Evidence for Use: The Role of Case Studies in Political Science Research

Abstract

In its most recent form, the debate about the relationship between quantitative and qualitative methodology in political science has been shaped by the publication of *Designing Social Inquiry: Scientific Inference in Qualitative Research* by Gary King, Robert O. Keohane, and Sidney Verba in 1994 (hereafter *DSI*). The focus of this debate has been case study research. *DSI* advocates that qualitative research, particularly case study research, be modeled on the template of quantitative research. The authors claim that all research has the same logic of inquiry and that this is most clearly exemplified in quantitative work.   I argue that the underlying philosophy of science of *DSI* is monistic and positivistic in ways not productive for understanding various different purposes that political science knowledge may have. Different methodologies have different strengths and so are suited to different ends. I examine this in relation to Julian Reiss's discussion of different concepts of causality and argue that case study research is suited to understanding causal mechanisms in ways that make such research better suited to inform policy decisions. I finish with an example using David Fearon's 2006 Congressional Testimony on Iraq.

0. Introduction.

My use of the phrase "evidence for use" originates in Nancy Cartwright's 2006

paper, "Well-Ordered Science: Evidence for Use." There she states that her goal

> is to urge us [philosophers of science] to direct our efforts away from the more abstract questions that usually entertain us – from highly general questions of warrant …to much more specific questions about particular methods and their problems of implementation, their range of validity, their strengths and weaknesses, and their costs and benefits" (2006, 982).

She goes on to suggest that one way to do this is to pay more attention to how

claims are used instead of focusing only on how they are tested.

This idea might be useful in looking at debates in political science about the

relative merits of quantitative and qualitative methods, which center in part on the

fact that different methodologies might be appropriate depending on the goals of

the science.   One of the sources of these debate may have to do with the different roles that quantitative and qualitative research play in knowledge production and use. Results that are achieved in quantitative research seem useful if one wants to make broad generalizations about correlations, and perhaps even causal connections between variables. But if one is seeking understanding of how such causes functioned "on the ground" in particular circumstances this may be the sort of specific information that only a case study could provide.

The quantitative/qualitative debate is more nuanced than this. Here is a brief sketch of some of those nuances. 1) There is a political element within the profession. What I mean by this is that "power" and prestige belong with the quantitative camp. This means that there is an asymmetry in the arguments with those who use case study research, a primary mode of qualitative method, on the defensive. 2) While case study researchers may be defensive we need to be clear about what they are defending.   Many do not work solely within the qualitative tradition but avail themselves of a variety of methods, both formal (deductive and mathematical), quantitative (essentially econometric), and qualitative and so what they are defensive about is not that case studies should play a role in political science research but what sort of role they should play. They disagree among themselves about the details. Multi-method or mixed method research is really what most practicing political science researchers find themselves engaged in and so to paint the debate as dichotomous – case study research or quantitative – is simply inaccurate. The following quote captures these points:

> Thus, a paradox: Although much of what we know about the empirical world is drawn from case studies and case studies continue to constitute a large

proportion of the work generated by the discipline, the case study *method is held in low regard or is simply ignored*. Even among its defenders there is confusion over the virtues and vices of this ambiguous research design. (Gerring 2004, 341)

The second issue is that something needs to be said about the underlying philosophy science assumed by those political scientists engaged in the debate. It is this second issue that will be the focus of the rest of the paper though the exploration of this question ultimately brings us back to the original topic of the paper.

1. Framing the Debate: Designing Social Inquiry

In its most recent form, the debate about the relationship between quantitative and qualitative methodology in political science has been shaped by the publication of *Designing Social Inquiry: Scientific Inference in Qualitative Research* by Gary King, Robert O. Keohane, and Sidney Verba (*DSI* hereafter) in 1994. The book is part polemical treatise, part handbook for the researcher, and a more general guide to making qualitative research more rigorous, which it turns out means "more like quantitative research".

