Journal of Theoretical Politics

Causation: A Unified Framework for the Social Sciences

John Gerring

The online version of this article can be found at:
http://jtp.sagepub.com/cgi/content/abstract/17/2/163

Published by:
⑤SAGE Publications
http://www.sagepublications.com

Additional services and information for *Journal of Theoretical Politics* can be found at:

Email Alerts: http://jtp.sagepub.com/cgi/alerts

Subscriptions: http://jtp.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 41 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://jtp.sagepub.com/cgi/content/refs/17/2/163

CAUSATION

A UNIFIED FRAMEWORK FOR THE SOCIAL SCIENCES

John Gerring

ABSTRACT

This paper offers four main arguments about the nature of causation in the social sciences. First, contrary to most recent work, I argue that there is a unitary conception of causation: a cause raises the probability of an event. This understanding of causation, borrowed from but not wedded to Bayesian inference, provides common semantic ground on which to base a reconstruction of causation. I argue, second, that rather than thinking about causation as a series of discrete types or distinct rules we ought to re-conceptualize this complex form of argument as a set of logical *criteria* applying to all arguments that are causal in nature (following the foregoing definition), across fields and across methods. Here, it is helpful to distinguish between the formal properties of a causal argument and the methods by which such an argument might be tested, the research design. Sixteen criteria apply to the former and seven criteria apply to the latter, as I show in the body of the paper. In summary, causation in the social sciences is both more diverse *and* more unified than has generally been recognized.

KEY WORDS ● causation ● epistemology ● methodology

Is causation in the social sciences unitary or plural? The unitary view, implicit in most work adopting a positivist epistemology (e.g. Hempel and Oppenheim, 1948), has recently been criticized by a number of writers. These writers claim that the unitary perspective betrays a narrow conception of causation and is not reflective of the broad range of causal arguments present in the fields of social science today. Instead, these writers suggest that the fields of the social sciences are characterized by a plurality of causal assumptions (epistemological, ontological, or logical) and an associated diversity of research designs. Causation is plural, not unitary.

The plural vision of causation has a long lineage. Aristotle divided the subject into four, quite different, types: formal causes (that into which an effect is made, thus contributing to its essence), material causes (the matter out of which an effect is fashioned), efficient causes (the motive force which made an effect), and final causes (the purpose for which an effect was produced). In the contemporary era, writers often distinguish between deterministic

and probabilistic causal arguments (Adcock, 2002; Goertz and Starr, 2002; Spohn, 1983; Waldner, 2002), 'correlational' causal analysis and the analysis of causal mechanisms or processes (Bhaskar, 1975/1978; Dessler, 1991; Elster, 1989; George and Bennett, 2005; Harre, 1970, 1972; Hedstrom and Swedberg, 1998: 7; Mahoney, 2001; McMullin, 1984; Ragin, 1987), 'top-down' and 'bottom-up' explanations (Glennan, 2000; Kitcher, 1989; Salmon, 1989),[1] and 'dependence' and 'production' causes (Hall, 2004).[2]

Other writers provide more differentiated typologies. Mario Bunge (1997: 412–13) identifies four types of causal explanation – covering law, interpretive, functional, and mechanismic. Covering law models involve a 'subsumption of particulars under universals'; interpretive causation focuses on the sense, meaning, or intention of an action; functional explanation focuses on the purpose (telos) of an action; and mechanismic causation focuses on 'the mechanism(s) likely to bring about the desired goal'. Henry Brady (2002) discerns four types of causation – a regularity theory associated with Hume, Mill, and Hempel, a counterfactual theory associated with the work of David Lewis (1973), a manipulation theory associated with the experimental tradition, and a mechanisms/capacities theory associated with the realist tradition in philosophy of science. Charles Tilly (2001) claims to have discovered five views of causal explanation, which he labels (a) skepticism, (b) covering law, (c) propensity, (d) system, and (e) mechanism and process. Even more differentiated typologies have been elaborated by philosophers (e.g. Sosa and Tooley, 1993).

Finally, one must take account of an ever-expanding menu of causal relationships. These include the *conjunctural cause* (also known as compound cause, configurative cause, combinatorial cause, conjunctive plurality of causes), where a particular combination of causes acting together to produce a given effect; *causal equifinality* (also known as a disjunctive plurality of causes, redundancy, or overdetermination), where several causes act independently of each other to produce, each on its own, a particular effect; the *cause in fact* (also known as actual cause), which explains a specific event as opposed to a class of events; the *non-linear* cause (e.g. a cause with a takeoff or threshold level); the *irreversible cause* (e.g. a cause with a ratchet effect); the *constant* cause, a cause operating continually on a given effect over a period of time; the *causal chain* (also known as causal path),

---

1. Stuart Glennan (2000: 1) explains, 'The top-down approach explains an event by showing it to be part of a larger nomological or explanatory pattern, while the bottom-up approach explains an event by describing the network of causes that are efficacious in bringing that event about' (see also Kitcher [1989] and Salmon [1989]).

2. The first understands causation in terms of 'counterfactual dependence between wholly distinct events'. The second understands as a cause any action that 'helps to *generate* or *bring about* or *produce* another event' (Hall, 2004).

where many intermediate causes lie between a structural X and an ultimate Y; and the *critical-juncture* or *path-dependent cause*, a cause or set of causes at a particular moment in time that has enduring effects.

Evidently, there are many ways to think about causation. Indeed, once we head down this analytic road it is not clear where we might stop or ought to stop. Causal relationships are, in principle, infinite in their diversity. There is potentially always some new way in which two relationships may co-vary or some new set of causal mechanisms between them. Causal frameworks are equally expansible; there are always new ways of conceptualizing the ways in which some class of variables might be understood. And theories of causation, while presumably not quite so open to multiplication, are still fairly open-ended.

It is equally evident that the type of causal argument one chooses to adopt is likely to affect one's choice of research design. 'Different types of causes require different approaches to empirical analysis' (Marini and Singer, 1988: 349). In Charles Tilly's (2001: 22) estimation,

> Behind many ostensibly theoretical disputes in political science lurk disagreements about the nature of valid explanations. Confrontations among advocates of realist, constructivists, and institutionalist approaches to international relations, for example, concern explanatory strategies more than directly competing propositions about how nations interact. Similarly, rampant debates about nationalism more often hinge on specifying what analysts must explain, and how, than on the relative validity of competing theories. Recent debates about democratization concern not only the choice of explanatory variables but also the very logic of explanation.

In this respect, causal pluralists offer an important corrective to a naïve positivism (see also Hall, 2003; Goertz and Starr, 2002; Ragin, 1987, 2000).

However, the pluralist view of causation, if we are to take it seriously, raises several difficulties. First, causal typologies such as those as sketched earlier may over-state the ontological, epistemological, and/or logical *different-ness* of causal explanations in the social sciences. Consider, for example, the distinction between causal arguments that are 'correlational' in nature and those that rely on the identification of causal mechanisms or processes. These are sometimes presented as distinct ways of understanding and demonstrating causal relationships, as previously noted (Bhaskar, 1975/1978; Dessler, 1991; Elster, 1989; George and Bennett, 2005; Harre, 1970, 1972; Hedstrom and Swedberg, 1998: 7; Mahoney, 2001; McMullin, 1984; Ragin, 1987). Stuart Glennan (1992: 50), in a widely cited article, argues explicitly that 'there should be a dichotomy in our understanding of causation' – between mechanism causes and correlation causes. Of course, much depends on how we choose to define the terms of this dichotomy. Let us suppose that correlations refer to covariational patterns between a cause and an effect and that mechanisms refer the connecting threads (pathways) between the purported cause and its effect. The question can then be posed

as follows. Are there causal explanations that feature only the associational patterns between an X and Y, without any consideration of what might link them together – or, alternatively, 'mechanismic' accounts that ignore patterns of association between cause and effect?

My sense is that such constrained forms of argument are relatively rare in the social science world. Granted, some correlational-style analyses slight the explicit discussion of causal mechanisms but this is usually because the author considers the causal mechanism to be clear and hence not worthy of explicit interrogation.[3] Similarly, a mechanismic argument without any appeal to covariational patterns between X and Y does not make any sense. The existence of a causal mechanism *presumes* a pattern of association between a structural X and an ultimate Y. Thus, to talk about mechanisms is also, necessarily, to talk about covariational patterns ('correlations'). Moreover, the suggested causal mechanism is itself covariational in nature since it presupposes a pattern of association between a set of intermediate variables. Granted, these patterns of intermediate association may not be directly observable; they may simply be assumed, based upon what we know about the world. Nonetheless, they might be probed and, in the best research designs, they are.

In any case, it seems overly simplistic and perhaps misleading to separate out correlational and mechanismic patterns of explanation if these are understood as dichotomous types. Similar difficulties apply to other typologies. As a result, the pluralist vision is suspect as a description of the facts of the case, i.e. as a description of the folkways of social science.

Second, and perhaps more importantly, whatever diversity of causal logic characterizes the contemporary social sciences, there is little *profit* in a plural account of causation. If causation means different things to different people then, by definition, causal arguments cannot meet. If A says that X_1 caused Y and B retorts that it was, in fact, X_2 or that Y is not a proper outcome for causal investigation, and they claim to be basing their arguments on different understandings of causation, then these perspectives cannot be resolved; they are incommensurable. Insofar as we value cumulation in the social sciences, there is a strong prima facie case for a unified account of causation.[4]

Consider our previous discussion of correlational and mechanismic accounts of causation. What are we to make of a situation in which one writer supports an argument based on $X{:}Y$ covariational patterns while another supports a contrary argument based on the existence of causal

---

3. This may or may not be a valid assumption; sometimes, writers are rather cavalier about the existence of causal mechanisms. My point is simply that in ignoring explicit discussion of causal mechanisms they are in no way challenging the significance of mechanisms in causal explanation.

4. This expresses the spirit of Brady's (2002) analysis but not of most others cited earlier.

mechanisms? Is this argument adjudicable if neither recognizes the relevance of the other's approach? Commonsense suggests that both X:Y correlations as well as mechanisms connecting X and Y are important elements of any causal argument. Thus, although we can easily imagine situations in which these two sorts of evidence lead to different conclusions, it does not follow that they are incommensurable. From a normative perspective, the pluralist vision of distinct causal types is not a useful one for the conduct of social science. Unity, not different-ness, should be the goal of any causal methodology.

Thus, my general argument: now that writers have succeeded in taking causation apart, the members of the social science community have strong reasons for trying to put it back together. The crucial caveat is that this unified account of causation must be sufficiently encompassing to bring together all (or most) of the arguments that we refer to as causal in the social sciences. Unity is useful but not if achieved by arbitrary definitional fiat.

I shall argue that there is an underlying concept of causation shared by all (or most) protagonists in this debate. The core, or minimal, definition of causation held implicitly within the social sciences is that a cause raises the probability of an event occurring. This understanding of causation, which is borrowed from but not wedded to Bayesian inference, provides common semantic ground on which to base a reconstruction of causation.

