political science articles in JSTOR between 1970 and 1979. For each article, we score a mention of either “causality or causal” as a one and no mention of these terms as a zero. We then regress these zero–one values of the “dependent variable” on zero–one values for “independent variables” measuring whether or not the article mentioned “regression,” “correlation,” or “behavioralism.” When we do this, we get the results in column one in Table 48.1. If we use the causal interpretation of regression analysis to interpret these results, we might conclude that all three factors led to the emphasis on “causal thinking” in political science because each coefficient is substantively large and statistically highly significant. But this interpretation ignores a multitude of problems.

Given the INUS model of causation which emphasizes the complexity of necessary and sufficient conditions, we might suspect that there is some interaction among these variables, so we should include interactions between each pair of variables. These interactions require that both concepts be present in the article (p. 1028) so that a “Regression X Correlation” interaction requires that both regression and correlation are mentioned. The results from estimating this model are in column two of the table. Interestingly, of the interaction terms, only the “behavior X regression” interaction is significant, suggesting that the combination of the behavioral revolution and the development of regression analysis helps “explain” the prevalence of causal thinking in political science. (The three-way interaction is not reported and is statistically insignificant.) Descriptively this result is certainly correct—it appears that a mention of behavioralism alone increases the probability of “causal thinking” in an article by about 11 percent, the mention of regression increases the probability by about 6 percent, the mention of correlation increases the probability by about 15 percent, and the mention of both behavioralism and regression together further increases the probability of causal thinking by about 13.5 percent.

But are these causal effects? This analysis is immediately open to the standard criticisms of the regression approach when it is used to infer causation: Maybe some other factor (or factors) causes these measures (especially “behavioral,” “regression,” and “causality”) to cohere during this period. Maybe these are all spurious relationships that appear to be statistically significant because the true cause is omitted from the equation. Or maybe causality goes both ways and all these variables are endogenous. Perhaps “causal thinking” causes mentions of the words “behavioral or behavior” and “regression” and “correlation.”

Although the problem of spurious relationships challenged the regression approach from the very beginning (see Yule 1907), many people (including apparently Yule) thought that it could be overcome by simply adding enough variables to cover all potential causes. The endogeneity problem posed a greater challenge, which only became apparent to political scientists in the 1970s. If all variables are endogenous, then there is a serious identification problem with cross-sectional data that cannot be overcome no matter how much data are collected. For example, in the bivariate case where “causal thinking” may influence “behavioralism” as well as “behavioralism” influencing “causal thinking,” the researcher only observes a single correlation, which cannot produce the two distinctive coefficients representing the impact of “behavioralism” on “causal thinking” and the impact of “causal thinking” on “behavioralism.”

The technical solution to this problem is the use of “instrumental variables” known to be exogenous and known to be correlated with the included endogenous variables, but the search for instruments proved elusive in many situations. The literature on this topic is now vast (Jackson 2008), and it includes tests for exogeneity, dealing with weak instruments, and the theoretical question of when structural equations must be used to understand a phenomenon. Heckman (2008, 5), for example, argues that the statistical models of causality (the Neyman–Rubin–Holland model discussed in more detail below and in Sekhon 2008) are incomplete because “[t]hey do not allow for simultaneity in choices of outcomes and treatment that are at the heart of game theory and models of social interactions and contagion.” As a result, statistical models produce parameters that are of limited use.

(p. 1029) This perspective suggests that the ongoing generalization of regression models is not just, as some proponents of the experimental model seem to suggest (Gerber, Green, and Kaplan 2004), a fruitless attempt to overcome insoluble problems with statistical tricks. Rather, these more sophisticated methods are needed in conjunction with a better eye toward finding defensible instruments and natural experiments. They are also needed because they often make us think harder about the nature of our data and the types of problems that we must overcome before making inferences. For example, the synthesis of factor analysis and causal modeling that produced what became known as LISREL, covariance structure, or structural equation models has increased our understanding of both measurement and causality. These approaches use factor analysis types of models to develop measures of latent concepts that are then combined with causal models of the underlying latent concepts (Bollen, Rabe-Hesketh, and Skrondal 2008). These techniques have been important at two levels. At one level, they simply provide a way to estimate more complicated statistical models that take into account both causal and measurement issues. At another level, partly through the vivid process of preparing “LISREL diagrams,” they provide a metaphor for understanding the relationships between concepts and their measurements, latent variables and causation, and the process of going from theory to empirical estimation. Unfortunately, the models have also sometimes led to baroque modeling adventures and a reliance on linearity and additivity that at the same time complicates and simplifies things too much. Perhaps the biggest problem is the reliance upon “identification” conditions that often require heroic assumptions about instruments.

One way out of the instrumental variables problem is to use time series data. At the very least, time series give us a chance to see whether a putative cause “jumps” before a supposed effect. We can also consider values of variables that occur earlier in time to be “predetermined”—not quite exogenous but not endogenous either. Time series methods such as simple time series regressions, ARIMA models, vector autoregression (VAR) models, and unit root and error correction models (ECM) take this approach (Pevehouse and Brozek 2008). The literature faces two tricky problems. One is the complex but tractable difficulty of autocorrelation, which typically means that time series have less information in them per observation than cross-sectional data and which suggest that some variables have been omitted from the specification (Beck and Katz 1996). The second is the more pernicious problem of unit roots and commonly trending (co-integrated) data which can lead to nonsense correlations. In effect, in time series data, time is almost always an “omitted” variable that can lead to spurious relationships which cannot be easily (or sensibly) disentangled by simply adding time to the regression. Hence the special adaptation of methods designed for these data.

For our exploration of the rise of causal thinking, we can estimate a time series autoregressive model for eighteen five-year periods from 1910 to 1999. The model regresses the proportion of articles mentioning “causal thinking” on the lagged proportions mentioning the words “behavioral or behavior,” “regression,” or “correlation.” Table 48.2 shows that mentions of “correlation” do not seem to matter (the (p. 1030)

Table 48.2. Mentions of “causal thinking” for five-year periods regressed on mentions of “behavioral or behavior,” “regression,” and “correlation” for five-year periods for 1910–99

Independent variables lagged	Regression coefficients (standard errors)
behavior	.283 (.065)***
Regression	.372 (.098)**
Correlation	–.159 (.174)
AR (1)	.276 (.342)
Constant	–.002 (.005)
N	17 (one dropped for lags)

Note: Significant: .05 (*), .01 (**), .001 (***).

coefficient is negative and the standard error is bigger than the coefficient), but mentions of “regression” or “behavioralism” are substantively large and statistically significant. (Also note that the autoregressive parameter is statistically insignificant.) These results provide further evidence that it might have been the combination of behavioralism and regression that led to an increase in causal thinking in political science.

A time series often throws away lots of cross-sectional data that might be useful in making inferences. Time series Cross-Sectional (TSCS) methods (Beck 2008) and event history models (Golub 2008) try to remedy this problem by using both sorts of information together. These techniques provide both fixes for problems and insights into subtle causal questions. They provide some elegant fixes for omitted variables problems because TSCS methods can use fixed unit effects to control for factors that are constant over time for a given unit, or fixed period effects to control for factors that are constant over units for a given time period. They also raise a host of important methodological questions such as distinguishing the degree of causal heterogeneity in units versus the amount of learning (“duration or state dependence”). This issue is important in many different research areas. In the study of welfare, solutions to these technical problems have implications for welfare policy. If recipients have different length welfare spells because they are heterogeneous (e.g. they might have different skill levels), then liberal job-training programs make sense; but if welfare spell lengths differ because being on welfare leads to welfare dependency (duration dependence), then conservatives are right about the need for time limits and “back-to-work” policies. In

the study of party identification, it makes a difference whether people are subject to heterogeneous exogenous forces leading to persistence in identification or whether they are subject to becoming more partisan simply as a result of a past experience as partisans (Bartels et al. 2007; Box-Steffensmeier and Smith 1996).

Other “regression” methods also increase our understanding of measurement and modeling issues. Discrete choice modeling (Glasgow and Alvarez 2008) deals with (p. 1031) dichotomous variables, ordered choices, and unordered choices. Ecological regression and its more sophisticated cousins (King 1997; Cho and Manski 2008) can be used whenever scholars are interested in the behavior of individuals but the data are aggregated at the precinct or census tract level (Cho and Manski 2008). Spatial analysis (Franzese and Hays 2008) and hierarchical modeling (Jones 2008; Steenbergen and Jones 2002) take into account the spatial and logical structure of data.

Consider what spatial analysis tells us about causal thinking. “Spatial interdependence” between units of analysis can be thought of as a nuisance just like autocorrelation in time series, but it can also be thought of as a sign that we must explain why the behavior of nearby units might be affected by similar unobserved variables. Spatial interdependence can be represented by a symmetric weighting matrix for the units of observation whose elements reflect the relative connectivity between unit i and unit j . By including this matrix in estimation in much the same way that we include lagged values of the dependent variable in time series, we can discover the impact of different forms of interdependence. But we are still left with questions about why this interdependence exists.

Hierarchical models also challenge our causal thinking. The classic use of multilevel models is in educational research where students are in classrooms situated in schools, which are in turn in school districts that are in states. Students may be affected by factors at all these levels—their own individual characteristics, the teaching in their classrooms, the culture of their schools, the policies of their school districts, and so forth. Many political problems have a similar structure. If we are to understand public opinion, shouldn't we consider more than people's individual characteristics? Shouldn't we consider how they are affected by their organizational affiliations, their communities, their political jurisdictions, and so forth? Don't we expect that two evangelicals with the same individual characteristics will behave differently if one is surrounded by other evangelicals while the other is surrounded by mainstream Protestants, Catholics, or nonbelievers?