The authors state their motivation for writing the book in the introduction:

Our main goal is to connect the traditions of what are conventionally denoted 'quantitative' and 'qualitative' research by applying a unified logic of inference to both. The traditions appear quite different; indeed they sometimes seem to be at war. Our view is that these differences are mainly ones of style and specific technique. The same underlying logic provides the framework for each research approach. (1994, 3)

Given that there is one logic underlying all social science research, they propose to reveal the key elements of that logic and provide a way to improve its use in

qualitative research. They do this primarily through formalization of qualitative

models that allow inference in qualitative research to mirror inference in

quantitative models. *DSI* appears to be committed to the view that the logic of

inference is exemplified in statistical/quantitative research[1] and if qualitative

methods are to be employed they should be designed using quantitative methods as

a template (Brady and Collier 2004).[2] "We argue that nonstatistical research will

produce more reliable results if researchers pay attention to the rules of scientific

inference – rules that are sometimes more clearly stated in the style of quantitative

research." (1994, 6) *DSI* goes on to specify what is intended by "rules of scientific

inference" and offers four characteristics of scientific research: (1) The goal is

inference; (2) The procedures are public; (3) The conclusions are uncertain; (4) The

content is the method. (1994, 8-9)

The first three are fairly self-explanatory and uncontentious, although if the

authors of *DSI* mean by (1) that the *only* goal of science is inference there may be

some debate and I will return to this later in the paper. (2) the procedures are

public, is really saying no more than that the methods should be replicable and not

subjective and the claim that the conclusions are uncertain is the claim that

scientific knowledge is always revisable in the light of new evidence. It is (4), that

the content is the method, that I will give the closest scrutiny.

> The content of "science" is primarily the methods and rules, not the subject
> matter, since we use these methods to study virtually anything. This point
> was recognized over a century ago when Karl Pearson (1892:16) explained
> that 'the field of science is unlimited; its material is endless; every group of
> natural phenomena, every phase of social life, every stage of past or present
> development is material for science. The unity of all science consists alone in
> its method, not its material." (1994, 9)

So *DSI* presents a version of strong unity of science thesis that identifies science with a particular method. Given that in (1) the authors identify the *one* goal of science as inference, this is probably not terribly surprising.

Though there are many other details to the account that *DSI* gives about how to properly use case study research in political science, I am going to stop here and focus on the three features the philosophy of science that *DSI* uses to frame the discussion. It includes these three key elements: (1) there is one logic of inference; (2) there is one goal of science; and (3) science is unified by method.

2. Critics and philosophy of science.

While *DSI* has framed the recent debate, the challenge to qualitative method and case study research in particular is one that many political scientists that use multi-method research have taken up. Two books from 2004 and 2005 give a pretty good overview of the range of responses to *DSI*. Henry E. Brady and David Collier's edited volume offers articles covering the spectrum and includes a response from King, Keohane, and Verba. *Case Studies and Theory Development in the Social Sciences* (2005) by Alexander L. George and Andrew Bennett also offers an alternative approach to qualitative methodology.

It is worth noting that many of these responses to *DSI* from within political science do not directly challenge the claim that there is one logic of inference. George and Bennett note an ambiguity in the claim and suggest that they might agree with this core thesis of *DSI* depending on the interpretation.

> If this logic of inference refers in a broad sense to the epistemological logic of deriving testable implications from alternative theories, testing these implications against quantitative or case study data, and modifying theories or our confidence in them in accordance with the results, then perhaps on a very general level there is one logic that is the modern successor of the still-evolving positivist tradition… .
>
> If, however, the logic of inference refers to specific methodological injunctions on such issues as the value of single-case studies, the procedures for choosing which cases to study, the role of process-tracing, and the relative importance of causal effects (the expected change in the dependent variable given a unit of change in an independent variable) and causal mechanisms as the bases for inference and explanation, …then we disagree with the overall argument… ." (1994, 10-11)

In either case, the authors of *DSI* are committed to the non-pluralistic view of science that I have already noted and it would seem that George and Bennett are as well, at least on the weaker interpretation.   As we will see in a moment, they reject *DSI'*s account of causal inference and do so on philosophical grounds, but they may nonetheless embrace the weaker claim.