I shall argue, second, that rather than thinking about causation as a series of discrete types or distinct rules, we ought to re-conceptualize this complex form of argument as a set of logical *criteria* applying to all arguments that are causal in nature (following the foregoing definition), across fields and across methods. Criteria, following Stanley Cavell (1979: 9), refer to 'specifications a given person or group sets up on the basis of which . . . to judge . . . whether something has a particular status or value'. In this case, we are looking to discover the shared norms that govern activity, implicitly or explicitly, in the community of the social sciences. Criteria may thus be understood as providing the normative ground for causal argument. Since all social science activity is norm-bound, we are able to capture the unity and diversity of causal argument by paying close attention to these norms, most of which are regarded as 'commonsense' within the disciplines of the social sciences.[5]

I shall argue, third, that, in coming to grips with causation, it is helpful to distinguish between the formal properties of a causal argument and the

---

5. Other attempts to specify the desiderata of causal argument, upon which this effort builds, can be found in Eckstein (1975: 88), Hempel (1991: 81), King et al. (1994: Ch. 3), Kuhn (1977: 322), Lakatos (1978), Laudan (1977: 68; 1996: 18, 131–2), Levey (1996: 54), Marini and Singer (1988), Przeworski and Teune (1970: 20–3), Simowitz and Price (1990), Stinchcombe (1968: 31), Van Evera (1997: 17–21), and Wilson (1998: 216). For further discussion of the criterial approach to social science methodology, see Gerring (2001: Chs 1–2).

methods by which such an argument might be tested. The two questions, *What are you arguing?* and *How do you know it is true?*, are logically distinct and call forth different criteria of adequacy.

My final argument, and the core of the paper, concerns the specific criteria that are commonly applied to causal argumentation and proof. I argue that 16 criteria apply to the formal properties of causal argument and eight apply to the choice of research design (the task of demonstration or proof). It will be argued, therefore, that causation in the social sciences is both more diverse *and* more unified than has generally been recognized. It is more diverse insofar as the criteria applied to causal argumentation surpass the implications of even the broadest of the causal typologies sketched earlier. There are multiple dimensions by which causal arguments are rightly judged. It is more unified insofar as these criteria apply across existing fields, methods, ontologies, and epistemologies. This argument does not dispute the diversity of empirical approaches evident in the social sciences today, including textual, ethnographic, experimental, statistical, and formal methods of analysis (often grouped into three large camps – interpretivist, behavioralist, and rational choice). What it does is to call attention to the remarkable commonalities that underlie these apparently heterogeneous approaches.

The argument proceeds in three parts. First, I discuss various definitional properties of 'cause'. Second, I lay out various criteria pertaining to the formal properties of causal argument. Third, I show the criteria pertaining to research design (our empirical investigation of causal relationships). Finally, I summarize my conclusions and discuss its implications with respect to the current methodological wars in the social sciences.

Three important caveats must be noted before I begin. First, this discussion concerns causation in the *social sciences*, not in everyday speech or in philosophy. I presume that the language region of social science is somewhat distinct – though not, of course, totally disconnected – from everyday language and from the language region of philosophy. Moreover, I presume that the practical needs of social science (specifically with reference to a unitary understanding of causation) are not necessarily duplicated in everyday speech or in philosophic analysis. Thus, I cite ordinary language and philosophic analysis sparingly in the following discussion. It is relevant only when it bears on the usage patterns and disciplinary practices of the social sciences.

Second, this is a *normative* discussion. I am interested in developing a framework that might improve the understanding of causation in the disciplines of the social sciences. More precisely, I build on existing practices with an aim to provide a more coherent and concise explanation of already-existing practices. This is quite different from the sort of methodological work that seeks simply to describe certain practices, to discuss trends, and to elucidate patterns.

Finally, I do not broach *ontological* questions pertaining to causal analysis. Of course, there is an implicit 'realist' ontological assumption in any endeavor that takes empirical analysis seriously. One is assuming that there is a reality out there and that we have at least a fair chance of figuring out what it is. And there are presumably other ontological assumptions embedded in the ways in which I (and other social scientists) have chosen to carve up and describe the empirical world. However, the argument rests on what might be called pragmatic grounds. Given a field of endeavor called social science which is understood to strive for knowledge of the humanly created world that is systematic, empirical, falsifiable, replicable, and, in some sense, 'objective', and given our need to understand that world (for policy and other purposes), how best should the question of causation be understood? This is the pragmatic ground upon which the argument rests (Gerring, 2001: epilogue). So understood, we may bracket questions about the ultimate nature of reality. To be perfectly clear, I agree that ontological presuppositions do, sometimes, guide social-scientific inquiry (e.g. Gerring, 2004; Hall, 2003). But I do not agree that we should cultivate these differences in the social sciences, for it would not advance our mission to do so.

'Cause': A Minimal Definition

The first step in restoring a unitary conception of causation is to arrive at a unitary definition of this polysemic word (at least within social science contexts). We need a definition that differentiates 'cause' from other near-synonyms but is also substitutable for all usages of the word. I shall refer to this as a minimal definition, for it constitutes the most general definition of that term by offering a minimal set of defining attributes (Gerring and Barresi, 2003).

Minimally, causes may be said to refer to events or conditions that *raise the probability* of some outcome occurring (under ceteris paribus conditions). X may be considered a cause of Y if (and only if) it raises the probability of Y. This covers all meanings of the word cause in social science settings, including those discussed at the outset of this essay. While it might seem as if this minimal definition is prejudicial to certain types of causal argument – e.g. probabilistic causes and Bayesian frameworks of analysis – this would be an incorrect conclusion. A minimal definition encapsulates all others without granting priority to any particular type of cause or causal argument. Evidently, a deterministic cause (understood as being necessary and/or sufficient to an outcome) is also a probabilistic cause. Thus, to say

that all causes raise the probability of an outcome is to include deterministic causes within the rubric of causation.

It should be clear that the phrase 'raise prior probabilities' presumes that the actual (ontological) probability of an event will be increased by X, not simply one's predictive capacity. The latter might be affected by correlative relationships that are not causal in nature. To be causal, the cause in question must *generate*, *create*, or *produce* the supposed effect.

With this minimal definition of causation, we can now address the meatier question of how to understand variation among causal arguments in the social sciences. Recall that my argument is opposed to two prevailing views – the positivist view (associated with Hempel) that all causes can be understood according to covering-laws, and the pluralist view (associated with a variety of authors cited earlier) that causal arguments are manifold and, by extension, incommensurable (insofar as the choice of causal argument plays a determinative role in causal evaluation). In contrast, I argue that causation is best understood as a *criterial* venture. Causal arguments strive to achieve many objectives but these objectives apply more or less uniformly to all causal argument. Hence, plurality and unity co-exist in the methodology of causal investigation.

Formal Criteria of Causal Argument

Let us begin by looking at the formal properties of causal argument. Later, we shall examine properties pertaining to research design (the application of evidence to a causal argument). The formal properties of causal argument may be summarized according to 16 criteria: (1) *specification*, (2) *precision*, (3) *breadth*, (4) *boundedness*, (5) *completeness*, (6) *parsimony*, (7) *differentiation*, (8) *priority*, (9) *independence*, (10) *contingency*, (11) *mechanism*, (12) *analytic utility*, (13) *intelligibility*, (14) *relevance*, (15) *innovation*, and (16) *comparison*, as adumbrated in Table 1. It is these criteria that distinguish a good causal argument from a poor or uninteresting one. (I employ argument, proposition, theory, and inference more or less interchangeably in the following discussion.)

1. Specification

One must know what a proposition is *about* in order to evaluate its truth or falsity. As Durkheim (1895/1964: 34) remarked a century ago, 'a theory . . . can be checked only if we know how to recognize the facts of which it is intended to give an account'. The first step down the road to empirical

Table 1. Causal Propositions: Formal Criteria

1. *Specification* (clarification, operationalization, falsifiability)
   (a) What are the positive and negative outcomes (the factual and the counterfactual, or the range of variation) that the proposition describes, predicts, or explains?
   (b) What is the set of cases (the population, context, domain, contrast-space, frame, or base-line), that the proposition is intended to explain?
   (c) Is the argument internally consistent (does it imply contradictory outcomes)?
   (d) Are the key terms operational?

2. *Precision*
   How precise is the proposition?

3. *Breadth* (scope, range, domain, generality, population)
   What range of instances are covered by the proposition?

4. *Boundedness* (non-arbitrariness, coherence)
   Is the specified population logical, coherent? Does the domain make sense?

5. *Completeness* (power, richness, thickness, detail)
   How many features, or how much variation, is accounted for by the proposition? How strong is the relationship?

6. *Parsimony* (economy, efficiency, simplicity, reduction, Ockham's razor)
   How parsimonious is the proposition?

7. *Differentiation* (exogeneity) (*antonym:* endogeneity)
   Is the X differentiable from the Y? Is the cause separate, logically and empirically, from the outcome to be explained?

8. *Priority*
   How much temporal or causal priority does X enjoy vis-a-vis Y?

9. *Independence* (exogeneity, asymmetry, recursiveness) (*antonyms:* endogeneity, reciprocality, symmetry, feedback)
   How independent is X relative to other Xs, and to Y?

10. *Contingency* (abnormality)
    Is the X contingent, relative to other possible Xs? Does the causal explanation conform to our understanding of the normal course of events?

11. *Mechanism* (causal narrative)
    Is there a plausible mechanism connecting X to Y?

12. *Analytic utility* (logical economy) (*antonyms:* idiosyncrasy, ad-hocery)
    Does the proposition fit with what we know about the world? Does it help to unify that knowledge?

13. *Intelligibility* (accessibility)
    How intelligible is the proposition?

14. *Relevance* (societal significance)
    How relevant is the proposition to a lay audience or to policymakers? Does it matter?

15. *Innovation* (novelty)
    How innovative is the proposition?

16. *Comparison*
    Are there better explanations for a given outcome? Is the purported X superior (along criteria 1–15) to other possible Xs? Have all reasonable counter-hypotheses been explored?

truth, therefore, is *specification*.[6] Specification in causal argument involves, first, the clarification of a positive and negative outcome – a *Y* and a not-*Y*, a factual and a counterfactual – or a *range* of variation on *Y* (high–low, weak–strong) that the proposition is intended to describe, predict, or explain. Without such variation, the statement cannot be disproven. One must also consider the definition of *cases* to which a causal proposition refers. A proposition is not fully specified until it has specified not only the positive and negative outcomes but also a set of cases (a context, background, baseline, or contrast-space) that it purports to explain. All arguments work by separating a foreground from a background, and the author of a proposition accomplishes this by constructing a particular context – spatial and temporal – within which an explanation can be proffered. This central fact has been variously labeled the background, baseline, breadth, contrast-space, counterfactual, frame, population, or scope of an explanation.[7] Specification also demands *internal consistency* in the proposition. If the argument shifts (e.g. from a descriptive to a causal proposition), if the outcomes are unstable, or if the definition of the population changes, then we cannot consider a proposition to be properly specified. Effectively, we have not one proposition but several. These several propositions must be reconciled, or separated (so that they do not contradict one another), before we can identify the truth or utility of any one of them. Specification, finally, demands that key concepts be fully *operationalized*. One cannot have a fully specified proposition without fully operational key concepts; the one is an extension of the other.