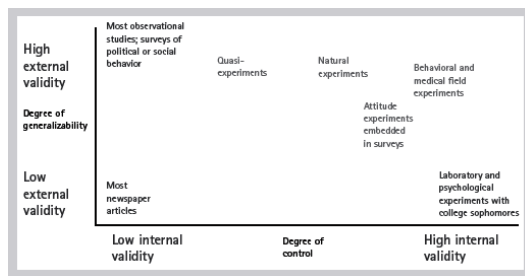
In addition to these innovations in statistical methods, there have been important innovations in statistical estimation methods. The classic book by Eric Hanushek and John Jackson (1977) introduced many political scientists to a much broader set of statistical methods. Gary King (1998) made R. A. Fisher's maximum likelihood methods popular in political science. More recently, Bayesian estimation methods (Gill 2007; Martin 2008) have vastly increased our ability to estimate complex models. Before the 1990s, many researchers could write down a plausible model and the likelihood function for what they were studying, but the model presented insuperable estimation problems. Bayesian estimation was often even more daunting because it required not only the evaluation of likelihoods, but the evaluation of posterior distributions that combined likelihoods and prior distributions. In the 1990s, the combination of Bayesian statistics, Markov Chain Monte Carlo (MCMC) methods, and powerful computers provided a technology for overcoming these problems. These methods make it possible to simulate even very complex distributions and to obtain estimates of previously intractable models. Recently, political scientists have also been contributing to the development of computational methods for function (p. 1032) maximization (Sekhon and Mebane 1998) and matching methods (Sekhon and Diamond 2005).

3.3 The Neyman–Rubin–Holland Model and Experimentation

In the last twenty years the spread of the Neyman–Rubin–Holland model of causal inference (Neyman 1990; Rubin 1974; Holland 1986) has revolutionized the teaching of causal inference (Brady 2008; Sekhon 2008) by emphasizing the importance of counterfactuals, manipulations, and above all, experimentation. This model also makes one important aspect of testing for a causal relationship a probabilistic one: whether or not the probability of the effect goes up when the cause is present.²³

Experiments have become the gold standard for establishing causality because of their strong claim to making valid inferences—to what Donald Campbell (Campbell and Stanley 1966; Cook and Campbell 1979) called their “internal validity.” Combining R. A. Fisher's notion of randomized experiment (1925) with the Neyman–Rubin model (Neyman 1923; Rubin 1974; 1978; Holland, 1986) provides a recipe for valid causal inference as long as several assumptions are met. At least one of these, the Stable Unit Treatment Value Assumption (SUTVA), is not trivial,²⁴

but some of the other assumptions are relatively innocuous so that when an experiment can be done, the burden of good inference is to properly implement the experiment.



Click to view larger

Fig. 48.5. External-internal validity trade-off

The number of experiments in political science has increased dramatically in the last thirty-five years (Morton and Williams 2008) because of their power for making causal inferences.²⁵ At the same time, experiments, especially those in highly controlled situations such as college classrooms, have an Achilles heel—their lack of generalizability across people, places, and things—what Donald Campbell (Campbell and Stanley 1966; Cook and Campbell 1979) labeled “external validity.” One of the challenges to modern political science is to find ways to undertake experiments in more general locations, with more representative populations, and with more realistic conditions. Another approach is to use field experiments and natural experiments to overcome the external validity limitations of laboratory experiments (Gerber and Green 2008). Despite early skepticism about what could be done with experiments, (p. 1033) social scientists are increasingly finding ways to experiment in areas such as criminal justice, the provision of social welfare, schooling, and even politics. But “there remain important domains of political science that lie beyond the reach of randomized experimentation” (Gerber and Green 2008, 361). Moreover, more realistic experiments also must cope with the real-world problems of “noncompliance” and “attrition.” Noncompliance occurs when medical subjects do not take the medicines they are assigned or citizens do not actually get the phone calls that were supposed to encourage their participation in politics. Attrition is a problem for experiments when people are more likely to be “lost” in one condition (typically, but not always, the control condition) than another. These difficulties make it hard to estimate causal impacts (if someone does not take the medicine, it is hard to estimate its impact), but they also present opportunities for making inferences about real-world situations (in the real world, people often do not take their medicine).

Figure 48.5 depicts the basic “external validity” versus “internal validity” inherent in experimentation versus observational studies. Observational studies, especially those using surveys, can make a strong claim to external validity—in fact that is the preeminent goal of the random sample. But it is often very hard to draw valid inferences from observational studies no matter how the sample is drawn. Randomized laboratory experiments can claim substantial internal validity, but often at the cost of external validity. In between are quasi-experiments, natural experiments, field experiments, and attitude experiments embedded in surveys.

3.4 Qualitative Tools for Causal Inference

What, then, is the role of qualitative research? When causal inference was defined solely as satisfying the neo-Humean regularity and antecedence conditions, then causal inference could be compressed into the tasks of showing that the cause (p. 1034) preceded the effect and that the probability of the effect increased with the presence (or strength) of the putative cause once all the plausible confounders had been controlled. The ingredients for this recipe for establishing causation were: (1) large numbers of observations to deal with random variation; (2) quantitative measures of causes, effects, and especially “control” variables to rule out confounding explanations; and (3) a computerized regression package to run the data. When randomization was added as a condition for inferring causality, then experiments seemed de rigueur. But when establishing causal relations is broken apart into the requirements for counterfactuals, manipulations, mechanisms, and necessary and sufficient conditions, other, more focused strategies seem plausible in many situations.

David Freedman (2008, 312), for example, has argued that “substantial progress also derives from informal reasoning and qualitative insights” even though he has written extensively on the Neyman–Rubin–Holland (NRH)

framework and he believes that it should be employed whenever possible because it sets the gold standard for causal inference. He suggests that another strategy relying upon “causal process observations” (CPOs) can be useful as a complement to the NRH framework (Brady and Collier 2004). CPOs rely upon detailed observations of situations to look for hints and signs that one or another causal process or mechanism might be at work, to look for cases where manipulations seem to have produced some effect, and to be open to natural experiments (where we observe two similar situations, one with and the other without the putative cause). Not just any cases or situations will do, but some provide inferential leverage because of their special qualities.

Thus Edward Jenner used fifteen case studies (involving thirty-one people) to conclude that cowpox inoculations could protect people against smallpox. Ignaz Semmelweis used cases to rule out “atmospheric, cosmic, telluric changes” as the causes for puerperal (also called childbed) fever, and he used the death of his colleague by “cadaveric particles” to identify the disease’s mode of transmission. Alexander Fleming observed an anomaly in a bacterial culture in his laboratory that led to the discovery of penicillin. John Snow was led to understand the method by which cholera was transmitted by thinking about the deaths of a poor soul in London who next occupied the same room as a newly arrived and cholera-infected seaman and the death of a lady who had drunk from the cholera-infected “Broad Street Pump” because she liked the taste of the water, even though she lived far from the pump.

This careful use of case study material has been codified in the “process tracing” of Alexander George and Andrew Bennett (George and Bennett 2005; Bennett 2008) and in the CPOs of Brady and Collier (2004). Process tracing is an analytic procedure through which scholars make fine-grained observations to test ideas about causal mechanisms and causal sequences. Bennett (2008) argues that the logic of process tracing has important features in common with Bayesian analysis: It requires clear prior expectations linked to the theory under investigation, examines highly detailed evidence relevant to those expectations, and then considers appropriate revisions to the theory in light of observed evidence. With process tracing, the movement from (p. 1035) theoretical expectations to evidence takes diverse forms, and Bennett reviews these alternatives and illustrates them with numerous examples.

A related, but somewhat more theoretically inclined approach, is to try to understand the mechanisms which are the underlying “cogs and wheels,” which connect cause and the effect (Hedstrom 2008). The mechanism, for example, which explains how vaccinations work to provide immunity from an illness is the interaction between a weakened form of a virus and the body’s immune system which confers longtime immunity. In social science, the rise in a candidate’s popularity after an advertisement might be explained by a psychological process that works on a cognitive or emotional level to process messages in the advertisement. Various authors have inventoried stylized mechanisms that underlie social phenomena (Hedstrom 2008; Elster 1998; Mahoney 2001), and game theorists and formal modelers are certainly at the ready with mechanisms to explain almost any phenomena.

For example, Levy (2008)—in a discussion of counterfactuals and case studies—argues that game theory is one (but not the only) approach that provides clear counterfactuals and mechanisms for understanding social phenomena. A game explicitly models all of the actors’ options including those possibilities that are not chosen. Game theory assumes that rational actors will choose an equilibrium path through the extensive form of the game, and all other routes are considered “off the equilibrium path”—counterfactual roads not taken. Levy argues that any counterfactual argument requires a detailed and explicit description of the alternative antecedent (i.e. the cause that did not occur in the counterfactual world), which is plausible and involves a minimal rewrite of history, and he suggests that one of the strengths of game theory is its explicitness about alternatives. Levy also argues that any counter-factual argument requires some evidence that the alternative antecedent would have actually led to a world in which the outcome is different from what we observe with the actual antecedent. With these ingredients useful causal arguments can be made.

4 The Future

4.1 Causal Inference and Interpretation: Bridging Alternative Approaches

Political methodology has become much more sophisticated and nuanced in the past thirty years, and these developments have met some of the challenges from critics of behavioralism. Nevertheless, there are still

competing perspectives on political science, and the following data suggest that one of the cleavages is between those who focus on explanation and hypothesis testing and those who are interested in interpretation and narrative. One of the challenges for political methodology is to get beyond this cleavage.

(p. 1036)

Table 48.3. Two dimensions of political science analysis, 1970–99

	Component	
	Causal	Interpretative
Narrative	.018	.759
Interpretive	.103	.738
Causal/causality	.700	.105
Hypothesis	.750	–.073
Explanation	.701	.131

Notes: Extraction method: principal component analysis. Rotation method: oblimin with Kaiser normalization.