The claims that I am grouping together as a view about the methodological unity of science reveal a lurking philosophy of science and so it is time to unmask it. Not surprisingly it is that old villain logical positivism, or some version of logical positivism and Popperian philosophy of science.   The key features are the unity of science thesis, the hints of some sort of demarcation thesis, and account of the use of case studies to test hypotheses. There may also be a hint of reductionism (the unity of science thesis is primarily a methodological theses in this view) but this is not explicit. Timothy McKeown, one of *DSI*'s critics, notes this explicitly in his 1999 review of the book:

> Although they disclaim any interest in philosophy of science, King, Keohane, and Verba adopt essentially Popperian positions on many important questions. In particular, they emphasize a clear distinction

between forming or stating hypotheses and testing them, an accompanying reluctance to treat hypothesis formation as anything other than an artform, their stress on the need for simplicity in theories, and their insistence on subsuming each case within a class of cases are all highly consistent with logical positivism or Karl Popper's reworking of it. (162)

Though McKeown does not offer an alternative philosophy of science explicitly, he does give an analysis of what he thinks has gone wrong.

Because the statistical worldview embodied in King, Keohane, and Verba's assumption is usually not the worldview that animates case studies, their approach leads to a series of misconceptions about the objectives of case studies and their accomplishments. These misconceptions are constructive, however, in the sense that exposing them leads to a clearer notion not only of the underlying logic behind case studies but also of the importance of nonstatistical thinking and research activity in all research domains – even those dominated by classical statistical data analysis. (162)

McKeown's criticism suggests that there are different goals that case study researchers have and so perhaps understanding what these goals are might aid us in getting a more complete idea of both what different methodologies are for, what sorts of evidence they produce, and so when and where in research they are appropriate. Given that the characterization of *DSI*'s philosophy of science as non-pluralistic and also because of the suggestion that McKeown seems to be making, I am going to suggest that a pluralistic philosophy of science might be helpful here. Without embracing the details of any one of the pluralist views that are currently available (Cartwright's, just to suggest one[3]), I will explore at least one way in which pluralism might help us better understand this debate.

3. Pluralism and causality.

I am going to narrow the remainder of the discussion to causal inference.   So when I turn to examine the question of what role pluralism might play in shedding light here, I will only be talking about pluralism about causality. To talk about causal pluralism and how it might help sort out the debate about causal inference in political science I am going to use the recent work of Julian Reiss.

King, Keohane, and Verba are clear that they do think that causal inferences are one of the goals of political science (they state that descriptive inferences is the other, but I am not going to discuss descriptive inference). They are also clear that causal inferences can be made using quantitative methods. This may not seem completely obvious nor is it completely agreed upon as the authors note at the beginning of their chapter on causal inference. "Many social scientists are uncomfortable with causal inference. They are so wary of the warning that 'correlation is not causation' that they will not state causal hypotheses or draw causal inferences, referring to their research as 'studying association and not causation.'" (1994, 75)   They go on to say that they think that this attitude is evading what they refer to as the Fundamental Problem of Causal Inference, a phrase that they take from Paul Holland (1986), which is that causal conclusions are always uncertain. We will always be uncertain about causal inferences, but this does not mean that we should not make them but rather that we should be "cautious about detailing our uncertainty when we do". Then they proceed to give a theoretical definition of causality that is fundamentally probabilistic. They note that they base the account on Holland's 1986 article with some modifications that they take from Suppes (1970).