2. Precision

In addition to clarity (specification), we also desire causal arguments to be precise. Naturally, the level of precision will vary; indeed, there is no limit, in principle, to the degree of precision a proposition might acquire (just as there is no limit, in principle, to the number of decimal points in a calculation). Ceteris paribus, a precise proposition will be looked upon as a more useful one.

---

6. One might also label this criterion *falsifiability*, and certainly the Popper oeuvre (e.g. Popper 1934/1968) is relevant to our discussion. See also Garfinkel (1981).

7. See Collier and Mahoney (1996: 67), Garfinkel (1981), Goodman (1965: Ch. 1), Hawthorn (1991), Putnam (1978: 41–5; 1987: 3–40), Ragin and Becker (1992), and van Fraassen (1980: 115–57). For most intents and purposes, context is also equivalent to the breadth of a proposition. To inquire about the context of an explanation is to inquire about the class of cases – real or imagined – to which it applies. Yet, the context of a proposition is rarely understood in this technical and precise manner (as a set of cases); usually, it is a more informal understanding about the background to a particular explanation.

3. Breadth

If the fundamental purpose of causal argument is to tell us about the world, then it stands to reason that a proposition informing us about many events is, by virtue of this fact, more useful than a proposition applicable to only a few. I shall refer to this desideratum as *breadth*, or generality. One wishes, when forming propositions, to capture as many events as possible. Let us note parenthetically that what idiographically-inclined writers cringe at in work of great breadth is not breadth per se but rather the sacrifice such breadth may bring along other criteria (accuracy, precision, strength, and so forth). It seems possible to conclude, therefore, that a proposition with a larger explanandum is always better, ceteris paribus.

4. Boundedness

An inference should be properly bounded, i.e. its scope should be neither too big nor too small. This speaks to the logic of the causal argument and the range of possible phenomena to which it might plausibly refer. An arbitrarily bounded inference, by contrast, is not convincing. Usually, this concerns inferences that are drawn in a manner that seems too restrictive, thus 'defining out' cases that seem otherwise relevant. Studies in the rational choice genre have been accused of this sort of inference-gerrymandering (an 'arbitrary domain restriction' [Green and Shapiro 1994: 45]). Whether 'rational choicers' are guilty of this sin is not important for present purposes. The point to remember is that the specification of the population is only the first step on the road to a meaningful empirical proposition. We must also make sure that the population (the breadth of an inference) 'makes sense'. (Granted, this is to some extent an empirical venture, as indeed are virtually all the 'formal' properties of an inference discussed herein.) If there are questions about the appropriate boundaries of a proposition, there will also be serious questions about its validity.

5. Completeness

The criterion of strength refers to the extent to which a causal proposition explains its intended target (the explanandum). One might also call it power, depth, informativeness, richness, thickness, or detail. Indeed, the virtues of strength have been trumpeted under many names – e.g. 'thick description' (Geertz, 1973), 'configurative analysis' (Heckscher, 1957: 46–51, 85–107; Katznelson, 1997), 'contrast of contexts' (Skocpol and Somers

1980), and 'holistic analysis' (Ragin 1987).[8] In statistical terms, it is the variation explained by a given model. Completeness, like breadth, stems from the goal of social science to tell us as much as possible about the empirical world. In this case, however, it is the population chosen for study not some broader set of phenomena, with which one is concerned.

6. Parsimony

Like a lever, a useful proposition carries great weight with a moderate application of force. It is powerful and its power derives from its capacity to explain a great deal with a minimal expenditure of energy. Such a proposition is *parsimonious*.[9] The argument for parsimony, as for other criteria, is not ontological; it does not presume that simplicity conforms to the natural order of things, as some have argued for the natural sciences.[10] It is, rather, a pragmatic argument. We need to bring knowledge together in reasonably compact form in order for that knowledge to serve a useful purpose. Reduction is useful.

7. Differentiation

A cause must be differentiable from the effect it purports to explain. This seems self-explanatory. Even so, differentiation is often a matter of degrees. To begin with, Xs and Ys are always *somewhat* differentiated from one another. The perfect tautology – e.g. 'The Civil War was caused by the Civil War' – is simply nonsense. Yet, one occasionally hears the following sort of argument: 'The Civil War was caused by the attack of the South against Fort Sumter'. This is more satisfactory. Even so, it is not likely to strike readers as a particularly acute explanation. Indeed, there is very little 'explanation' occurring here, since the X is barely differentiated from the Y (the attack against Fort Sumter is generally considered as *part of* the Civil War). Consider a second example, this one classical in origin. To say that this man (X) is father to this child (Y) is to infer that the father *caused* the child to exist; he is a necessary (though not, of course, sufficient)

---

8. See also Almond and Genco (1977/1990), Collier (1993), Eidlin (1983), Feagin et al. (1991), and Hirschman (1970).

9. For work on the inter-related questions of reduction, simplicity, and parsimony, see Friedman (1972), Glymour (1980), Hesse (1974), King et al. (1994: 20), Popper (1934/1968), Quine (1966), Sober (1975, 1988).

10. King et al. (1994: 20, 104) take this view of parsimony and reject it on those grounds. If interpreted as a *pragmatic* norm, however, it might not be rejected by the authors. See, e.g., their discussion of the importance of 'leverage' ('explaining as much as possible with as little as possible' [p. 29]).

cause of the child. We are less impressed, however, by the argument that a *foetus* is the cause of a child or a child the cause of an adult. There is something wrong with these formulations, even though X clearly raises the probability of Y. What is wrong is that there is no demonstrable differentiation between these Xs and Ys; they are the same object, observed at different points in time. In short, we have stipulated a causal relationship to a 'continuous self-maintaining process' and this violates the precept of differentiation.[11]

8. Priority

Causes, Hume thought, must be prior to their effects. Certainly, one might say, a cause may not arrive *after* its effect.[12] Yet, it is not clear that causes are always, or necessarily, prior. If two cars collide, it is not clear that either car came 'first'; and even if it were the case, it is not clear that this would give the prior car greater claim to causality. Yet, from a pragmatic perspective, one might point out that it is more *useful* to identify factors that are prior to the event to be explained (Miller, 1987: 102–4). Consider the following path diagram.

$$X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \rightarrow Y$$

We are apt to consider X_1 to be *the* cause and causal factors X_{2-4} intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to X_1. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are 'ultimate' or 'fundamental'.

Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this

---

11. Marini and Singer (1988: 364). We should be clear that differentiation is not the same as *priority*. It is possible for a cause to be quite close to an effect – spatially or temporally – and still be clearly differentiated. Causal arguments hinging on human motivations usually have this quality. So long as we have an adequate indicator of a person's motive, and a separate indicator of her actions, we shall be satisfied that differentiation between X and Y has been achieved.

12. The apparent exception to this dictum – the anticipated cause – is not really an exception at all, I would argue. If person A alters her behavior because of the expected arrival of person B, one could plausibly claim that person B was causing person's A's behavioral change. However, it would be more correct to say that person B – the causal agent – was already present, in some sense, or that person A's behavior change was stimulated by an expectation that was, in any case, already present. (This sort of question is debated with respect to functional arguments in social science.)

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

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

9. Independence

The search for causal explanation is also a search for causes that are independent. Two species of independence must be distinguished. The first refers to the independence of a cause relative to other causes.[13] A cause which is co-dependent, one might say – a cause whose effect is dependent upon the existence of a range of other factors – is less useful than a cause that acts alone. (Of course, all causal arguments depend upon a background of necessary conditions but these may be fairly obvious and ubiquitous. If so, we can afford to ignore them and can claim independence for the cause that acts alone.) The second, and more usual, meaning of independence refers to the

---

13. The research design analogue would be collinearity (discussed in Chapter 8). Yet, here we are dealing only with the formal properties of causes – i.e. with causal relationships as they really are – rather than with matters of proof or demonstration.

independence of a cause relative to the outcome that is being explained. Dependence, in this sense, means that Y has a reciprocal effect on X, and is often referred to as symmetry, feedback, circularity, or endogeneity.[14]

Causal differentiation, priority, and independence are apt to co-vary. An X that is high on one dimension is apt to enjoy a similar status on the other: the further away X is from Y, the easier it may be to distinguish X from Y, and the less likely it is that Y will affect X. But these three criteria are nonetheless distinctive. Sometimes, for example, a cause with little priority is differentiated from, and in dependent of, its effect.

10. Contingency

Proper explanations seek out the most contingent cause from a field of possible factors. Thus the so-called cause is distinguished from background conditions – those factors which constitute the 'normal course of events'.[15] One can think about the logic of this selection criterion in the following way. A causal argument necessarily assumes a good deal about the world – namely, that it continues to revolve in about the same way (or with about the same rate and direction of variation) over the temporal scope of the inference. This is often referred to as the ceteris paribus assumption. Within this structure, one or two things seem unusual, unexpected. These are the contingent factors to which we commonly assign causal status.

Consider the proposition, common among journalists and popular commentators, that we do not have better government in Washington because politicians are more interested in getting re-elected than in solving problems of public policy. In other words, self-interested behavior is the cause of our current mess. This may be true, in the sense that a wholesale change of perspective on the part of the Washington establishment with regards to their re-election might lead to significant policy changes. But does this explanation identify the most contingent element in this causal relationship? Can advocates of this position think of other examples where members of

---

14. See King et al. (1994: 94, 108, 185) for a mathematical treatment of this problem, and Mackie (1974: Ch. 7) for a philosophical treatment.

15. The concept of *contingency* reiterates the same basic idea expressed by other writers under a variety of rubrics – 'abnormality' (Hart and Honore, 1959: 32–3), the 'normal course of events' (Hart and Honore, 1966: 225), 'in the circumstances' (Marini and Singer, 1988: 353), a 'causal field' (Mackie, 1965/1993: 39), and 'difference in a background' (Einhorn and Hogarth, 1986). See also Holland (1986), Taylor (1970: 53). All point to the task of differentiating a foreground (the cause) from a background (the structure, against which the causal relationship can be understood) by identifying that causal feature which is most contingent.

a legislature have decided, as a group, to forget about their re-election pros-
pects and simply pursue what they considered to be right for the nation?
The causal proposition of self-interest violates our sense of normality; it con-
fuses a more or less permanent background condition (self-interest) with a
contingent cause.