Based upon our qualitative understanding of methodological perspectives in American political science, we searched among all articles in JSTOR from 1970 to 1999 for five words that we suspected might have a two-dimensional structure. The words were “narrative,” “interpretive,” “causal or causality,” “hypothesis,” and “explanation.” After obtaining their correlations across articles,²⁶ we used principal components and an oblimin rotation to clarify the structure. We found two eigenvalues with sizes larger than one which suggested the two-dimensional principal components solution reported in Table 48.3. There is clearly a “causal dimension” which applies to roughly one-third of the articles and an “interpretive” dimension which applies to about 6 percent of the articles.²⁷ Although we expected this two-dimensional structure, we were somewhat surprised to find that the word “explanation” was almost entirely connected with “causal or causality” and with “hypothesis.” And we were surprised that the two dimensions were completely distinctive since they are essentially uncorrelated at .077. Moreover, in a separate analysis, we found that whereas the increase in “causal thinking” occurred around 1960 or maybe even 1950 in political science (see Figure 48.1), the rise in the use of the terms “narrative” and “interpretive” came in 1980.²⁸

We view these findings not as a “fact” to be accepted, but as a methodological divide to overcome. “Causal thinking” is clearly not the only approach to political analysis. Modern political methodology, as demonstrated in this chapter, has a repertoire that recognizes the interpretative problems inherent in conceptualization, historical narrative (Mahoney and Terrie 2008), intensive interviewing (Rathbun 2008), and other (p. 1037) methods (Goodin and Tilly 2006, ch. 28). In an effort to come to grips with different conceptions of political research, modern methodology is developing techniques that go beyond simple behavioralism.

4.2 Organizations, Institutions, and Movements in the Field of Methodology

We see strong reasons to believe that political methodology will continue to advance far beyond simple behavioralism. One of the most interesting features of American political science during the last fifty years has been the attention paid to methodology and the creation of organizations devoted to its advancement. The initial step in this direction was the creation in 1962 of the Inter-university Consortium for Political Research (now the Inter-university Consortium for Political and Social Research), which runs an important summer training program in methods and that has become an internationally renowned facility for data archiving (Converse 1964; Franklin 2008). It is hard to over-emphasize the impact of the ICPSR on American political science through its training and

archiving of digital data.

In the last twenty-five years, several methodological movements have developed within political science, which involved the present authors at first hand.²⁹ The two methodology sections of the American Political Science Association are among the largest of the discipline's thirty-eight sections. The Political Methodology Section, formed in 1984, has annual summer meetings that have grown to hundreds of attendees, and the section's journal, *Political Analysis*, publishes some of the best articles on political methodology (Franklin 2008; Lewis-Beck 2008). The APSA organized section for Qualitative Methods (now "Qualitative and Multi-Method Research") became an APSA organized section in 2003, and works in parallel with the Institute for Qualitative and Multi-Method Research (initially at Arizona State University, but now at Syracuse University) that has run a training program since 2002.

Recently, the National Science Foundation, under the leadership of James Granato and Frank Scioli (Granato and Scioli 2004), created the Empirical Implications of Theoretical Models (EITM) initiative to create a new generation of scholars who knew enough formal theory and enough about methods to do two things: (1) build theories that could be tested; and (2) develop methods for testing theories (Aldrich, Alt, and Lupia 2008). Two major EITM summer programs (one that has rotated among Harvard, Duke, Michigan, Berkeley, and UCLA since 2002 and another at Washington (p. 1038) University in St. Louis that has run since 2003) have trained several hundred students and faculty members. NSF has also provided support for the qualitative methods training institute, as well as for a new initiative to explore ways of archiving qualitative data.

Through these various institutes and organizations, the discipline has expanded its ability to train its own graduate students, and there is an increasing capacity to train both graduate and undergraduate students within political science departments. More attention is being paid to the training of undergraduates as a fundamental base needed for the discipline, including discussions of a political methodology Wiki led by Philip Schrodtt. Political methodology is also finding more and more connections with theory. Beck (2000) draws the contrast between statisticians and political methodologists in that "statisticians work hard to get the data to speak, whereas political scientists are more interested in testing theory." The focus on theory draws both quantitative and qualitative political scientists into the substance of politics, and it helps unite political methodologists with the political science community. Finally, the range and scope of outlets for publishing work in political methodology has increased dramatically in the last forty years (Lewis-Beck 2008). Based on the vibrancy of these institutions, the future of political methodology looks bright indeed.

4.3 Political Science at the Disciplinary Crossroads

In two different ways, political methodology straddles a disciplinary crossroads with all the hubbub, excitement, and diversity of the traditional bazaar. Within the discipline of political science, methodology serves the many different substantive and methodological interests of political scientists. In doing this, it helps to unite the discipline by focusing on the importance of providing methods that offer conceptual clarity, better interpretations of meanings, inferential leverage, and better explanations. In the words of two of us, it provides "diverse tools, shared standards" (Brady and Collier 2004).

Political methodology is also at the crossroads of other disciplines because it has evolved from borrowing, to welcoming, to creating new wares and new methods. Historically, quantitative methodology borrowed heavily from statistics, sociology, econometrics, psychometrics, statistics again, and most recently, biostatistics. Qualitative methodology borrowed from anthropology, history, and sociology. However, political methodology has recently begun to come into its own. Beck (2000) characterizes the field as one of welcoming methods from other disciplines, while also making important advances tailored to the unique features of political science data and problems. This represents a substantial improvement on Achen's assessment of political methodology as a field that "has so far failed to make serious theoretical progress on any of the major issues facing it" (1983, 69). Achen lamented that a major preoccupation (and limitation) of political methodologists was simply to teach methods developed in other fields.

(p. 1039) Bartels and Brady (1993) noted approvingly that beyond describing important methodological developments in other fields, political scientists were routinely applying advanced quantitative techniques in every substantive area of enquiry in political science. Indeed, in looking specifically at the two areas Achen (1983) identified as ripe for methodological contributions (the nature of survey response and economic voting), Bartels

and Brady concluded that “political methodologists have made significant progress on both of these problems in the intervening decade” and suggested that “further investment in basic methodological research will continue to pay handsome dividends in terms of our substantive understanding of politics” (1993,146).

Recent contributions in qualitative methodology, interestingly enough, have been able to draw upon canonical works from within political science. Examples are Sartori (1970), Przeworski and Teune (1970), Verba (1967; 1971a; 1971b), Lijphart (1971; 1975), Eckstein (1975), Almond and Genco (1977), and George (1979). Notwithstanding significant contributions from other disciplines, these studies retained a foundational status for qualitative researchers well into the 1990s, and even with the renaissance of writing on qualitative methods that began slowly in the 1990s, and expanded greatly after 2000 (Collier and Elman 2008), this earlier work by political scientists retains its importance.

While political methodologists should always seek to introduce methodological advances from other fields and to teach the “canon” of well-known methods, it is heartening that the field is beginning to make its own contributions, and that the evolution from borrowing, to welcoming, to a discipline that is welcomed by other social sciences is occurring.

We believe that in doing this, political methodologists should keep in mind three principles.

- Techniques should be the servants of improved data collection, measurement, and conceptualization and of better understanding of meanings and enhanced identification of causal relationships. Methodologists should develop strong research designs that ensure “that the results have internal, external, and ecological validity” (Educational Psychology 2008), and these principles are more important than whether researchers use qualitative or quantitative methods.

- These tasks can be undertaken in diverse ways: description and modeling, case-study and large-*n* designs, and quantitative and qualitative research.

- Techniques should cut across boundaries and be useful for many different kinds of researchers.

Methodologists should ask how their methods can be used by, or at least inform, the work of those outside those areas where they are usually employed. For example, those describing large-*n* statistical techniques should provide examples of how their methods inform, or even are adopted by, those doing case studies or interpretive work. Similarly, authors explaining how to do comparative historical work or process tracing should reach out to explain how it could inform those doing cross-sectional or time series studies.

(p. 1040) From our survey of the literature on qualitative and quantitative methods in the social sciences, these principles seem to be taking hold, and we come to three general conclusions. First, there is a lot of interest in multimethod research—integrating, combining, or mixing methods—however it is termed. Increasing numbers of researchers argue that mixed methods research is an attractive alternative to quantitative-only or qualitative-only research, e.g. Johnson and Onwuegbuzie (2004). Even more quantitatively oriented disciplines, such as counseling psychology, are incorporating what used to be less-favored qualitative methods in their work (Hanson et al. 2005), and some feminist researchers have begun to use more and more quantitative research methods, which in the past have been criticized as associated with masculinity, e.g. Westmarland (2001). In fact, some researchers go so far as to argue that mixed methods should be a separate research movement as opposed to just a fusion of the qualitative and quantitative approaches (Tashakkori and Teddlie 2002).

While there is excitement about multimethod work, not everyone agrees that it is the best way to proceed (see in particular, Symposium: Multi-Method Work, *Dispatches from the Front Lines* 2007). Bennett (2007) points out that multiple methods might ensure that one method’s weaknesses will be offset by the strength of another, but it is also possible that errors might simply accumulate because users might not have had enough time to master all the methods. Indeed, some scholars conclude that it is better to master just one method (see Wittenberg 2007).

Second, “Quantitative versus Qualitative” debates are for the most part over. There are not many quantitative–qualitative “debate” articles. Across the social sciences, this type of article was mostly published in the 1980s and early 1990s. There can be little doubt that research that involves the integration of quantitative and qualitative research has become increasingly common in recent years as a result (Taylor 2006). Recently published articles are generally moving beyond debates (whether one method wins over the other or vice versa) to how to do multimethod research or how to build upon insights or gaps of one or the other.