I am not going to discuss the details of the account, but do want to emphasize three aspects of it that will be relevant 1) that it is probabilistic and 2) that they give both a quantitative and qualitative example and that 3) that they focus is on causal effects and not causal processes or mechanisms. They do later discuss causal mechanisms and processes and note how they have an "intuitive" appeal but they caution, "However, identifying causal mechanisms requires causal inference…. That is, to demonstrate the causal status of each potential linkage in such a posited mechanism, the investigator would have to define and then estimate the causal effect underlying it. To portray an internally consistent causal mechanism requires using our more fundamental definition of causality…for each link in the chain of causal events." (1994, 86) So, indeed, King, Keohane, and Verba reconfirm their commitment to one logic of (causal) inference.

In their chapter on the philosophy of science, George and Bennett tackle this point about causal inference directly. This is not completely surprising because many case study researchers believe that rather than just having an intuitive appeal, the identification of causal mechanisms that takes place through process-tracing (or perhaps other means) in *case studies* is where the evidence for causal inferences actually lies. George and Bennett make a philosophical argument to that effect, using the work of Wesley Salmon on causal mechanisms and causal interaction. On this account a probabilistic or statistical account can never fully establish causal connections because of a variety of issues (for instance, common causes and conjunctive forks). So we have a confrontation about causal inference that seems to be a key feature of the debate about the role of case studies.

So, let me just summarize before I consider how pluralism might help. *DSI* proposes that there is just one definition of causality and one logic of (causal) inference and that the way it works in quantitative cases should provide a template for qualitative inference. The authors of *DSI* take their definition to be fundamental and so anything that we might want to say about causal mechanisms or processes in case studies depends on being established through this one legitimate mode of causal inference. George and Bennett, among others, argue that *DSI* only succeeds in capturing one sense of causality and that there are others that are legitimate in political science and that case study research is the best way to get at them.

Can a philosophy of science that is pluralistic about causes shed any light on this debate? I turn to Julian Reiss's account in "Causation in the Social Sciences: Evidence, Inference, and Purpose", (2009) for some insight. Reiss argues in this paper for a pluralism about the concept of cause in the social sciences. This would seem as though it might be good news for researchers who use case studies.

Reiss's argument for causal pluralism begins by showing that all extant accounts of causality face unresolvable counter-examples, but of course the conclusion that there are a plurality of concepts of cause does not follow from this. Reiss points out that there could be another response.

> One possible and straightforward response is to loosen the relationship between causation and what one might call the "manifestations" of causation such as counterfactual dependence, correlation, stability under intervention, and so on. The manifestations of causation, according to this response, are not regarded as *defining* causation or as expressing *characteristics universally associated with* causal relationships but rather as providing *evidence* or *test conditions* for the existence of causal relationships. The relation between causation and its manifestations can thus be regarded as roughly equivalent to the relation between theoretical entities in science and their observable counterparts. (2009, 27)

Reiss calls this sort of pluralism "evidential pluralism". He contrasts evidential pluralism with conceptual pluralism, the view that there are many different concepts of cause. If multi-method researchers in political sciences are evidential pluralists, then the various different methods in the social sciences are all worth pursuing because they are all various ways of confirming or disconfirming the causal relation. So while we might have the appropriate statistical evidence to draw the conclusion about a causal connection through multiple regression analysis, a case study which also showed the causal mechanism through which that connection occurred would make us more certain that we had indeed established the case. Something like this view seems to underlie the work of at least some multi-method researchers.

As appealing as this idea might be, Reiss gives an example of a sort of case in which it is not possible to combine conceptual monism and evidential pluralism. He also argues that one cannot save conceptual monism through privileging some one concept of causality that we can infer to from other sorts of concepts of cause that are supported through other sorts of evidence.

Ultimately Reiss concludes that conceptual pluralism is probably correct and that the methodologies that establish particular sorts of causes are specific to those causes. So case study research may be a methodology that reveals causal mechanisms whereas statistical analysis does not, but there may be different purposes for which we would use knowledge of causal mechanisms as opposed to cause in the sense that is established through statistical analysis.