The principle of contingency applies equally to counterfactuals. If we are
interested in the outbreak of the Civil War, we do not want to hear about the
possible poisoning of Abraham Lincoln during his 1860 campaign. (Had he
been poisoned, we might speculate, the Republicans might not have won the
presidential campaign, which might have averted the secession of the South
from the union.) This is so far outside the normal course of events (*normally*
people are not poisoned) that it would be downright silly of any historian to
mention it. One can think of numerous other examples of things that, were
they to have occurred (or *not* occurred), would have 'altered the course of
history', as the phrase goes. By the same token, to say that the course of
Reconstruction was altered by Lincoln's assassination is an inference that
seems very real and very useful. Precisely, one might add, for the same
reasons: because assassination is such an *unusual* event (Lieberson, 1985:
Ch. 3).

'All possibilities for a world', writes Geoffrey Hawthorn (1991: 158),
'should . . . start from a world as it otherwise was. They should not require
us to unwind the past. And the consequences we draw from these alternatives
should initially fit with the other undisturbed runnings-on in that world.
Neither the alternative starting points nor the runnings-on we impute to
them should be fantastic.' Thus, we arrive at the following formulation: in
identifying causes we look to find those inferences that do least violence to
the normal, expected course of events. The more normality a given inference
assumes (in light of what we know of the world and of a given era or context),
the more readily we bestow causal status on that inference.

This means that for actually occurring events, *contingency* is prized: we
want to identify the contingent event amidst the broad current of non-
contingent events that constitute the normal course of history and human
behavior. Lincoln's assassination is such an event. By contrast, with counter-
factual events – events not actually occurring – we want to find those that
are *least* contingent (most normal). Thus, causal arguments based upon
Lincoln's actual assassination (1865) are a lot more meaningful than causal
arguments about his non-assassination (1860).

11. Mechanism

In order to be at all convincing, a causal proposition must be accompanied
by an explanation of the mechanism, or mechanisms, that connect a putative

cause with a purported effect.[16] In the hypothetical example of night and day, we lack such a mechanism: we cannot explain how night can generate day. Many correlative relationships in social science mirror this problem. To be sure, the distinction between a 'mechanism' and a 'cause' is a blurry one. All mechanisms might also be regarded as causes. Otherwise stated, a number of mechanisms link any cause with its effect (the regress is infinite); as soon as these mechanisms are identified, they may be regarded as intermediate causes (King et al., 1994: 86). Generally, we choose to explore this infinite chain of causation only so far as is necessary to demonstrate the plausibility of a given causal relationship. (Certain things are always taken for granted in any causal explanation, as discussed later.) It is this at this level that we identify intermediate causes as mechanisms and thereby cease to pursue the infinite causal regress. Thus, within the rubric of any causal argument, there is always a causal mechanism, explicit or implied. Even in the most perfect experimental context where a causal relationship between two entities seems validated, one is unlikely to feel that an argument is clinched until a causal mechanism has been identified. Indeed, it is not possible to fully test a causal hypothesis without some notion of the mechanism that might be at work, for all experiments are premised on certain assumptions about causal mechanisms. These assumptions identify the factors that must be held constant and those that should be manipulated.

12. Analytic Utility

No causal argument of any sort (indeed, no argument of any sort) could be made without assuming a good deal about how the world works. Some of these assumptions are commonsensical, others more sophisticated ('theoretical'). Some views of the world are held with a high degree of certainty, others are held with suspicion or pending further evidence to the contrary. The point is that we are more inclined to accept a causal argument if it dovetails with current understandings of the world, imperfect and uncertain as that knowledge may be. Indeed, part of the task of causal explanation involves showing how an explanation fits with the order of things, as presently understood. Causal arguments rest delicately upon a skein of truisms and theories. It is this background knowledge, the knowledge that makes causal explanation possible at all, that composes the non-empirical part of any causal argument. I refer to as the analytic utility of a proposition, an idea closely connected to the 'coherence' theory of truth.[17]

---

16. Blalock (1964), Hedstrom and Swedberg (1998), Little (1998), Mahoney (2001), Marini and Singer (1988: 379), Salmon (1984).

17. Kirkham (1992), Laudan (1996: 79).

In addition to general knowledge, we must also consider the more specialized realms of knowledge to which a proposition might belong. Often, a new proposition is part-and-parcel of a research tradition or a theory of a higher order of magnitude. This greatly enhances the logical economy of a field. By contrast, the proposition that sits by itself in a corner is likely to be dismissed as 'ad hoc' or 'idiosyncratic'. It does not fit with our present understanding of the world. It refuses to cumulate. It has little analytic utility. Of course, deviant propositions may be extremely useful in the long run. Indeed, the first sign of breakdown in a broad theory or paradigm is the existence of observations, or small-order theories, that cannot easily be made sense of. Yet, until such time as a new theory or paradigm can be constructed, it may truly be said of the wayward proposition that it is not useful, or less useful.[18]

In sum, the utility of a proposition cannot be understood as a purely inductive exercise, in communication with the empirical facts and in contrast with other alternative hypotheses (criterion 16). It must also be situated within a larger theoretical framework. James Johnson (2003) refers to this, following Larry Laudan (1977: 45), as 'conceptual problems' in theory formation. In terms of the intellectual history of the discipline, we may think of this more capacious and ambitious understanding of theory as an attempt to transcend the myopic, just-the-facts approach associated with early work in the behavioral mode, where one simply tested isolated hypotheses without self-conscious attention to theory-building.[19]

13. Intelligibility

It is difficult to imagine a relevant work that is not also, at least minimally, understandable. Indeed, intelligibility may be one of the chief criteria distinguishing the social sciences from the humanities and the natural sciences. The subjects of social science are different in that they require understanding and/or specific decisions on the part of the lay public. This means that they must be understood by lay citizens or, at the very least, by policy-makers.[20]

---

18. Comte (1975: 73), Einstein (1940/1953: 253), Friedman (1974), Homans (1967; cited in Rule 1997: 176), Mach (1902/1953: 450–1), Mill (1843/1872: 143–4), Wilson (1998: 291).

19. For recent defenses (implicit or explicit) of 'barefoot empiricism', see Green and Shapiro (1994), Van Evera (1997).

20. 'The economist who wants to influence actual policy choices must in the final resort convince ordinary people, not only his confreres among the economic scientists', notes Gunnar Myrdal (1970: 450–1).

14. Relevance

No matter how virtuous a proposition may be on other criteria, if it cannot pass the *So what?* test it is not worth very much. Propositions, large or small, have various levels of relevance. There are some things which, however true they are, we are not bothered to argue about. Here, I am not talking about relevance within a theoretical or classificatory framework (which I have called analytic utility) but rather relevance to the lay public, an issue I have addressed elsewhere (Gerring, forthcoming). Similarly, we are likely to identify as significant a causal factor which 'it is in our power to produce or prevent, and by producing or preventing which we can produce or prevent that whose cause it is said to be', for such a causal factor will be more relevant to human concerns than a causal factor that lies beyond our control.[21]

15. Innovation

'An author is little to be valued', says Hume (1985: 254) in his characteristically blunt fashion, 'who tells us nothing but what we can learn from every coffee-house conversation'. Indeed, we should like a proposition to be not just true but also new, or reasonably so. To be sure, we need to know if things that were once true are still equally so. There *is* a point to re-testing old hypotheses on new data. But there is less point to such a study if the result is entitled *Plus ça change . . .* The more a study (or a single proposition) differs from present views of a given subject – either in empirical findings or in theoretical framework – the greater its level of innovation and, ceteris paribus, its general utility.

16. Comparison

Causal demonstration is conducted to some extent by a process of *comparison*. A causal argument is falsified or verified to the extent that one causal story is demonstrated to be superior to others that might be constructed around the same event(s). Indeed, one places little faith in a causal conjecture based upon the investigation of a single cause. Nor have we much faith in a causal argument that eliminates other possible causes but does not give these causal hypotheses a fair shake – the 'straw man' argument. If, however, the writer has scrupulously investigated other possible causes for the event – as indicated by commonsense, by secondary work on a subject, or the writer's

---

21. R.G. Collingwood, cited in Garfinkel (1981: 138). See also Cook and Campbell (1979: 25), Gardiner (1952/1961: 12), Gasking (1955), Harre and Madden (1975), Suppes (1970), von Wright (1971), Whitbeck (1977).

own best guess – and none of these causes seem as secure, then we shall be much more confident in proclaiming X to be truly a cause of Y.[22]

Criteria of Demonstration or Proof (Research Design)

The debate over causation concerns not merely formal properties of an argument but also methods of proof or demonstration. Assumptions about causation and approaches to the empirical world are inextricably linked. Thus, we are led to consider not only the formal properties of causal argument but also the adjoining problem of induction (also known as the problem of confirmation or empirical testing), the probative elements of causal argument.

   Recall that from the pluralist perspective different kinds of causal argument call forth correspondingly different approaches to research design. In contrast, I shall argue that all attempts to prove or demonstrate a causal argument are subject to the same set of desiderata. Research design is, therefore, understood as a criterial venture rather than a series of rules or specific methods.[23] Seven factors characterize goodness in research design across fields and methods in the social sciences: (1) *plenitude*, (2) *comparability*, (3) *independence*, (4) *representativeness*, (5) *variation,* (6) *transparency*, and (7) *replicability*, as summarized in Table 2.[24]

1. Plenitude

All knowledge is comparative. It follows that the more comparative reference points one has at one's disposal, the better one can test the veracity of a given proposition. The accumulation of comparative reference points constitutes

---

22. The criterion of *comparison*, as I have called it, might also be labelled 'inference to the best explanation' (Harman, 1965), the 'method of residues' (Mill, 1843/1872: 259), 'abduction', 'the method of hypothesis', 'hypothetic inference', 'the method of elimination', 'eliminative induction', or 'theoretical inference' (most of the foregoing are discussed in Harman [1965]). See also Campbell (1988: Ch. 6), Campbell and Stanley (1963), Cook and Campbell (1979: 23), Marini and Singer (1988: 386), Miller (1987: Ch. 4), Neuman (1997: 51), Popper (1969), and Stinchcombe (1968: 20–2).

23. For a review of the 'rulebook' approach to social science methodology see Gerring (2001: 9–12).

24. Many of these features can be viewed as characteristics of the experimental research design, where the factor of interest is arbitrarily manipulated so as to provide pre-/post-comparisons and perhaps experimental/control comparisons.

Table 2. Causal Propositions: Criteria of Research Design

1. *Plenitude* (evidence)
   How many cases? How large is the sample (*N*)?

2. *Comparability* (equivalence, unit homogeneity, cross-case validity) (*antonym:* uniqueness)
   (a) *Descriptive comparability* (conceptual validity): How comparable are the *X*s and the *Y*?
   (b) *Causal comparability:* How similar are the cases with respect to factors that might affect *Y* or the *X:Y* relationship of interest? Finally, can any remaining dis-similarities be taken into account (controlled, modelled)?