Third, despite signs of convergence, there still remains an unproductive bifurcation between qualitative and quantitative methods in political science. The *Oxford Handbook of Political Methodology* was designed to help fill that gap, but there are still many instances where researchers pursue one or the other method when a combination would be much more powerful.

4.4 The Future

We end with some speculation about the specific kinds of methods that might be developed in the next thirty years. This exercise is no doubt foolhardy, but it might be amusing to those who read this chapter now and in the future. We organize the possibilities by the two major sections in this review.

(p. 1041) 4.4.1 Conceptualization, Measurement, and Data Collection

1. *Conceptualization.* Just as the last thirty years have seen substantial changes in our understanding of the nature of conceptualization and in our gunny sack of useful concepts, the future should bring innovations in:

(a) *Our understanding of conceptualization.* Social science concepts still remain mysterious things with their essential contestability, their constructed nature, and their position between political practice and political science research. We expect social theorists to continue to think hard and carefully about concepts and we expect that methodologists will develop new methods that will take us beyond validity, reliability, factor analyses, and our current understandings.

(b) *Micro-concepts.* We will learn much more about nature of attitudes such as identity, efficacy, liberalism-conservatism, and their roots in neurobiology. We will learn more about the nature of interactions through developments in theoretical and experimental game theory, and the role of social networks through developments in sociology and other fields.

(c) *Macro-concepts.* We will improve our understanding of democracy, revolutions, transitions, legitimacy, collective action, and many other concepts by having better theories, more data, and better linkages between micro-concepts and macro-concepts (e.g. the role of regime approval, repression, and ignorance in the microfoundations of legitimacy, and the role of context and social networks in collective action or revolutions).

2. *A better understanding of how people conceptualize politics.* The “left-right” and multidimensional conceptualization of politics has been a major innovation in our comprehension of political cleavages, mass politics, and elite politics in organizations, bureaucracies, legislatures, and courts. Cognitive and neurobiological science may tell us more about how people think about politics and about whether our spatial metaphors really make sense at a fundamental level (Gardenfors 2004). The result will be important for how we model and think about politics.

3. *Better measurements.* We expect that there will be better indicators of emotions, values, social networks, racism, ethnic identity, and so forth. Moreover, these methods will increasingly use neurobiological and cognitive science techniques, game theory methods (e.g. “Trust games,” Camerer 2003), and the power of the web.

4. *Automatic text, audio, and video coders.* As computer scientists and linguists get better at parsing text, there will be more and better automatic text decoders which will allow researchers to code the vast bodies of data now available on the web in a reproducible and fully documented fashion. These tools will be invaluable for both quantitative and qualitative researchers.

5. *New kinds of large data-sets.* There will be more linking of data over time and space; across modalities of data (video, written, behavioral), and across different sources. There will be continuous sampling of information through modern (p. 1042) telecommunications (e.g. cellphone read-outs of people’s locations). Scholars will be increasingly able to analyze large collections of text, audio, and video data, which will greatly extend the horizons of qualitative and quantitative researchers.

6. *World indicators.* We see substantial use of indicators to measure the level of democracy, human rights, corruption, happiness, feelings of efficacy, quality of service delivery, and other features of governments. There will be an increasing use of these indicators, and we hope that there will be greatly expanded attention to their reliability and validity, which at times is problematic.

7. *More use of the web for data collection.* Surveys, experiments, and simulations will be increasingly done on the web (or whatever serves as the system that integrates the web, cellphones, and television) (Berman

and Brady 2005).

4.4.2 Causal Inference

1. *New techniques for making causal statements.* Statistics courses in the social sciences will increasingly start by studying what is meant by a causal inference. We will see a better understanding of how to use case study and other information, and there will be new statistical techniques that tell us how to establish the inevitably complicated linkages of information across studies (akin to “meta-analysis”).
2. *Advances from neurobiology and cognitive sciences.* New methods will be developed for linking political thinking and behavior to the brain. The work using the fMRI will be extended in ways that will provide much more detailed information about what is happening in the brain.
3. *Better linking of formal models to testing procedures.* More sophisticated methods will emerge to link formal models to testing procedures. Often this will involve teams of social scientists with specialties in methods, modeling, or substantive areas of research.
4. *Large-scale field experiments.* Experimental studies of voting turnout and deliberative polls are some of the first-generation work using large-scale field experiments. The next generation will see experiments in countries as they try different public policies and different political systems.
5. *New designs for causal inference.* There will be even more sophisticated ways developed to make causal inferences that combine counterfactuals, manipulations, and interventions in the operation of putative mechanisms. These new methods will make novel uses of temporal and spatial variation.

In sum, we are convinced that political methodologists will continue a vigorous program of innovation on many fronts, and we are convinced that such innovation will make invaluable contributions to the shared goals of our discipline: to advance the substantive understanding of politics.

References

- Abdelal, R., Herrera, Y. M., Johnston, A. I., and McDermott, R. (eds.) 2009. *Measuring Identity*. Cambridge, Mass.: Cambridge University Press.
- Achen, C. H. 1983. Towards theories of data. In *Political Science: The State of the Discipline*, ed. A. Finifter. Washington, DC: American Political Science Association.
- 1986. *The Statistical Analysis of Quasi-experiments*. Los Angeles: University of California Press.
- 2002. Toward a new political methodology: microfoundations and ART. *Annual Review of Political Science*, 5: 423–50.
- Adcock, R. and Collier, D. 2001. Measurement validity: a shared standard for qualitative and quantitative research. *American Political Science Review*, 95: 529–46.
- Aldrich, J. H., Alt, J. E., and Lupia, A. 2008. The EITM approach: origins and interpretations. Ch. 37 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Allsop, D. and Weisberg, H. 1988. Measuring change in party identification in an election campaign. *American Journal of Political Science*, 32: 996–1017.
- Almond, G. and Genco, S. J. 1977. Clouds, clocks, and the study of politics. *World Politics*, 29: 489–522.
- and Verba, S. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton, NJ: Princeton University Press.
- Anderson, B. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Arjomand, S. A. 1986. Iran’s Islamic revolution in comparative perspective. *World Politics*, 38: 383–414.
- Arkes, H. R. and Tetlock, P. E. 2004. Attributions of implicit prejudice, or “would Jesse Jackson ‘fail’ the implicit association test?” *Psychological Inquiry*, 15: 257–78.

- Arrow, K. J. 1951. *Social Choice and Individual Values*. New York: Wiley.
- Azar, E. E. 1970. The analysis of international events. *Peace Research Reviews*, Nov.: 1–113.
- Cohen, S. H., Jukam, T. O., and McCormick, J. M. 1972. The problem of source coverage in the use of international events data. *International Studies Quarterly*, 16: 373–88.
- Banks, A. S. *Cross-Polity Time-Series Data*. Cambridge, Mass.: MIT Press.
- and Textor, R. B. 1963. *A Cross Polity Survey*. Cambridge, Mass.: MIT Press.
- Bartels, B., Box-Steffensmeier, J. M., Smidt, C., and Smith, R. M. 2007. Microfoundations of partisanship. Typescript, Ohio State University.
- Bartels, L. 1987. Candidate choice and the dynamics of the presidential nominating process. *American Journal of Political Science*, 31: 1–30.
- and Brady, H. E. 1993. The state of quantitative political methodology. Vol. 2, pp. 121–59, in *Political Science: The State of the Discipline*, ed. A. Finifter. Washington, DC: American Political Science Association.
- Barton, A. H. 1955. The concept of property-space in social research. In *The Language of Social Research: A Reader in the Methodology of Social Research*, ed. P. F. Lazarsfeld and M. Rosenberg. Glencoe, Ill.: Free Press.
- and Lazarsfeld, P. F. 1969. Some functions of qualitative analysis in social research. In *Issues in Participant Observation*, ed. G. J. McCall and J. L. Simmons. Reading, Mass.: Addison-Wesley.
- Baumgartner, F. R., Jones, B. D., and MacLeod, M. C. 1998. Ensuring quality, reliability, and usability in the creation of a new data source. *Political Methodologist*, 8: 1–10.
- Beck, N. 2000. Political science: a welcoming discipline. *Journal of the American Statistical Association*, 95: 651–64.
- (p. 1044)** Beck, N. 2008. Time series cross-sectional methods. Ch. 20 in **Box-Steffensmeier, Brady, and Collier 2008**.
- and Katz, J. N. 1996. Nuisance vs. substance: specifying and estimating time series–cross-section models. *Political Analysis*, 6: 1–36.
- Beissinger, M. 2002. *Nationalist Mobilization and the Collapse of the Soviet State*. Cambridge, Mass.: Cambridge University Press.
- Belli, R. F., Traugott, M. W., Young, M., and McGonagle, K. A. 1999. Reducing vote overreporting in surveys: social desirability, memory failure, and source monitoring. *Public Opinion Quarterly*, 63: 90–108.
- Ben-Dak, J. D. and Azar, E. 1972. Research perspectives on the Arab–Israeli conflict: introduction to a symposium. *Journal of Conflict Resolution*, 16: 131–4.
- Bennett, A. 2007. Introduction for symposium: “Multi-Method Work, Dispatches from the Front Lines.” *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 5: 9.
- 2008. Process tracing: a Bayesian perspective. Ch. 30 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Berger, P. L. and Luckmann, T. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor.
- Berinsky, A. J. 2002. Political context and the survey response: the dynamics of racial policy opinion. *Journal of Politics*, 64: 567–84.
- Berman, F. and Brady, H. E. 2005. Final report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences. <<http://vis.sdsc.edu/sbe/reports/SBE-CISE-FINAL.pdf>>

Bevir, M. 2008. Meta-methodology: clearing the underbrush. Ch. 3 in **Box-Steffensmeier, Brady, and Collier 2008**.