So ultimately the debate about the role of case study research would seem to require that we return to some considerations about philosophy of science. If Reiss is right, then *DSI* is just wrong about there being one sort of causal inference, many of the multi-method researchers who believe that they are using alternative methods to discover the same cause are also wrong. The normative argument that we can take from this philosophy of science would seem to be that researchers must be clear about what their purposes are and what evidence serves those purposes. This means being clear about what sort of cause they are looking for in order to determine what method is most likely to given the evidence that they need. This would be a norm for those operating under the sort of philosophy of science that *DSI* uses.

Though I have not argued it, I want to close by making a preliminary claim about my original hypothesis. If one of the things that political science want knowledge for is to advise on policy then it does seem that they want to be able to say something about causal mechanisms and though there are lots of gaps in the argument, if case studies are a way of getting at them then they should clearly continue to be a core element of political science research. As I said, the argument would most certainly need to be spelled out. I offer only this intuitive motivation for it here, though I can offer an example that is suggestive of how to flesh out that motivation into an argument.

4. *Different purposes: an example.*

In September 2006, James Fearon, a political scientist at Stanford University and an

expert on civil war, testified before U.S. House of Representatives, Committee on

Government Reform, Subcommittee on National Security, Emerging Threats, and

International Relations on "Iraq: Democracy or Civil War?" This testimony took

place prior to the "surge" as was part of the public debate surrounding it.   Fearon's

expert testimony supports that Congress should not deploy additional troops and

his purpose in presenting the arguments is clearly to use his political science

knowledge to help shape policy.

      Fearon is himself an advocate of multi-method research, though his

allegiance lies more closely with the version advocated by *DSI* and his own work is

primarily quantitative.   Reading the testimony, we can see that what is most

compelling are not the statistical arguments but the cases.   He begins by noting that

the conflict in Iraq (at that time) qualified as a civil war on generally agreed upon

understanding of "civil war". He then proceeds to give an account of the likelihood

of achieving successful power-sharing agreements as a resolution to such conflicts

by summarizing statistical arguments in the following way:

> Civil wars typically last much longer than international wars. For civil wars
> beginning since 1945, the average duration has been greater than 10 years,
> with fully half ending in more than seven years (the median). The numbers
> are fairly similar whether we are talking about wars for control of a central
> government, or wars of ethnic separatism.

> When they finally do end, civil wars since 1945 have typically concluded with
> *a decisive military victory* for one side or the other. In contests for control of
> the central state, either the government crushes the rebels (at least 40% of
> 54 cases), or the rebels win control of the center (at least 35% of 54 cases).
> (2006, 3)

What is *not* compelling about such information is that it does not tell us the most

crucial piece of information that we need to know. These are the probabilities, but

as is always the case with probabilities, we want to know if this information is going

to be useful to us in making a decision about *Iraq*. We need to know whether Iraq is

in the relevant reference class.   This is a version of the "reference class problem":

"The reference class problem arises when we want to assign a probability to a single

proposition, *X*, which may be classified in various ways, yet its probability can

change depending on *how* it is classified" (Hájek 2006, 565).[4]   In this case, we want

to be able to assign a probability to the outcome of our intervention in Iraq and we

need to know whether we should classify it as a civil war (Fearon makes a

compelling case that we should), and we need to know *which* type of civil war it is.

In other words, we want a finer-grained partition in order to more accurately assess

the probabilities for *this X* .

> Fearon's next arguments attempt to do just this.