3. *Independence* (*antonyms:* autocorrelation, Galton's problem, contamination)
   How independent are the cases with respect to factors that might affect *Y* or the *X:Y* relationship of interest? Can any remaining interdependencies be taken into account (controlled, modelled)?

4. *Representativeness* (external validity) (*antonyms:* sample bias, selection bias)
   Are the cases representative of the population with respect to all factors that might affect *Y* or the *X:Y* relationship of interest? Can any remaining un-representative elements be taken into account (controlled, modeled)?

5. *Variation* (variance)
   Do the cases offer variation (a) on *Y*, (b) on relevant *X*s, (c) without collinearity, and (d) within a particular case(s)?

6. *Transparency* (process-tracing)
   Does the research design offer evidence about the process (i.e. the intermediate factors) by which *X* affects *Y*?

7. *Replicability* (reliability)
   Can the research design be replicated? Are the results reliable?

evidence and more evidence is better (all other things being equal). For this reason, plenitude in a sample is desirable: the more cases one has to demonstrate a posited causal relationship, the more confidence one is likely to place in the truth of that proposition.[25] One case is better than none. Indeed, it is a quantum leap, since the absence of cases means that there is no empirical support whatsoever for a proposition. Two cases are better than one, and so forth. The same logic that compels us to provide empirical support for our beliefs also provides the logic behind plenitude in research design.

A large sample may also help in specifying a proposition: clarifying a positive and negative outcome, a set of cases which the proposition is intended to explain (a context or contrast-space) and operational definitions of the foregoing. The same basic point has been frequently articulated in statistical

---

25. 'Cases' are understood as observations that provide independent evidence of a proposition. Cases may be located across units under study or within a unit under intensive study, as is typical in case study research (Gerring, 2004).

terms. As Eckstein (1975: 113) observes, 'through any single point an infinite number of curves or lines can be drawn'. In other words, the causal relationships that one might intuit from a single case (data point) are, in principle, infinite. Large-N studies will also have an easier time identifying causal *contingency* in a research situation. To see why this is so, let us consider the N'1 research design. With a single event under our microscope, any number of things might be considered part of the normal course of events. The writer has wide latitude to construct her own definition of contingency. However, if we broaden the number of cases under investigation, we have *automatically* provided a way to distinguish between what is normal and what is abnormal. Whatever varies among these cases may be considered contingent; whatever does not may be considered part of the normal course of events. With respect to other criteria of research design, I should note briefly that plenitude usually enhances the boundedness, representativeness, and variation of a sample; it is, by the same token, often at odds with the criterion of comparability. These points are taken up later.

There is one notable exception to the dictum of plenitude (more cases are better). This concerns the falsification of necessary or sufficient causal arguments, either of which may be disproven by a single case (Dion, 1998). In this respect at least, the pluralist view of causation seems justified, for different causal arguments presume a somewhat different attitude toward research design. However, in practice, this is a relatively small exception. Few causal arguments are understood in such deterministic (invariant) terms as to be effectively falsified with a single case. More commonly, and realistically, the existence of a single counter-instance prompts the reformulation of a theory, rather than its rejection. If one were to discover a single instance of democracies fighting wars with one another, one would be more likely to revise the theory (along probabilistic lines) than to jettison it.

As a final note, it may be worth pointing out that while scarcity is intrinsically problematic (for reasons explored later), plenitude is not. The existence of multiple cases may bring other problems in its train – e.g. cases that are of questionable comparability – but it is still, ceteris paribus, a desirable quality in a research design.

2. Comparability

The problematic status of a case is that in order for it to serve its function, in order for it to be a case *of* something, cases must be similar to one another in some (though by no means all) respects. 'Comparability' refers to three elements of a research sample: (a) *descriptive comparability:* the comparability of the relevant Xs and the Y (such that 'X' and 'Y' mean roughly the same thing across cases); (b) *causal comparability:* the comparability of the $X:Y$ relationship (such that X and Y do not interact in idiosyncratic

ways in different cases); and (c) *control:* the extent to which remaining dis-
similarities (of both sorts) may be taken into account.[26]

Descriptive comparability is easiest to explain, albeit often difficult to
determine. It is essentially equivalent to concept validity across the chosen
cases. *Causal* comparability is somewhat more complex since it refers to a
hypothetical relationship among causes and effects. Evidently, there must
be some similarity between entities in order for phenomenon *A* to tell us any-
thing about phenomenon *B*. But perfect similarity (identity, uniformity) is
not necessary. Indeed, strictly speaking, it is not even possible, since every
case (every entity) is, in some minimal sense, unique. (At the very least, it
occupies a unique position in time and/or space.) In some situations, cases
need only bear a distant resemblance to one another. If we are investigating
the role of gravity, a piano is as good a case as an apple. Either will do.
Generally speaking, one can say that two cases are comparable, or 'unit
homogeneous', when they respond in similar ways to similar stimuli.
Achieving causal comparability in a research design means choosing cases
that are similar to each other in whatever ways might affect the *Y* or the
posited *X:Y* relationship, *or whose remaining differences can be taken into
account* by the analysis. All causal investigations attempt to hold certain
features of the cases 'constant', even if this extends only to one very basic
characteristic, such as having mass. Again, it is important to stress that the
particular features (*X*s) one has to worry about are contingent upon the par-
ticular outcome of interest (*Y* or *X:Y* relationship). It is neither necessary,
nor possible, nor desirable for *all* features be similar, as the criterion of
variation suggests (see later).

3. Independence

The problem of case independence is stated in general terms by Zelditch: 'If a
unit is not independent, no new information about [a variable] is obtained by
studying it twice, and no additional confirmation of [a theory] is obtained
by counting it twice'.[27] This refers to diachronic case-independence (serial

---

26. On the issue of comparability see Achen and Snidal (1989), DeFelice (1986), Dion (1998),
Eidlin (1983), Ember (1991), Frendreis (1983), Geddes (1990), Hammel (1980), Holt and Turner
(1970), Meckstroth (1975), Munck (1998: 29), Przeworski and Teune (1970), Smelser (1976),
Vallier (1971). What I have called comparability has also been referred to as 'equivalence'
(Van Deth, 1998; Van Vijver and Leung, 1997: 7) and 'unit homogeneity' (King et al., 1994).

27. Zelditch (1971: 282–3), quoted in Lijphart (1975: 171). Zelditch may be overstating the
case somewhat. One *can* obtain new information by studying a non-independent case but this
information gain will be less than if the case were wholly independent (Craig Thomas, personal
communication). See also Lijphart (1975: 171), Przeworski and Teune (1970: 52).

autocorrelation) as well as synchronic case-independence (spatial autocorrelation). Usually, problems of case-dependence create problems of comparability as well. For example, if certain countries' government expenditures are influenced by membership in the EU, then they may no longer be comparable to countries outside the EU (with respect to their public expenditures). With problems of independence, as with problems of comparability, these issues can sometimes be corrected, either statistically or through some more informal method of control. Even so, one would obviously prefer perfectly independent cases, obviating the need for controls – whose efficacy is necessarily somewhat suspect.

It should be noted that some causal questions concern the very issue of case-independence – e.g. of processes of diffusion (e.g. Collier and Messick, 1975). Even so, one is attempting to test *certain* hypotheses about diffusion (or non-independence), holding other factors constant. The latter factors constitute the sort of case-independence that one must retain if the study is to have any validity. Thus, independence is properly regarded as a desirable characteristic of all research designs.

4. Representativeness

Comparability refers to the internal properties of the sample. *Representativeness* refers to comparability between the sample and the population (sometimes referred to as 'external validity'). Specifically, are we entitled to generalize from a given case or cases under study to a larger universe of cases? Do we have reason to believe that what is true for the sample is also true for the population that we wish to describe?

The distinction between comparability and representativeness becomes clear when one considers the use of laboratory experiments on rats as evidence about human physiology or behavior. There is no problem of comparability here: rats of a given type are quite similar to each other. But there are formidable problems of representativeness. What happens to a rat, under a given set of circumstances, may or may not happen to a human. Attempts to shed light on contemporary western societies by examining 'primitive' societies in other parts of the world, small-scale human interactions (e.g. PTA meetings), or, for that matter, computer models of human societies, all meet with the same basic objection. Do the results of such studies illustrate a broader phenomenon or are they limited to the cases under study?

Unlike problems of comparability, problems of representativeness must remain at the level of assumption. Representativeness refers, by definition, to *un-studied* cases. To be sure, we may rely on other studies of these cases, we may have good theoretical reasons for supposing that the population will conform to patterns observed within our sample and we may have a

good track record with samples of this sort. Sometimes, we have the opportunity to test the assumption of representativeness. Experiments conducted on laboratory rats may be repeated on human beings. Results from a computer simulation may be repeated in the real world. Polling results are routinely put to the test of an election. In these circumstances, what is happening, methodology-wise, is that potential cases are transformed into actual cases. The point is that every time we make assumptions about a broader class of phenomena than we have directly studied we are raising questions of representativeness. No methodological procedure will overcome this basic assumption, which must be dealt with in light of what we know about particular phenomena and particular causal relationships.

It is doubtless already apparent that representativeness, like other criteria, is a matter of degrees. If we are studying the response of billiard balls when struck by other billiard balls and they are all produced by some standard set of specifications, then we might safely assume that one billiard ball is, for all intents and purposes, the same as another. Yet, small defects will appear in certain balls (rendering them unfit as cases) and even among perfect specimens there will be minute differences of mass, volume, etc. If our measurements of the outcome are sufficiently precise, these differences might turn out to be quite significant: under these circumstances, one billiard ball may not be equal to another, which is to say, the problem of representativeness becomes real. If our project is altered in a more elemental fashion – suppose we are now interested in the effects of differently colored balls on the responses of a rat – then a blue ball is no longer representative of a larger collection of (multi-colored) balls. In short, here as elsewhere, everything depends on the intent of the research. There is no such thing as representativeness *in general*.[28]

5. Variation

Empirical evidence of causal relationships is covariational in nature. In observing two billiard balls collide, we observe that *A* and *B* *covary*: where *A* hits *B*, *B* responds by moving. Prior to *A*'s arrival, *B* was stationary, and after *A*'s departure, *B* becomes stationary once again. Covariation, as the term implies, means that where *X* appears, *Y* appears as well; where *X* does not appear, *Y* is also absent; where *X* is strong so is *Y*; and so forth.

---

28. See Dewey (1938: 480). Problems of representativeness encountered in natural science are generally less acute than those usually encountered in social science. One billiard ball is apt to be fairly representative of other billiard balls, for most intents and purposes – likewise, with rats. Not so, however, for human beings and human institutions, which tend to vary enormously from one unit to another.