—and Kedar, A. 2008. Concept formation in political science: an anti-naturalist critique of qualitative methodology. *Perspectives on Politics*, 6: 503–17.

Beyle, H. C. 1931. *Identification and Analysis of Attribute-Cluster-Blocs: A Technique for Use in the Investigation of Behavior in Governance, including Report on Identification and Analysis of Blocs in a Large Non-Partisan Legislative Body, the 1927 Session of the Minnesota State Senate*. Chicago: University of Chicago Press.

Black, D. 1948. The decisions of a committee using a special majority. *Econometrica*, 16: 245–61.

Blalock, H. M., Jr. 1964, *Causal Inference in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Blumer, H. 1969. *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, NJ: Prentice Hall.

Boas, T. and Gans-Morse, J. 2009. Neoliberalism: from new liberal philosophy to anti-liberal slogan. *Studies in Comparative International Development*.

Bollen, K. A., Rabe-Hesketh, S., and Skrondal, A. 2008. Structural equation models. Ch. 18 in **Box-Steffensmeier, Brady, and Collier 2008**.

Box-Steffensmeier, J. M., Brady, H. E., and Collier, D. (eds.) 2008. *Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

—and Smith, R. M. 1996. The dynamics of aggregate partisanship. *American Political Science Review*, 90: 567–80.

Brady, H. E. 2008. Causation and explanation in social science. Ch. 10 in **Box-Steffensmeier, Brady, and Collier 2008**.

—and Collier, D. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. New York: Rowman and Littlefield.

(p. 1045) —and Kaplan, C. 2000. Categorically wrong? Nominal versus graded measures of ethnic identity. *Studies in Comparative International Development*, 35: 56–91.

—2009. Conceptualizing and measuring ethnic identity. In **Abdelal et al. 2009**.

—and Johnston, R. 2006. The rolling cross-section and causal attribution. Pp. 164–95 in *Capturing Campaign Effects*, ed. H. E. Brady and R. Johnston. Ann Arbor: University of Michigan Press.

Bradburn, N., Sudman, S., and Wansink, B. 2004. *Asking Questions: The Definitive Guide to Questionnaire Design—For Market Research, Political Polls, and Social Health Questionnaires*. San Francisco: Jossey-Bass.

Brass, P. 1991. *Ethnicity and Nationalism*. Beverly Hills, Calif.: Sage.

Brimhall, D. R. and Otis, A. S. 1948. Consistency of voting by our Congressmen. *Journal of Applied Psychology*, 32: 1–14.

Camerer, C. F. 2003. *Behavioral Game Theory*. Princeton, NJ: Princeton University Press.

Campbell, A. and Kahn, R. L. 1952. *The People Elect a President*. Survey Research Center, Institute for Social Research, University of Michigan.

Campbell, D. and Fiske, D. 1959. Convergent and discriminant validation by the multitrait multimethod matrix. *Psychological Bulletin*, 56: 81–105.

Campbell, D. T. and Stanley, J. C. 1966. *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand-McNally.

- Cho, W. K. T. and Manski, C. F. 2008. Cross-level/ecological inference. Ch. 24 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Clinton, J., Jackman, S., and Rivers, D. 2004. The statistical analysis of roll call data. *American Political Science Review*, 98: 355–70.
- Cohen, M. R. and Nagel, E. 1934. *An Introduction to Logic and Scientific Method*. New York: Harcourt, Brace, and Company.
- Collier, D. 1979. *The New Authoritarianism in Latin America*. Princeton, NJ: Princeton University Press.
- and Adcock, R. 1999. Democracy and dichotomies: a pragmatic approach to choices about concepts. *Annual Review of Political Science*, 2: 537–65.
- and Elman, C. 2008. Qualitative and multimethod research: organizations, publication, and reflections on integration. Ch. 34 in **Box-Steffensmeier, Brady, and Collier 2008**.
- and Gerring, J. (eds.) 2008. *Concepts and Method in Social Science: The Tradition of Giovanni Sartori*. Oxford: Routledge.
- Hidalgo, F. D., and Maciuceanu, A. O. 2006. Essentially contested concepts: debates and applications. *Journal of Political Ideologies*, 11: 211–46.
- and Levitsky, S. 1997. Democracy with adjectives: conceptual innovation in comparative research. *World Politics*, 49: 430–51.
- 2008. Democracy: conceptual hierarchies in comparative research. Ch. 10 in **Collier and Gerring 2008**.
- LaPorte, J., and Seawright, J. 2008. Typologies: forming concepts and creating categorical variables. In **Box-Steffensmeier, Brady, and Collier 2008**.
- and Mahon, Jr., J. E. 1993. Conceptual stretching revisited: adapting categories in comparative analysis. *American Political Science Review*, 87: 845–55.
- Collier, R. B. and Collier, D. 1979. Inducements versus constraints: disaggregating “corporativism.” *American Political Science Review*, 73: 967–86.
- 1991. *Shaping the Political Arena*. Princeton, NJ: Princeton University Press.
- Converse, P. E. 1964. The nature of belief systems in mass publics. In *Ideology and Discontent*, ed. D. Apter. New York: Free Press.
- (p. 1046)** Cook, T. D. and Campbell, D. T. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton-Mifflin.
- Coppedge, M. and Reinicke, W. H. 1990. Measuring polyarchy. *Studies in Comparative International Development*, 25: 51–72.
- Cruse, D. A. 2004. *Meaning and Language: An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.
- Dahl, R. 1961. The behavioral approach in political science: epitaph for a monument to a successful protest. *American Political Science Review*, 55: 763–72.
- 1971. *Polyarchy: Participation and Opposition*. New Haven, Conn.: Yale University Press.
- Davis, O. A., Hinich, M. J., and Ordeshook, P. C. 1970. An expository development of a mathematical model of the electoral process. *American Political Science Review*, 64: 426–48.
- Donsbach, W. and Traugott, M. 2007. *The SAGE Handbook of Public Opinion Research*. Newbury Park, Calif.: Sage.

- Doran, C. F., Pendley, R. E., and Atunes, G. 1973. A test of cross-national event reliability. *International Studies Quarterly*, 17: 175–203.
- Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Eckstein, H. 1975. Case study and theory in political science. Pp. 79–138 in *Handbook of Political Science*, ed. F. Greenstein and N. Polsby. Reading, Mass.: Addison-Wesley.
- Eldersveld, S. J. 1964. *Political Parties: A Behavioral Analysis*. Chicago, Illinois: Rand-McNally.
- Elkins, Z. 2000. Gradations of democracy? Empirical tests of alternative conceptualizations. *American Journal of Political Science*, 44: 293–300.
- Elman, C. 2005. Explanatory typologies in qualitative studies of international politics. *International Organization*, 59: 293–326.
- Elster, J. 1998. A plea for mechanisms. In *Social Mechanisms*, ed. P. Hedstrom and R. Swed-berg. Cambridge: Cambridge University Press.
- Emerick, C. F. 1910. A neglected factor in race suicide. *Political Science Quarterly*, 25: 638–65.
- Eulau, H. 1963. *The Behavioral Persuasion in Politics*. New York: Random House.
- 1969. *Behavioralism in Political Science*. New York: Atherton.
- Farr, J., Dryzek, J. S., and Leonard, S. T. (eds.) 1995. *Political Science in History: Research Programs and Political Traditions*. New York: Cambridge University Press.
- and Seidelman, R. 1993. *Discipline and History: Political Science in the United States*. Ann Arbor: University of Michigan Press.
- Fazio, R. H. and Olson, M. A. 2003. Implicit measures in social cognition research: their meaning and use. *Annual Review of Psychology*, 54: 297–327.
- Feierabend, I. K. and Feierabend, R. L. 1966. Aggressive behaviors within polities, 1948–1962: a cross-national study. *Journal of Conflict Resolution*, 10: 149–79.
- Fisher, R. A. 1925. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Franklin, C. H. 2008. Quantitative methodology. Ch. 35 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Franzese, R. J., Jr. and Hays, J. C. 2008. Empirical models of spatial interdependence. Ch. 25 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Freeden, M. 1994. Political concepts and ideological morphology. *Journal of Political Philosophy*, 2: 140–64.
- Freedman, D. 1997. From association to causation via regression. *Advances in Applied Mathematics*, 18: 59–110.
- 2008. On types of scientific enquiry: the role of qualitative reasoning. Ch. 12 in **Box-Steffensmeier, Brady, and Collier 2008**.
- (p. 1047) Freeman, J. R. and Goldstein, J. S. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.
- Gage, N. L. and Shimberg, B. 1949. Measuring senatorial “progressivism.” *Journal of Abnormal and Social Psychology*, 44: 112.
- Gallie, W. B. 1956a. Essentially contested concepts. *Proceedings of the Aristotelian Society*, 56: 167–98.
- 1956b. Art as an essentially contested concept. *Philosophical Quarterly*, 6: 97–114.
- Gardenfors, P. 2004. *Conceptual Spaces: The Geometry of Thought*. Boston: MIT Press.

- Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
- Geanakoplos, J. 1992. Common knowledge. *Journal of Economic Perspectives*, 6: 53–82.
- George, A. L. 1979. Case studies and theory development: the method of structured, focused comparison. In *Diplomacy: New Approaches in History, Theory, and Policy*, ed. P. G. Lauren. New York: Free Press.
- and Bennett, A. 2005. *Case Studies and Theory Development in the Social Sciences*. Boston: MIT Press.
- Gerber, A. S. and Green, D. P. 2008. Field experiments and natural experiments. Ch. 15 in **Box-Steffensmeier, Brady, and Collier 2008**.
- and Kaplan, E. H. 2004. The illusion of learning from observational research. Pp. 251–73 in *Problems and Methods in the Study of Politics*, ed. I. Shapiro, R. Smith, and T. Massoud. New York: Cambridge University Press.
- Gerring, J. 2001. *Social Science Methodology: A Criterial Framework*. New York: Cambridge University Press.
- Giddens, A. 1986. *The Constitution of Society*. Berkeley: University of California Press.
- Gill, J. 2007. *Bayesian Methods: A Social and Behavioral Sciences Approach*, 2nd edn. Boca Raton, Fla.: Chapman and Hall.
- Glasgow, G. and Alvarez, R. M. 2008. Discrete choice methods. Ch. 22 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Goertz, G. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.
- and Levy, J. 2007. *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*. New York: Routledge.
- and Starr, H. (eds.) 2002. *Necessary Conditions: Theory, Methodology, and Applications*. New York: Rowman and Littlefield.
- Goffman, E. 1959. *The Presentation of Self in Everyday Life*. New York: Anchor.
- Goldstein, J. S. and Freeman, J. R. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.
- Golub, J. 2008. Survival analysis. Ch. 23 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Goodin, R. E. and Tilly, C. (eds.) 2006. *Oxford Handbook of Contextual Political Analysis*. Oxford: Oxford University Press.
- Gosnell, H. F. and Gill, N. N. 1935. An analysis of the 1932 presidential vote in Chicago. *American Political Science Review*, 29: 967–84.
- Granato, J. and Scioli, F. 2004. Puzzles, proverbs, and omega matrices: the scientific and social significance of empirical implications of theoretical models (EITM). *Perspectives on Politics*, 2: 313–23.
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, 74: 1464–80.
- Gurr, T. R. 1968. A casual model of civil strife: a comparative analysis using new indices. *American Political Science Review*, 62: 1104–24.
- (p. 1048) Gurr, T. R. 1970a. Sources of rebellion in Western societies: some quantitative evidence. *Annals of the American Academy of Political and Social Science*, 391: 128–44.
- 1970b. *Why Men Rebel*. Princeton, NJ: Princeton University Press.
- Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., and Creswell, J. D. 2005. Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology*, 52: 224–35.

- Hanushek, E. A. and Jackson, J. E. 1977. *Statistical Methods for Social Scientists*. Orlando, Fla.: Academic Press.
- Heckman, J. 2008. Econometric causality. NBER Working Paper 13934.
- Hedstrom, P. 2008. Studying mechanisms to strengthen causal inferences in quantitative research. Ch. 13 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Hempel, C. G. 1952. *Fundamentals of Concept Formation in Empirical Science*. Chicago: University of Chicago Press.
- Hirsch, F. 2005. *Empire of Nations: Ethnographic Knowledge and the Making of the Soviet Union*. Ithaca, NY: Cornell University Press.
- Holbrook, T. M. 1996. *Do Campaigns Matter?* Newbury Park, Calif.: Sage.
- Holland, P. W. 1986. Statistics and causal inference (in theory and methods). *Journal of the American Statistical Association*, 81: 945–60.
- Hopkins, A. H. and Weber, R. E. 1976. Dimensions of public policies in the American states. *Polity*, 8: 475–89.
- Hroch, M. 1986. *Social Preconditions of National Revival in Europe: A Comparative Analysis of the Social Composition of Patriotic Groups among the Smaller European Nations*. New York: Cambridge University Press.
- Jackman, S. 2001. Multidimensional analysis of roll call data via Bayesian simulation: identification, estimation, inference and model checking. *Political Analysis*, 9: 227–41.
- Jackson, J. E. 2008. Endogeneity and structural equation estimation in political science. Ch. 17 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Jennings, M. K. and Niemi, R. G. 1974. *The Political Character of Adolescence: The Influence of Families and Schools*. Princeton, NJ: Princeton University Press.
- 1981. *A Panel Study of Young Adults and their Parents*. Princeton, NJ: Princeton University Press.
- Jerit, J., Barabas, J., and Bolsen, T. 2006. Citizens, knowledge, and the information environment. *American Journal of Political Science*, 50: 266–82.
- Johnson, R. B. and Onwuegbuzie, A. J. 2004. Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, 33: 14–26.
- Johnston, R., Blais, A., Brady, H. E., and Crete, J. 1992. *Letting the People Decide: The Dynamics of a Canadian Election*. Stanford, Calif.: Stanford University Press.
- Jones, B. S. 2008. Multilevel models. Ch. 26 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Key, V. O., Jr. 1949. *Southern Politics in State and Nation*. New York: Alfred A. Knopf.
- King, G. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual behavior from Aggregate Data*. Princeton, NJ: Princeton University Press.
- 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. New York: Cambridge University Press.
- Keohane, R. O., and Verba, S. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
- and Lowe, W. 2003. An automated information extraction tool for international conflict data with performance as good as human coders: a rare events evaluation design. *International Organization*, 57: 617–42.
- (p. 1049) —Murray, C. J. L., Salomon, J. A., and Tandon, A. 2003. Enhancing the validity and cross-cultural comparability of survey research. *American Political Science Review*, 97: 567–83. Reprinted with printing errors

corrected February 2004.

Koopmans, R. and Rucht, D. 2002. Protest events analysis. Pp. 231–59 in *Methods of Social Movement Research*, ed. B. Klandermans and S. Staggenborg. Minneapolis: University of Minnesota Press.

Kotowski, C. M. 1984. Revolution. In *Social Science Concepts: A Systematic Analysis*, ed. G. Sartori. Beverly Hills, Calif.: Sage.

Krysan, M. and Couper, M. P. 2003. Race in the live and the virtual interview: racial deference, social desirability, and activation effects in attitude surveys. *Social Psychology Quarterly*, 66: 364–83.

Kuklinski, J. H., Cobb, M. D., and Gilens, M. 1997. Racial attitudes and the “New South.” *Journal of Politics*, 59: 323–49.

Kurtz, M. J. 2000. Understanding peasant revolution: from concept to theory to case. *Theory and Society*, 29: 93–124.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lasswell, H. D. 1941. The World Attention Survey. *Public Opinion Quarterly*, 5: 456–62.

—and Blumenstock, D. 1939. The volume of communist propaganda in Chicago. *Public Opinion Quarterly*, 3: 63–78.

—and Sereno, R. 1937. Governmental party leaders in Fascist Italy. *American Political Science Review*, 31: 914–29.

Lazarsfeld, P. F. and Barton, A. H. 1951. Qualitative measurement in the social sciences: classification, typologies, and indices. In *The Policy Sciences*, ed. D. Lerner and H. D. Lasswell. Stanford, Calif.: Stanford University Press.

—Berelson, B., and Gaudet, H. 1944. *The People’s Choice: How the Voter Makes up his Mind in a Presidential Campaign*. New York: Duell, Sloan and Pearce.

—and McPhee, W. N. 1986. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.

Lee, S. M. 1993. Racial classifications in the U.S. Census, 1890 to 1990. *Ethnic and Racial Studies*, 16: 75–94.

Lee, T. 2007. From shared demographic categories to common political destinies: immigration and the link from racial identity to group politics. *Du Bois Review*, 4: 433–56.

Levitsky, S. 1998. Peronism and institutionalization: the case, the concept, and the case for unpacking the concept. *Party Politics*, 4: 77–92.

Levy, J. S. 2008. Counterfactuals and case studies. Ch. 27 in **Box-Steffensmeier, Brady, and Collier 2008**.

Lewis, J. and Poole, K. 2004. Measuring bias and uncertainty in ideal point estimates via the parametric bootstrap. *Political Analysis*, 12: 105–27.

Lewis-Beck, M. 2008. Forty years of publishing in quantitative methodology. Ch. 36 in **Box-Steffensmeier, Brady, and Collier 2008**.

Lieberman, M., Schreiber, D., and Ochsner, K. 2003. Is political cognition like riding a bicycle? How cognitive neuroscience can inform research on political thinking. *Political Psychology*, 24: 681–704.

Lijphart, A. 1971. Comparative politics and comparative method. *American Political Science Review*, 65: 682–93.

—1975. The comparable cases strategy in comparative research. *Comparative Political Studies*, 8: 158–77.

(p. 1050) Lipset, S. M. and Rokkan, S. 1967. Cleavage structures, party systems and voter alignments: an