> Quite often, in perhaps 50% of these cases, what makes decisive victory
> possible is the provision or withdrawal of support from a foreign power to
> the government or rebel side. For example, the long civil war in Lebanon
> ended in 1991 after the US and Israel essentially changed their positions and
> became willing to see the Syrian-backed factions win control if this would
> lead to peace. International intervention in civil wars is extremely common
> and often determines the outcome. Power-sharing agreements that divide up
> control of the central government among the combatants are far less
> common than decisive victories. I code at most 9 of 54 cases, or 17%, this
> way. Examples include El Salvador in 1992, South Africa in 1994 and
> Tajikistan in 1998. (2006, 3)

While he does continue to give statistical arguments, he also begins to point to

examples that support a partition of he reference class that would do so more

clearly. These examples come from case studies of particular civil wars and part of

his argument is clearly that Iraq is more like *these* examples than others and so the import of the statistical argument is clarified *for this case*.

This particular use of political science knowledge is suggestive that when the purpose for which we seek the knowledge is policy advice there are not rhetorical reasons for pursuing case study research (we tend to find such arguments more psychologically compelling), but ultimately the knowledge produced from case study research may serve a more directly epistemological purpose, which is to support the examination of a particular partition of the reference class.

The example is at least suggestive that a more careful exploration of methodology and its relation to goals and purposes is worth exploring. Most particularly, it is suggestive that case studies may play a particularly important role in evidence for use.

References

Brady, H. E. and Collier, D. (eds.) 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowen & Littlefield Publishers, Inc.

Cartwright, N. 2007. *Hunting Causes and Using Them: Studies in Philosophy and Economics*. New York: Cambridge University Press.

_____. "Well-ordered Science: Evidence for Use," *Philosophy of Science*, Vol. 73, no. 5, 981-990.

 _____. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Fearon, D. "Iraq: Democracy or Civil War?" 2006 (September 15). U.S. House of Representatives, Committee on Government Reform, Subcommittee on National Security, Emerging Threats, and International Relations downloaded from
http://fsi.stanford.edu/publications/iraq_democracy_or_civil_war/

Gerring, J. 2004. "What is a Case Study and What is it Good For?" *American Political Science Review* Vol. 98, No. 2, 341-354.

George, A. L. and Bennett, A. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Hájek, A. 2007. "The reference class problem is your problem too," *Synthese* 156: 563-585.

Holland, P. 1986. "Statistics and Causal Inference," *Journal of the American Statistical Association* 81: 945–970

King, G., Keohane, R.O., and Verba, S. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

McKeown, T. J. 1999. "Case Studies and the Statistical Worldview: Review of King, Keohane, and Verba's Designing Social Inquiry: Scientific Inference in Qualitative Research," *International Organization*, Vol. 53, No. 1, pp. 161-190.

Reiss, J. 2009. "Causation in the Social Sciences: Evidence, Inference, and Purpose", *Philosophy of the Social Sciences* 39:1, 20-40.

Suppes, P. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.

---

[1] My statement of these theses interprets their views in a narrow sense. It may be that *DSI* interprets the logic of justification more broadly than this statement implies, but I will work with this narrow interpretation as a way of more clearly identifying the possible points of disagreement and the ways in which the defenders of case study research have proceeded.

[2] There are several features of *DSI* that make it difficult to sort out precisely what claims the authors are making. First, it is both a methodological tract and intended as a practical guide for researchers and the tone shifts back and forth from a more theoretical discussion to that of a handbook. Second, though the authors claim that their goal is to connect the two traditions, it is clear that the way in which they are doing so is through reforming qualitative research to conform to their conception of scientific inference. In other words, it is not really through and integration of the two approaches. However, throughout the book they also make comments about qualitative work being useful for hypothesis formation and the identification of variables. These uses of qualitative research are not strictly speaking what the authors describe as part of the logic of inference, but rather might be characterized

as the logic of discovery. This is indeed how some of their critics do characterize their discussion.

[3] I am thinking of Cartwright's *The Dappled World* (1999) , but also more specifically *Hunting Causes and Using Them* (2007)*.*

[4] The name "reference class problem" is originally due to Reichenbach, but I have used Hájek's account here because it seems to me particularly clear and Hájek also argues that the problem is not just a problem for frequentists but for all interpretations of probability.