It might also be called correlation (common in statistical work), constant
conjunction (Hume), association (Neuman), congruity (Bennett), or con-
comitant variation (Mill).[29] All these terms convey a view of causation
expressed by Galileo: 'that and no other is to be called cause, at the presence
of which the effect always follows, and at whose removal the effect dis-
appears'.[30] Of course, the precise relationship between X and Y may not
be immediately proximate in time and space. It may be lagged; it may also
be a matter of waning and waxing, rather than appearing and disappearing
entirely. It may also be probabilistic, rather than invariant. The central
point is that, whatever the temporal and spatial relationship, it is *regular*
(or can assumed to be regular).

Granted, one is concerned not simply with the range of variation but also
the distribution of variation. Hence, the terms *variance* and *co-variance*
express more precisely what is at issue. However, since these terms are not
commonly employed in non-statistical work, I shall retain the more familiar
terms, variation and co-variation.

Four sorts of covariational patterns are important for the empirical
demonstration of causal relationships: (a) variation on the outcome of
interest (Y-variation), (b) variation on the causal variable(s) of interest
(X-variation), (c) the *avoidance* of covariation among putative causes
(collinearity), and (d) variation *within* one or more units ('within-case'
variation).

These four considerations affect all research designs. The experimental
method achieves the first three desiderata by arbitrary manipulation; it
may or may not offer evidence drawn from within-case variation (this
might require an additional experimental test). The counterfactual method
(the 'thought-experiment') deals with real and imagined variation. Variation
in a research design may be observed spatially (across units) and/or tempo-
rally (on the same units).

The general point to note is that the right sort of variation – involving only
the variables of interest – is *always useful*, and always preferable to non-

---

29. See Bennett (1999), Hume (1960: 219), Marini and Singer (1988), Neuman (1997: 50), and
Mill (1843/1949: 263). These terms are not exactly synonymous (no two terms are). Yet, for my
purposes the differences between them are so slight as to hinder, rather than enhance, clarity.

30. Quoted in Bunge (1959: 33). Bowley, an early pioneer of statistical modelling, put it this
way: 'It is never easy to establish the existence of a causal connection between two phenomena or
series of phenomena; but a great deal of light can often be thrown by the application of algebraic
probability . . . When two quantities are so related that . . . an increase or decrease of one is found
in connection with an increase or decrease (or inversely) of the other, and the greater the mag-
nitude of the changes in the one, the greater the magnitude of the changes in the other, the quan-
tities are said to be *correlated. Correlation* is a quantity which can be measured numerically'
(quoted in Morgan 1997: 62). See also Frendreis (1983).

variation research designs. The only exception is the research design whose purpose is to *dis*prove an invariant ('deterministic') causal argument; in this rather unusual instance, a no-variation research design is acceptable (Dion, 1998).

6. Transparency

A research design is also to be prized if it offers evidence about causal mechanisms – e.g. a series of dominoes colliding into one another. Naturally, evidence of this sort is also covariational in nature (one observes the covariation of a succession of intermediary variables, in this case the individual dominoes that lie between the beginning and the end of this causal chain). But it is worth distinguishing the intermediary processes by which a causal relationship takes place from the inputs and outputs. Thus, I refer to a research design that offers process-tracing evidence of this sort as *transparent* – relative that is to another research design which may be more opaque (a 'black-box'). Process-tracing has many near-synonyms, including 'discerning', 'process analysis', 'pattern-matching', 'microfoundations', 'causal narrative', 'congruence', 'colligation', 'contiguity', and 'intermediate processes'.[31] All concern the ways in which one might elucidate a clear causal path between X and Y.

7. Replicability

A good research design is one that produces *reliable* results (the results do not vary greatly from iteration to iteration) and is *replicable* by future researchers (King, 1995; King et al., 1994: 23, 26, 51). Reliability and replicability may be viewed as two aspects of the same general goal – one applying to different results from a given study and the other applying to results obtained by different studies.

---

31. Terms and associated works are as follows: 'process-tracing' (George and McKeown, 1985: 34ff), 'discerning' (Komarovsky, 1940: 135–46), 'process analysis' (Barton and Lazarsfeld, 1969), 'pattern-matching' (Campbell, 1975), 'microfoundations' (Little, 1998), 'causal narrative' (Abbott, 1990, 1992; Abrams, 1982; Aminzade, 1992; Bates et al., 1998; Griffin, 1992, 1993; Katznelson, 1997; Kiser, 1996; Mink, 1987; Quadagno and Knapp, 1992; Roth, 1994; Rueschemeyer and Stephens, 1997; Sewell, 1992, 1996; Somers, 1992; Stone, 1979; Stryker, 1996; Watkins, 1994), 'congruence' (George and Bennett, 2004), 'colligation' (Roberts, 1996), 'intermediate processes' (Mill, 1843/1872). For general discussion, see Bennett (1999), Brown (1984: 228), Collier and Mahoney (1996: 70), Goldstone (1997). For philosophical discussion in the 'realist' tradition, see Bhaskar (1975/1978), Harre (1972), McMullin (1984), and Salmon (1984).

The key point is that if a finding is obtained under circumstances that are essentially un-repeatable – e.g. a 'natural' experiment that comes around only once – then we rightfully entertain doubts about its veracity. We are cognizant that any number of factors might have interfered with the validity of the original study, including (among other things) measurement error and the willful mis-reporting of data. Verification involves repetition; claims to truth, therefore, involve assurances of replicability. Replicability is simply the formal acknowledgement of the ongoing collaborative project that is science.

Conclusions

I began this essay by noting that there are two traditional approaches to causation in the social sciences, unitarism and pluralism. I argued that existing unitarian views are too constrained while pluralist views are either unconvincing or, to the extent that they are true, unfortunate. We need a single framework within which to understand causal relationships in the social sciences. At the same time, this framework must be sufficiently broad to encompass all (or most) of the arguments we typically describe as causal and it must do so in a coherent fashion. The desiderata must be logically derived from some central postulates about what causation is, or should be, within the realm of social science, not a jerrybuilt (ad hoc) assortment of dos and don'ts.

It seems appropriate at this juncture to say a few words about how the parts of this puzzle fit together. As a point of departure, I proposed a minimal definition of causation. That which raises the probability of an event occurring may be considered a 'cause'. This definition is substitutable across all usages and contexts within the social sciences, providing a reasonably bounded concept (one that will not be confused with other key terms in the lexicon). Subsequent sections laid out a criterial approach to the intertwined problems of (a) forming causal propositions (second section) and (b) testing those propositions (third section). All causal arguments, I argued, are liable to 16 formal criteria (see Table 1) and seven criteria of research design (see Table 2).

To be sure, writers may choose to privilege some criteria over others. Forming a proposition and testing that proposition involves choosing among competing goods. But such choices usually involve costs and it is in this sense that we may describe the framework as intrinsic to the enterprise of causal analysis within the social sciences. A writer striving for breadth, who thereby compromises precision, is still liable to accusations of imprecision. The terms of the negotiated trade-off are persistent and irreducible, and traverse the territories often ascribed to different causal

traditions (according to the pluralist view). They are *not* conditional upon specific models of causal inference.[32]

Insofar as one values cumulation in the social sciences – which is to say, the ability to adjudicate between rival arguments – one must strive for a unified conception of causation. It is the nature of the proposition at issue, and the nature of the evidence available for and against that proposition, rather than one's theory of causation, that ought to drive the research agenda. Thus, for practical as well as for epistemological reasons, we are well advised to foreground the commonalities in causal inference within the social sciences rather than the differences.

Let me be slightly more specific and considerably more programmatic. In recent years, it has become quite common to distinguish among causal arguments that are 'mechanismic' and those that are 'correlational', as discussed at the outset of this essay. I have argued that *all* causal arguments strive for evidence of covariational (correlational) relationships between the putative X and Y *as well as* evidence of causal pathways between X and Y. The latter, indeed, are also covariational in nature insofar as they seek to establish empirical links between the structural cause (X), one or more intervening causes, and the outcome (Y). Thus, to label causal arguments as either mechanismic or correlational does violence to the work of good social science. To be sure, it may be an accurate description of bad social science – work that slights one or the other – and perhaps this pejorative sense of these terms is justified. However, the implicit claim of most authors who employ this crude typology is that causes are justifiably mechanismic or correlational; that these are distinct ways of making a causal argument and that, perforce, we must choose among them. They are, in this usage, incommensurable. This does not seem justified, in light of the great majority of arguments that comfortably integrate both these elements of causal argument. Of course, this is only one of many examples of causal-pluralist arguments, which come in many shapes and sizes and are discussed in some detail at the outset of this essay. But it will do nicely for purposes of illustration since it is probably more familiar to readers than some of the more differentiated typologies.

In sum, this paper comes down with the positivists on the question of causal unity. We can, and ought, to embrace a uniform framework of causation in the social sciences – what causation means, the formal criteria that apply to causal arguments, and the general criteria that apply to their empirical testing. However, this paper has defended a much more pluralistic

---

32. The importance of *trade-offs* in the work of social science is emphasized in the work of David Collier, Larry Laudan, Adam Przeworski, Giovanni Sartori, Rudra Sil, and Henry Teune. For further discussion, see Gerring (2001: Ch. 2).

view of what those criteria consist of than most so-called positivists would presumably agree with. Indeed, I have argued that the formal and empirical criteria are a good deal more wide-ranging than even the multiple typologies of the pluralists might suggest. There is plenty of room for different styles of causal argument, emphasizing different criteria along various dimensions (pertaining to the formal elements of causal argument and/or research design). The point is that we should be able to evaluate them within the same general rubric of causation. Arguments, so framed, can meet. Cumulation is possible.

  More generally, this paper argues against the common tendency to divide up work in the social sciences into distinct schools, each with its own ontology, epistemology, and toolkit of approved methods. The 'separate tables' (Almond, 1990) vision of political science, while lamentably true, must be resisted. Reaching across these tables requires that we be able to conceptualize important research tasks like causal analysis in a unitary, rather than fragmentary, manner – *without* sacrificing the diversity of causal arguments and evidence that now enriches the social sciences. The criterial framework, which I have explored in greater detail elsewhere (Gerring, 2001), may provide the conceptual scaffolding for such a re-thinking.

REFERENCES

Abbott, Andrew (1990) 'Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approaches', *Historical Methods* 23(4): 140–50.

Abbott, Andrew (1992) 'From Causes to Events: Notes on Narrative Positivism', *Sociological Methods and Research* 20(4): 428–55.

Abrams, Philip (1982) *Historical Sociology*. Ithaca, NY: Cornell University Press.

Achen, Christopher H. and Duncan Snidal (1989) 'Rational Deterrence Theory and Comparative Case Studies', *World Politics* 41(Jan.): 143–69.

Adcock, Robert (2002) 'Probability, Contingency, and Determinism in Comparative Historical Analysis', paper presented at the annual meetings of the American Political Science Association, Boston MA, August–September.

Almond, Gabriel A. (1990) *A Discipline Divided: Schools and Sects in Political Science*. Newbury Park, CA: Sage.