- introduction. In *Party Systems and Voter Alignments: Cross-National Perspectives*, ed. S. M. Lipset and S. Rokkan. New York: Free Press.
- Lord, F. M. and Novick, M. R. 1968. *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- Luks, S. and Brady, H. E. 2003. Defining welfare spells: coping with problems of survey responses and administrative data. *Evaluation Review*, 27: 395–420.
- McKinney, J. C. 1966. *Constructive Typology and Social Theory*. New York: Meredith.
- MacRae, D., Jr. 1970. *Issues and Parties in Legislative Voting: Methods of Statistical Analysis*. New York: Harper and Row.
- Mahoney, J. 2001. *The Legacies of Liberalism: Path Dependence and Political Regimes in Central America*. Baltimore: Johns Hopkins University Press.
- and Terrie. P. L. 2008. Comparative-historical analysis in contemporary political science. Ch. 32 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Marcus, G. E. and Mackuen, M. B. 1993. Anxiety, enthusiasm, and the vote: the emotional underpinnings of learning and involvement during presidential campaigns. *American Political Science Review*, 87: 672–85.
- Martin, A. D. 2008. Bayesian analysis. Ch. 21 in **Box-Steffensmeier, Brady, and Collier 2008**.
- and Quinn, K. M. 2007. Assessing preference change on the US Supreme Court. *Journal of Law, Economics, and Organization Advance Access*, 23: 365–85.
- Mavor, J. 1895. Labor and politics in England. *Political Science Quarterly*, 10: 486–517.
- Mill, J. S. 1888. *A System of Logic, Ratiocinative and Inductive*, 8th edn. New York: Harper and Brothers.
- Morris, J. P., Squires, N., Taber, C. S., and Lodge, M. 2003. Activation of political attitudes: a psychophysiological examination of the hot cognition hypothesis. *Political Psychology*, 24: 727–45.
- Morton, R. B. and Williams, K. C. 2008. Experimentation in political science. Ch. 14 in **Box-Steffensmeier, Brady, and Collier 2008**.
- Munck, G. and Verkuilen, J. 2002. Conceptualizing and measuring democracy: evaluating alternative indices. *Comparative Political Studies*, 35: 5–34.
- Neyman, J. 1923. On the application of probability theory to agricultural experiments: essay on principles, trans. D. M. Dabrowska and T. P. Speed. *Statistical Science*, 5 (1990): 463–80.
- Nobles, M. 2000. *Shades of Citizenship: Race and the Census in Modern Politics*. Palo Alto, Calif.: Stanford University Press.
- Odegard, P. H. 1935. Political parties and group pressures. *Annals of the American Academy of Political and Social Science*, 179: 68–81.
- O'Donnell, G. and Schmitter, P. C. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore: Johns Hopkins University Press.
- Oliver, P. E. and Maney, G. M. 2000. Political processes and local newspaper coverage of protest events: from selection bias to triadic interactions. *American Journal of Sociology*, 106: 463–505.
- Patterson, M. and Monroe, K. R. 1998. Narrative in political science. *Annual Review of Political Science*, 1: 315–31.
- Paxton, P. 2000. Women in the measurement of democracy: problems of operationalization. *Studies in Comparative International Development*, 35: 92–111.
- Pearson, K. 1896. Mathematical contributions to the theory of evolution: III. Regression, heredity, and panmixia.

Philosophical Transactions of the Royal Society of London, 187: 253–318.

—1909. Determination of the coefficient of correlation. *Science*, 30: 23–5.

(p. 1051) Pevehouse, J. C. and Brozek, J. D. 2008. Time series analysis. Ch. 19 in **Box-Steffensmeier, Brady, and Collier 2008**.

Phelps, E. A. and Thomas, L. A. 2003. Race, behavior, and the brain: the role of neuroimaging in understanding complex social behaviors. *Political Psychology*, 24: 747–58.

Pierson, P. 2000. Path dependence, increasing returns, and the study of politics. *American Political Science Review*, 94: 251–67.

Pollock, J. K., Jr. 1929. The German party system. *American Political Science Review*, 23: 859–91.

Poole, K. and Rosenthal, H. 1997. *A Political-economic History of Roll Call Voting*. New York: Oxford University Press.

—1985. A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29: 357–84.

Presser, S. and Stinson, L. 1998. Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review*, 63: 137–45.

Price, V. and Zaller, J. 1993. Who gets the news? Alternative measures of news reception and their implications for research. *Public Opinion Quarterly*, 57: 133–64.

Pritchett, C. H. 1941. Divisions of opinion among justices of the U.S. Supreme Court 1939– 1941. *American Political Science Review*, 35: 890–8.

—1945. Dissent on the Supreme Court, 1943–44. *American Political Science Review*, 3: 42–54.

Przeworski, A. and Teune, H. E. 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley.

Radcliff, B. 1992. The general will and social choice theory. *Review of Politics*, 54: 34–49.

Ragin, C. 1987. *The Comparative Method*. Berkeley: University of California Press.

Raichle, M. 2003. Social neuroscience: a role for brain imaging. *Political Psychology*, 24: 759–63.

Rathbun, B. C. 2008. Interviewing and qualitative field methods: pragmatism and practicalities. Ch. 29 in **Box-Steffensmeier, Brady, and Collier 2008**.

Rice, S. A. 1928. *Quantitative Methods in Politics*. New York: Alfred A. Knopf.

Riker, W. H. 1982. *Liberalism against Populism: A Confrontation between the Theory of Democracy and the Theory of Social Choice*. San Francisco: W. H. Freeman.

—1986. *The Art of Political Manipulation*. New Haven, Conn.: Yale University Press.

Robinson, J. P., Shaver, P. R., and Wrightsman, L. S. 1993. *Measures of Political Attitudes*. Volume 2 of *Measures of Social Psychological Attitudes*. San Diego: Academic Press.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66: 688–701.

—1978. Bayesian inference for causal effects: the role of randomization. *Annals of Statistics*, 6: 34–58.

Rummel, R. J. 1963. Dimensions of conflict behavior within and between nations. *General Systems Yearbook*, 8: 1–50.

—1966. Dimensions of conflict behavior within nations, 1946–59. *Journal of Conflict Resolution*, 10: 65–73.

- 1966. Some dimensions in the foreign behavior of nations. *Journal of Conflict Resolution*, 10: 201–24.
- 1972. *The Dimensions of Nations*. Beverly Hills, Calif.: Sage.
- Russett, B., Alker, H., Jr., Deutsch, K. W., and Lasswell, H. 1964. *World Handbook of Political and Social Indicators*. New Haven, Conn.: Yale University Press.
- (p. 1052) Sartori, G. 1970. Concept misformation in comparative politics. *American Political Science Review*, 64: 1033–53.
- 1975. The Tower of Babel. In G. Sartori, F. W. Riggs, and H. Teune, *Tower of Babel: On the Definition and Analysis of Concepts in the Social Sciences*. Occasional Paper No. 6, International Studies Association, University of Pittsburgh.
- (ed.) 1984. *Social Science Concepts: A Systematic Analysis*. Beverly Hills, Calif.: Sage.
- 1991. Comparing and miscomparing. *Journal of Theoretical Politics*, 3: 243–57.
- Schattschneider, E. E. 1952. Political parties and the public interest. *Annals of the American Academy of Political and Social Science*, 280: 18–26.
- 1960. *The Semisovereign People: A Realist's View of Democracy in America*. New York: Holt, Rinehart and Winston.
- Schrodt, P. and Gerner, D. J. 1994. Validity assessment of a machine-coded event data set for the Middle East, 1982–92. *American Journal of Political Science*, 38: 825–54.
- Sekhon, J. 2004. Quality meets quantity: case studies, conditional probability and counterfactuals. *Perspectives on Politics*, 2: 281–93.
- 2008. The Neyman–Rubin model of causal inference and estimation via matching methods. Ch. 11 in **Box-Steffensmeier, Brady, and Collier 2008**.
- and Diamond, A. 2005. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. Presented at the Political Methodology Summer Conference, Florida State University, July 21–3.
- and Mebane, W. 1998. Genetic optimization using derivatives: theory and application to nonlinear models. *Political Analysis*, 7: 189–213.
- Shaw, D. R. 1999. The study of presidential campaign event effects from 1952 to 1992. *Journal of Politics*, 61: 387–422.
- 2006. *The Race to 270: The Electoral College and the Campaign Strategies of 2000 and 2004*. Chicago: University of Chicago Press.
- Simon, H. A. 1954. Spurious correlation: a causal interpretation. *Journal of the American Statistical Association*, 24: 467–74.
- Singer, J. D. and Small, M. 1972. *The Wages of War: 1816–1965: A Statistical Handbook*. New York: John Wiley.
- Skocpol, T. 1979. *States and Social Revolution*. New York: Cambridge University Press.
- Sniderman, P. and Grob, D. 1996. Innovations in experimental design in general population attitude surveys. *Annual Review of Sociology*, 22: 377–99.
- Steenbergen, M. R. and Jones, B. S. 2002. Modeling multilevel data structures. *American Journal of Political Science*, 46: 218–37.
- Stouffer, S., Guttman, L. Suchman, E. A., Lazarsfeld, P. F., Star, S. A., and Clausen, J. A. 1950. *Studies in Social Psychology in World War II: The American Soldier. 4, Measurement and Prediction*. Princeton, NJ: Princeton

University Press.

Suppes, P. 1984. *Probabilistic Metaphysics*. Oxford: Blackwell.

Symposium: Multi-Method Work, Dispatches from the Front Lines 2007. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 5/1.

Tajfel, H. 1970. Experiments in intergroup discrimination. *Scientific American*, 223: 96–102.

Tanter, R. 1966. Dimensions of conflict behavior within and between nations, 1958–60. *Journal of Conflict Resolution*, 10: 41–64.

Tashakkori, A. and Teddlie, C. B. (eds.) 2002. *Handbook of Mixed Methods Social and Behavioral Research*. Beverly Hills, Calif.: Sage.

Taylor, G. R. 2006. *Integrating Quantitative and Qualitative Methods in Research*, 2nd edn. Lanham, Md.: University Press of America.

(p. 1053) Taylor, C. L. and Hudson, M. C. 1972. *World Handbook of Political and Social Indicators*, 2nd edn. New Haven, Conn.: Yale University Press.

Taylor, J. R. 2003. *Linguistic Categorization*, 3rd edn. Oxford: Oxford University Press.

Thagard, P. 1990. Concepts and conceptual change. *Synthese*, 82: 255–74.

—1992. *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.

Tilly, C., Tilly, L., and Tilly R. 1975. *The Rebellious Century, 1830–1930*. Cambridge, Mass.: Harvard University Press.

Tiryakian, E. A. 1968. Typologies. Pp. 177–86 in *International Encyclopedia of the Social Sciences*, 16. New York: Macmillan Company and the Free Press.

Tourangeau, R., Kips, L. J., and Rasinski, K. A. 2000. *The Psychology of Survey Response*. Cambridge, Mass.: Cambridge University Press.

Verba, S. 1967. Some dilemmas in comparative research. *World Politics*, 2: 111–27.