Almond, Gabriel A. and Stephen J. Genco (1977/1990) 'Clouds, Clocks, and the Study of Politics', in Gabriel A. Almond (ed.) *A Discipline Divided: Schools and Sects in Political Science*. Newbury Park, CA: Sage.

Aminzade, Ronald (1992) 'Historical Sociology and Time', *Sociological Methods and Research* 20: 456–80.

Barton, Alan H. and Paul F. Lazarsfeld (1969) 'Some Functions of Qualitative Analysis in Social Research', in George J. McCall and J.L. Simmons (eds) *Issues in Participant Observation*. Reading, MA: Addison-Wesley.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal and Barry Weingast (1998) *Analytic Narratives*. Princeton, NJ: Princeton University Press.

Bennett, Andrew (1999) 'Causal Inference in Case Studies: From Mill's Methods to Causal Mechanisms', paper presented at the annual meetings of the American Political Science Association, Atlanta, GA, September.

Bhaskar, Roy (1975/1978) *A Realist Theory of Science*. Brighton, Sussex: Harvester Press.

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

Brady, Henry E. (2002) 'Models of Causal Inference: Going Beyond the Neyman–Rubin–Holland Theory', paper prepared for the annual meeting of the Midwest Political Science Association, Chicago, IL.

Brown, Robert (1984) *The Nature of Social Laws: Machiavelli to Mill*. Cambridge: Cambridge University Press.

Bunge, Mario (1959) *Causality*. Cambridge, MA: Harvard University Press.

Bunge, Mario (1997) 'Mechanism and Explanation', *Philosophy of the Social Sciences* 27(Dec.): 410–65.

Campbell, Donald T. (1975) ' ''Degrees of Freedom'' and the Case Study', *Comparative Political Studies* 8(Jul.): 178–93.

Campbell, Donald T. (1988) *Methodology and Epistemology for Social Science*, ed. E. Samuel Overman. Chicago: University of Chicago Press.

Campbell, Donald T. and Julian Stanley (1963) *Experimental and Quasi-Experimental Designs for Research*. Boston, MA: Houghton Mifflin.

Cavell, Stanley (1979) *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. Oxford: Oxford University Press.

Collier, David (1993) 'The Comparative Method', in Ada W. Finifter (ed.) *Political Science: The State of the Discipline*, Vol. II. Washington DC: American Political Science Association.

Collier, David and James Mahoney (1996) 'Insights and Pitfalls: Selection Bias in Qualitative Research', *World Politics* 49 (Oct.): 56–91.

Collier, David and Richard E. Messick (1975) 'Prerequisites Versus Diffusion: Testing Alternative Explanations of Social Security Adoption', *American Political Science Review* 69(4): 1299–1315.

Comte, Auguste (1975) *Auguste Comte and Positivism: The Essential Writings*, ed. Gertrud Lenzer. Chicago: University of Chicago Press.

Cook, Thomas and Donald Campbell (1979) *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin.

DeFelice, E. Gene (1986) 'Causal Inference and Comparative Methods', *Comparative Political Studies* 19(3): 415–37.

Dessler, David (1991) 'Beyond Correlations: Toward a Causal Theory of War', *International Studies Quarterly* 35: 337–55.

Dewey, John (1938) *Logic: The Theory of Inquiry*. New York: Henry Holt.

Dion, Douglas (1998) 'Evidence and Inference in the Comparative Case Study', *Comparative Politics* 30(2): 127–45.

Durkheim, Emile (1895/1964) *The Rules of Sociological Method*. New York: The Free Press.

Eckstein, Harry (1975) 'Case Studies and Theory in Political Science', in Fred I. Greenstein and Nelson W. Polsby (eds) *Handbook of Political Science, Vol. 7. Political Science: Scope and Theory*. Reading, MA: Addison-Wesley.

Eidlin, Fred (1983) 'Area Studies and/or Social Science: Contextually-Limited Generalizations versus General Laws', in Fred Eidlin (ed.) *Constitutional Democracy: Essays in Comparative Politics*. Boulder, CO: Westview.

Einhorn, Hillel J. and Robin M. Hogarth (1986) 'Judging Probable Cause', *Psychological Bulletin* 99(3): 3–19.

Einstein, Albert (1940/1953) 'The Fundaments of Theoretical Physics', in Herbert Feigl and May Brodbeck (eds) *Readings in the Philosophy of Science*, pp. 253–62. New York: Appleton-Century-Crofts.

Elster, Jon (1989) *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Ember, Melvin (1991) 'The Logic of Comparative Research', *Behavior Science Research* 25: 143–54.

Feagin, Joe R., Anthony M. Orum and Gideon Sjoberg (1991) *A Case for the Case Study*. Chapel Hill, NC: University of North Carolina Press.

Frendreis, John P. (1983) 'Explanation of Variation and Detection of Covariation: The Purpose and Logic of Comparative Analysis', *Comparative Political Studies* 16(2): 255–72.

Friedman, Kenneth S. (1972) 'Empirical Simplicity as Testability', *British Journal for the Philosophy of Science* 23: 25–33.

Friedman, Michael (1974) 'Explanation and Scientific Understanding', *Journal of Philosophy* LXXI: 5–19.

Gardiner, Patrick (1952/1961) *The Nature of Historical Explanation*. Oxford: Oxford University Press.

Garfinkel, Alan (1981) *Forms of Explanation: Rethinking the Questions of Social Theory*. New Haven, CT: Yale University Press.

Gasking, D. (1955) 'Causation and Recipes', *Mind* 64: 479–87.

Geddes, Barbara (1990) 'How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics', in James A. Stimson (ed.) *Political Analysis*, Vol. 2. Ann Arbor: University of Michigan Press.

Geertz, Clifford (1973) *The Interpretation of Cultures*. New York: Basic Books.

George, Alexander L. and Andrew Bennett (2005) *Case Studies and Theory Development*. Cambridge, MA: MIT Press.

George, Alexander L. and Timothy J. McKeown (1985) 'Case Studies and Theories of Organizational Decision Making', *Advances in Information Processing in Organizations*, Vol. 2. Santa Barbara: JAI Press.

Gerring, John (2001) *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.

Gerring, John (2004) 'What is a Case Study and What is it Good For?', *American Political Science Review* 98(2): 341–54.

Gerring, John (forthcoming) 'A Normative Turn in Political Science?', in Richard Ned Lebow (ed.) *Theory and Evidence*.

Gerring, John and Paul A. Barresi (2003) 'Putting Ordinary Language to Work: A Min–Max Strategy of Concept Formation in the Social Sciences', *Journal of Theoretical Politics* 15(2): 201–32.

Glennan, Stuart S. (1992) 'Mechanisms and the Nature of Causation', *Erkenntnis* 44: 49–71.

Glennan, Stuart S. (2000) 'Rethinking Mechanistic Explanation', unpublished manuscript, Department of Philosophy and Religious Studies, Butler University, 1–31 December.

Glymour, Clark. 1980. *Theory and Evidence*. Princeton: Princeton University Press.

Goertz, Gary and Harvey Starr, eds (2002) *Necessary Conditions: Theory, Methodology and Applications*. New York: Rowman and Littlefield.

Goldstone, Jack A. (1997) 'Methodological Issues in Comparative Macrosociology', *Comparative Social Research* 16: 121–32.

Goodman, Nelson (1965) *Fact, Fiction, Forecast*, 2nd edn. Indianapolis: Bobbs-Merrill.

Green, Donald P. and Ian Shapiro (1994) *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.

Griffin, Larry J. (1992) 'Temporality, Events, and Explanation in Historical Sociology: An Introduction.' *Sociological Methods and Research* 20(4): 403–27.

Griffin, Larry J. (1993) 'Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology', *American Journal of Sociology* 98: 1094–133.

Hall, Peter A. (2003) 'Aligning Ontology and Methodology in Comparative Politics', in James Mahoney and Dietrich Rueschemeyer (eds) *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Hall, Ned (2004) 'Two Concepts of Causation', in John Collins, Ned Hall and L. A. Paul (eds) *Causation and Counterfactuals*, pp. 225–76. Cambridge, MA: MIT Press.

Hammel, E.A. (1980) 'The Comparative Method in Anthropological Perspective', *Comparative Studies in Society and History* 22(2): 145–55.

Harman, Gilbert (1965) 'The Inference to the Best Explanation', *Philosophical Review* lxxiv: 88–95.

Harre, Rom (1970) *The Principles of Scientific Thinking*. Chicago: University of Chicago Press.

Harre, Rom (1972) *The Philosophies of Science*. London: Oxford University Press.

Harre, Rom and E.H. Madden (1975) *Causal Powers: A Theory of Natural Necessity*. Oxford: Blackwell.

Hart, H.L.A. and A.M. Honore (1959) *Causality in the Law*. Oxford: Oxford University Press.

Hart, H.L.A. and A.M. Honore (1966) 'Causal Judgment in History and in the Law', in William H. Dray (ed.) *Philosophical Analysis and History*. New York: Harper & Row.

Hawthorn, Geoffrey (1991) *Plausible Worlds: Possibility and Understanding in History and the Human Sciences*. Cambridge: Cambridge University Press.

Heckscher, Gunnar (1957) *The Study of Comparative Government and Politics*. London: George Allen and Unwin.

Hedstrom, Peter and Richard Swedberg, eds (1998) *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Hempel, Carl G. (1991) 'Empiricist Criteria of Cognitive Significance: Problems and Changes', in Richard Boyd, Philip Gasper and J.D. Trout (eds) *The Philosophy of Science*. Cambridge, MA: MIT Press.

Hempel, Carl G. and Paul Oppenheim (1948) 'Studies in the Logic of Explanation', *Philosophy of Science* 15(2): 135–75.

Hesse, Mary (1974) *The Structure of Scientific Inference*. London: Macmillan.

Hirschman, Albert O. (1970) 'The Search for Paradigms as a Hindrance to Understanding', *World Politics* 22(3): 329–43.

Holland, Paul W. (1986) 'Statistics and Causal Inference', *Journal of the American Statistical Association* 81: 945–60.

Holt, R.T. and J.E. Turner (1970) 'The Methodology of Comparative Research', in R.T. Holt and J.E. Turner (eds) *The Methodology of Comparative Research*. New York: Free Press.

Homans, George C. (1967) *The Nature of Social Science*. New York: Harcourt Brace Jovanovich.

Hume, David (1960) 'The Idea of Necessary Connexion' [from *An Enquiry Concerning Human Understanding*, Sect. 7], in Edward H. Madden (ed.) *The Structure of Scientific Thought: An Introduction to Philosophy of Science*. London: Routledge & Kegan Paul.

Hume, David (1985) *Essays: Moral, Political, and Literary*, ed. Eugene F. Miller. Indianopolis: LibertyClassics.

Johnson, James (2003) 'Conceptual Problems as Obstacles to Theoretical Progress in Political Science: Political Culture among the "Sociologists" ', *Journal of Theoretical Politics* 15(1): 87–115.

Katznelson, Ira (1997) 'Structure and Configuration in Comparative Politics', in Mark Irving Lichbach and Alan S. Zuckerman (eds) *Comparative Politics: Rationality, Culture, and Structure*. New York: Cambridge University Press.

King, Gary (1995) 'Replication, Replication', *PS: Political Science and Politics* 28(3): 443–99.

King, Gary, Robert O. Keohane and Sidney Verba (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Kirkham, Richard L. (1992) *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press.

Kiser, Edgar (1996) 'The Revival of Narrative in Historical Sociology: What Rational Choice Can Contribute', *Politics and Society* 24: 249–71.

Kitcher, Phillip (1989) 'Explanatory Unification and the Causal Structure of the World', in Phillip Kitcher and Wesley Salmon (eds) *Scientific Explanation: Minnesota Studies in the Philosophy of Science*, Vol. XIII. Minneapolis: University of Minnesota Press.

Komarovsky, Mirra (1940) *The Unemployed Man and His Family: The Effect of Unemployment upon the Status of the Man in Fifty-nine Families*. New York: Dryden Press.

Kuhn, Thomas S. (1977) *The Essential Tension*. Chicago: University of Chicago Press.

Lakatos, Imre (1978) *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press.

Laudan, Larry (1977) *Progress and its Problems: Toward a Theory of Scientific Growth*. Berkeley: University of California Press.

Laudan, Larry (1996) *Beyond Positivism and Relativism: Theory, Method, and Evidence*. Boulder, CO: Westview Press.

Levey, Geoffrey Brahm (1996) 'Theory Choice and the Comparison of Rival Theoretical Perspectives in Political Sociology', *Philosophy of the Social Sciences* 26(1): 26–60.

Lewis, David K. (1973) *Counterfactuals*. Oxford: Blackwell.

Lieberson, Stanley (1985) *Making It Count: The Improvement of Social Research and Theory*. Berkeley, CA: University of California Press.

Lijphart, Arend (1975) 'The Comparable Cases Strategy in Comparative Research', *Comparative Political Studies* 8: 158–77.

Little, Daniel (1998) *Microfoundations, Method, and Causation*. New Brunswick: Transaction.

Mach, Ernest (1902/1953) 'The Economy of Science', in Philip P. Wiener (ed.) *Readings in Philosophy of Science*, pp. 446–52. New York: Charles Scribner's.

Mackie, John L. (1965/1993) 'Causes and Conditions', in Ernest Sosa and Michael Tooley (eds) *Causation*. Oxford: Oxford University Press.

Mackie, John L. (1974) *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.

McMullin, Ernan (1984) 'Two Ideals of Explanation in Natural Science', in Peter A. French, Theodore E. Uehling, Jr and Howard K. Wettstein (eds) *Midwest Studies in Philosophy Volume IX: Causation and Causal Theories*. Minneapolis: University of Minnesota Press.

Mahoney, James (2001) 'Beyond Correlational Analysis: Recent Innovations in Theory and Method', *Sociological Forum* 16(3): 575–93.

Marini, Margaret and Burton Singer (1988) 'Causality in the Social Sciences', in Clifford Clogg (ed.) *Sociological Methodology* 18: 347–409.

Meckstroth, Theodore (1975) ' "Most Different Systems" and "Most Similar Systems": A Study in the Logic of Comparative Inquiry', *Comparative Political Studies* 8(2): 133–77.

Mill, John Stuart (1843/1872) *System of Logic*, 8th edn. London: Longmans, Green.

Miller, Richard W. (1987) *Fact and Method: Explanation, Confirmation and Reality in the Natural and the Social Sciences*. Princeton, NJ: Princeton University Press.

Mink, Louis (1987) 'History and Fiction as Modes of Comprehension', in Brian Fay, Eugene Golob and Richard Van (eds) *Historical Understanding*. Ithaca, NY: Cornell University Press.

Morgan, Mary S. (1997) 'Searching for Causal Relations in Economic Statistics', in Vaughn R. McKim and Stephen P. Turner (eds) *Causality in Crisis?: Statistical Methods and the Search for Causal Knowledge in the Social Science*. Notre Dame, IN: Notre Dame Press.

Munck, Gerardo L. (1998) 'Canons of Research Design in Qualitative Analysis', *Studies in Comparative International Development* 33(3): 18–45.

Myrdal, Gunnar (1970) *The Challenge of World Poverty: A World Anti-Poverty Program in Outline*. New York: Pantheon.

Neuman, W. Lawrence (1997) *Social Research Methods: Qualitative and Quantitative Approaches*, 2nd edn. Boston, MA: Allyn & Bacon.

Popper, Karl (1934/1968) *The Logic of Scientific Discovery*. New York: Harper & Row.

Popper, Karl (1969) *Conjectures and Refutations*. London: Routledge and Kegan Paul.

Przeworski, Adam and Henry Teune (1970) *The Logic of Comparative Social Inquiry*. New York: John Wiley.

Putnam, Hilary (1978) *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.

Quadagno, Jill and Stan J. Knapp (1992) 'Have Historical Sociologists Forsaken Theory:? Thoughts on the History/Theory Relationship.' *Sociological Methods and Research* 20: 481–507.

Quine, Willard van Orman (1966) 'Simple Theories of a Complex World', in *the Ways of Paradox and Other Essays*. New York: Random House.

Ragin, Charles C. (1987) *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California.

Ragin, Charles C. (2000) *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. and Howard S. Becker (eds) (1992) *What Is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press.

Roberts, Clayton (1996) *The Logic of Historical Explanation*. University Park: Pennsylvania State University Press.

Roth, Paul A. (1994) 'Narrative Explanations: The Case of History', in Michael Martin and Lee C. McIntyre (eds) *Readings in the Philosophy of Social Science*. Cambridge, MA: MIT Press.

Rueschemeyer, Dietrich and John D. Stephens (1997) 'Comparing Historical Sequences – A Powerful Tool for Causal Analysis', *Comparative Social Research* 16: 55–72.

Rule, James B. (1997) *Theory and Progress in Social Science*. Cambridge: Cambridge University Press.

Salmon, Wesley C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Salmon, Wesley C. (1989) 'Four Decades of Scientific Explanation', in Phillip Kitcher and Wesley Salmon (eds) *Scientific Explanation: Minnesota Studies in the Philosophy of Science*, Vol. XIII. Minneapolis: University of Minnesota Press.

Sewell, William H., Jr (1992) 'Introduction: Narratives and Social Identities', *Social Science History* 16: 479–88.

Skocpol, Theda and Margaret Somers (1980) 'The Uses of Comparative History in Macrosocial Inquiry', *Comparative Studies in Society and History* 22(2): 147–97.

Simowitz, Roslyn and Barry L. Price (1990) 'The Expected Utility Theory of Conflict: Measuring Theoretical Progress', *American Political Science Review* 84(2): 439–60.

Smelser, Neil J. (1976) *Comparative Methods in the Social Sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Sober, Elliott (1975) *Simplicity*. Oxford: Oxford University Press.

Sober, Elliott (1988) *Reconstructing the Past: Parsimony, Evolution and Inference*. Cambridge: MIT Press.

Somers, Margaret R. (1992) 'Narrativity, Narrative Identity, and Social Action: Rethinking English Working-Class Formation', *Social Science History* 16: 591–630.

Sosa, Ernest and Michael Tooley (eds) (1993) *Causation*. Oxford: Oxford University Press.

Spohn, Wolfgang (1983) 'Deterministic and Probabilistic Reasons and Causes', *Erkenntnis* 19: 371–96.

Stinchcombe, Arthur L. (1968) *Constructing Social Theories*. New York: Harcourt, Brace.

Stone, Lawrence (1979) 'The Revival of Narrative: Reflections on a New Old History', *Past and Present* 85 (Nov.): 3–24.

Stryker, Robin (1996) 'Beyond History versus Theory: Strategic Narrative and Sociological Explanation', *Sociological Methods and Research* 24(3): 304–52.

Suppes, Patrick C. (1970) *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.

Taylor, Charles (1970) 'The Explanation of Purposive Behavior', in Robert Borger and Frank Cioffi (eds) *Explanation in the Behavioral Sciences*. Cambridge: Cambridge University Press.

Tilly, Charles (2001) 'Mechanisms in Political Processes', *Annual Review of Political Science* 4: 21–41.

Vallier, I., ed. (1971) *Comparative Methods in Sociology: Essays on Trends and Applications*. Berkeley, CA: University of California Press.

Van Deth, Jan W., ed. (1998) *Comparative Politics: The Problem of Equivalence*. New York: Routledge.

Van Evera, Stephen (1997) *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Van Fraassen, Bas C. (1980) *The Scientific Image*. Oxford: Clarendon Press.

Van Vijver, Fons and Kwok Leung (1997) *Methods and Data Analysis for Cross-Cultural Research*. Thousand Oaks, CA: Sage.

Von Wright, Georg Henrik (1971) *Explanation and Understanding*. Ithaca, NY: Cornell University Press.

Waldner, David (2002) 'Anti Anti-Determinism: Or What Happens When Schrodinger's Cat and Lorenz's Butterfly Meet Laplace's Demon in the Study of Political and Economic Development', paper presented to the annual meetings of the American Political Science Association, Boston MA, August–September.

Watkins, J.W.N. (1994) 'Historical Explanation in the Social Sciences', in Michael Martin and Lee C. McIntyre (eds) *Readings in the Philosophy of Social Science*. Cambridge, MA: MIT Press.

Whitbeck, C. (1977) 'Causation in Medicine: The Disease Entity Model', *Philosophy of Science* 44: 619–37.

Wilson, Edward O. (1998) *Consilience: The Unity of Knowledge*. New York: Alfred A. Knopf.

Zelditch, M., Jr (1971) 'Intelligible Comparisons', in I. Vallier (ed.) *Comparative Methods in Sociology: Essays on Trends and Applications*, pp. 267–307. Berkeley, CA: University of California Press.

JOHN GERRING is associate professor of political science at Boston University. His work has appeared in *American Political Science Review*, *British Journal of Political Science*, *International Organization*, *Journal of Policy History*, *Journal of Theoretical Politics*, *Party Politics*, *Political Research Quarterly*, *Polity*, *Social Science History*, and *Studies in American Political Development*. His books include *Party Ideologies in America, 1828–1996* (1998), *Social Science Methodology: A Criterial Framework* (2001), *Case Study Research: Principles and Practices* (forthcoming, 2006), *Global Justice: A Pragmatic Approach* (forthcoming), and *Centripetalism: A Theory of Democratic Governance* (forthcoming; with Strom Thacker). ADDRESS: Boston University, Department of Political Science, 232 Bay State Road Boston MA 02215, 617–353–2756, USA [email: jgerring@bu.edu].