—1971a. Cross-national survey research: the problem of credibility. In *Comparative Methods in Sociology: Essays on Trends and Applications*, ed. I. Vallier. Berkeley: University of California Press.

—1971b. Sequences and development. In *Crises and Sequences in Political Development*, ed. L. Binder et al. Princeton, NJ: Princeton University Press.

—and Nie, N. H. 1972. *Political Participation in America*. Chicago: University of Chicago Press.

Webb, S. 1889. Socialism in England. *Publications of the American Economic Association*, 4: 7–73.

Weeks, D. 1930. The Texas-Mexican and the politics of south Texas. *American Political Science Review*, 24: 606–27.

Westmarland, N. 2001. The quantitative/qualitative debate and feminist research. *Forum: Qualitative Social Research*, 2/1, Art. 13. <<http://nbn-resolving.de/urn:nbn:de:0114-fqs0101135>>

Wittenberg, J. 2007. Peril and promise: multimethods research in practice. In Symposium: Multi-Method Work, Dispatches from the Front Lines 2007.

Woolley, J. T. 2000. Using media-based data in studies of politics. *American Journal of Political Science*, 44: 156–73.

Yanow, D. and Schwartz-Shea, P. 2006. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. New York: M. E. Sharpe.

Yule, G. U. 1907. On the theory of correlation for a number of variables, treated by a new system of notation. *Proceedings of the Royal Society of London*, 79: 182–93.

Notes:

⁽¹⁾ Behavioralism was a complex phenomenon with many different currents. For proponents see Dahl (1961) and Eulau (1963; 1969); for histories see Farr, Dyrzek, and Leonard (1995), Farr and Seidelman (1993); and for a critique, Bevir (2008).

⁽²⁾ For a discussion and defense of why we use these specific words see our introduction to the *Oxford Handbook of Political Methodology*.

⁽³⁾ As we make clear in the “Introduction” to the stand-alone methods volume (Box-Steffensmeier, Brady, and Collier 2008) in the Oxford series of handbooks, we mostly deferred to the Goodin and Tilly (2006) handbook which made a wide-ranging contribution to this line of investigation.

⁽⁴⁾ We are vastly simplifying a very complicated topic: see Radcliff (1992).

⁽⁵⁾ For example: “lines of social cleavage due to differences in industrial and social status” or “the cleavage line between city and country” (Emerick 1910, 649–50).

⁽⁶⁾ For example: “there will not improbably appear a fresh cleavage when ‘social’ legislation reaches an active phase” (Mavor 1895, 505) or “Socialism seems destined to produce in the near future a perfectly new moral ‘line of cleavage’ in English society” (Webb 1889, 36).

⁽⁷⁾ “A decade of self-government under this liberal and democratic charter [the Weimar Constitution] has rather definitely established the lines of party cleavage and the general features of the party system ...” (Pollock 1929, 859).

If we think of societies being composed of people with sociodemographic characteristics X , who have issue positions I , and who belong to groups G , then political cleavages were variously described as differences in politically relevant characteristics (South versus North or farmers versus business), differences in political issue positions (pro-slavery versus anti-slavery or pro-silver versus anti-silver), or differences in political groups (Democrats versus Republicans). Neither the considerations for diagnosing a factor as political nor the exact way of measuring this was made clear.

⁽⁸⁾ In terms of the notation in the previous footnote, Rice showed that given distinct characteristics X_1 and X_2 , the value of $\text{Prob}(\text{Voting for } G|X_1)$ was very different from $\text{Prob}(\text{Voting for } G|X_2)$.

⁽⁹⁾ In fact, he elides the question by noting that “In reality narrow limits to inquiry are imposed by official models of grouping, recording, and reporting votes and voters, and by the possibilities of statistical analysis of records” (167). Therefore, a substantive problem becomes merely one for which data are available. In effect, Rice presumed that a characteristic X was political if for political parties G_1 and G_2 , we have $\text{Prob}(\text{Support } G_1|X)$ is very different from $\text{Prob}(\text{Support } G_2|X)$.

⁽¹⁰⁾ Defined as having an “ideal point” representing maximum utility from which utilities decline monotonically on each side from this point.

⁽¹¹⁾ Downs went on to argue that if the parties were mobile, then they would both converge to the middle of the space if ideal points were distributed somewhat like a normal distribution. Downs was apparently unaware of Black’s work, which would have led him to the stronger conclusion that the parties would converge to the median voter in the space no matter what the distribution of people’s ideal points. Hence, the line of cleavage would be at the median voter.

⁽¹²⁾ As with the Downsian model, the major concern of this paper was with the possibilities of equilibrium and not with the definition of cleavages.

⁽¹³⁾ These data are based upon a search of JSTOR for the *APSR* for the periods indicated. The search terms were

“interview or interviews or interviewing and not (survey or surveying).” Then each article found was examined by the authors to determine if multiple interviews (but not a survey) were used as the basis for the research or if the article advocated the use of such interviews. In most cases, it was easy to make this determination.

(¹⁴) Both of these studies were inspired by Lazarsfeld, Berelson, and Gaudet, *The People's Choice* (1944), which reported on a 1940 panel study in Erie County, Ohio.

(¹⁵) The American National Election Studies provides access to many other studies as well: <http://www.electionstudies.org/other_election_studies.htm>.

(¹⁶) World Values Surveys: <<http://www.worldvaluessurvey.org/>>; International Social Survey Programme: <<http://www.issp.org/>>; Comparative Study of Electoral Systems: <<http://www.issp.org/>>; Pew Global Attitudes Survey: <<http://pewglobal.org/>>; Voice of the People: <<http://www.voice-of-the-people.net/>>. For a listing of many others see: <http://www.gesis.org/EN/data_service/eurobarometer/handbook/index.htm>.

(¹⁷) Search on “Voteview” or go to Poole’s website: <<http://www.voteview.com/>>.

(¹⁸) Other scaling methods include Jackman (2001); Lewis and Poole (2004); and Clinton, Jackman, and Rivers (2004).

(¹⁹) Search on “Voteworld” or go to: <http://ucdata.berkeley.edu:7101/new_web/VoteWorld/voteworld/>.

(²⁰) In the chart, we double the number of articles for the 2000–4 period to make it comparable to the previous decadal periods. Hence, “endogeneity” or “endogenous” are extrapolated to appear in 150 total articles.

(²¹) Figure 48.4 only includes the data on selection bias, but a search for the words “omitted variable(s),” “confounding variable(s),” or “confounder(s)” yields almost exactly the same pattern. There are no mentions of these terms until the 1960s when there are four, then seven in the 1970s, nine in the 1980s, fifteen in the 1990s, and fourteen in the first half of the 2000s.

(²²) In effect, the question is how one can reconcile a metaphysically (or ontologically) deterministic world with the need for a probabilistic epistemology. Some authors, however, argue that the world is inherently probabilistic: See Suppes (1984).

(²³) Thus if C is cause and E is effect, a necessary condition for causality is that $\text{Prob}(E|C) > \text{Prob}(E|\text{not } C)$. Of course, this also means that the expectation goes up $\text{Exp}(E|C) > \text{Exp}(E|\text{not } C)$.

(²⁴) SUTVA means that a subject’s response depends only on that subject’s assignment, not the assignment of other subjects. SUTVA will be violated if the number of units getting the treatment versus the control status affects the outcome (as in a general equilibrium situation where many people getting the treatment of more education affects the overall value of education more than when just a few people get education), or if there is more communication of treatment to controls depending on the way assignment is done.

(²⁵) The observant reader will note that these authors make a causal claim about the power of an invention (in this case experimental methods) to further causal discourse.

(²⁶) We constructed variables for each word with a zero value if the word was not present in an article and a one if it was mentioned at least once. Then we obtained the ten correlations between pairs of the five variables with articles as the unit of analysis.

(²⁷) Each word appears in a different number of articles, but one or the other or both of the words “narrative” or “interpretive” appear in about 5.9% of the articles and the words “hypothesis” or “causal” or “causality” appear in almost one-third (31.3%). “Explanation” alone appears in 35.4% of the articles.

(²⁸) In 1980–4, the words “narrative” or “interpretive” were mentioned only 4.1% of the time in political science journals; in the succeeding five-year periods, the words increased in use to 6.1%, 8.1%, and finally 10.1% for 1995–9.

(²⁹) Brady was a founding member and early president of the Political Methodology Society. He was a co-principal

investigator (with PI Paul Sniderman and Phil Tetlock) of the Multi-Investigator Study which championed the use of experiments in surveys and which provided the base for the Time-sharing Experiments for the Social Sciences (TESS) program. He was present at the meeting convened by Jim Granato at NSF which conceived of the EITM idea, and he is a co-PI of one of the two EITM summer programs. Janet Box-Steffensmeier was an early graduate student member of the Political Methodology Society and a recent president. David Collier was the founding president of the APSA Qualitative Methods section, and the chair of CQRM's Academic Council.

Henry E. Brady

Henry E. Brady is Professor of Political Science and Public Policy at University of California, Berkeley. He received his Ph.D. in Economics and Political Science from MIT, and his areas of interest include Quantitative Methodology, American and Canadian Politics, and Political Behavior. He teaches undergraduate courses on political participation and party systems and graduate courses on advanced quantitative methodology.

David Collier

David Collier is Professor of Political Science at University of California, Berkeley and former President of the American Political Science Association. His fields are comparative politics, Latin American politics, and methodology. His latest book is *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, of which he is co-editor and co-author with Henry E. Brady.

Janet M. Box-Steffensmeier

Janet M. Box-Steffensmeier is Vernal Riffe Professor of Political Science and Sociology, Director of the Program in Statistics Methodology, Ohio State University.

