Ontology, Epistemology, and Multi-Methods

https://doi.org/10.5281/zenodo.939486

Abhishek Chatterjee
University of Virginia
ac7y@virginia.edu

Enthusiasm for multi-methods research can possibly be ascribed to the prima facie promise it holds for moving beyond, if not resolving, seemingly intractable debates on the relative merits of "qualitative" (historical, interpretive, etc.) versus "quantitative" (i.e. inferential statistical) research methods.[1] The justification of multi-methods rests on the claim that combining a few case studies with a larger *inferential*—and not descriptive—statistical study manages to capture the strengths of both insofar as the discovery of causal relations is concerned. This in turn lends greater confidence that the relationships being asserted are indeed causal. The specific argument is that since inferential statistics allows for generalization (while case studies normally do not), and case studies are better at tracing what are called "causal mechanisms," combining the two affords us the best of both worlds.

The trouble with this is that scholars seeking to justify multi-methods seem to assume that the question of what constitutes "cause" or "causal mechanism" is unproblematic, and the problem is limited to that of making causal claims. The problem, in this view, is solely epistemological. Epistemologies however do not exist in vacuum; they are both supported by and in turn support ontologies (or metaphysics), which can roughly be defined as presuppositions or innate conceptions about the nature of the world. An insufficient appreciation of this leads to mutually contradictory arguments in favor of multi-method research designs; arguments, which on reflection could not possibly support such designs. Arguments conceding the usual weaknesses of case studies—but nonetheless attempting to justify them—imply a metaphysics that makes it impossible to portray case studies as either necessary or sufficient in causal analysis, which in turn also precludes any justification of multi-method research. In other words, some fundamental concessions—implicitly based on a specific ontology—negate almost all subsequent justifications that could be made in favor of case studies, and by extension, multi-method designs.

The causal ontology often accepted in pointing out the deficiencies of case studies—implicitly or explicitly—is "reductionist" and "regularist," i.e. one which respectively defines causes in terms of non-causal relations and states of affair and affirms that such non-causal relations are regularities in nature. The origins of this metaphysical view can be traced to David Hume (1999 [1748]: 136)—hence often referred to as "Humean."[2] The particular conception of what it means to make 'causal generalizations' is a logical implication of this ontology of causality. Moreover the idea of 'generalization' cannot be separated from the definition of causality here; in other words to say that something is caused by something else is also to generalize in a certain way, namely, by referring to regularities. Though inferential statistics finds sufficient justification in (and in turn sufficiently justifies) this ontology of causality, explanations based on case studies are not consistent with it. Case studies and inferential statistics cannot logically mix if the definition of causality is reductionist and regularist. This also applies to arguments claiming that case studies illuminate causal mechanisms, since the only definition of "mechanism" that is consistent with this ontology is one that sees them as concatenation of variables that occur with some regularity, something that case studies are not equipped to handle. Multi-methods using case studies can therefore never be justified under this metaphysical view.

Yet (1) referring to regularities is not the only way to generalize, (2) causes do not necessarily have to contain generalizations, and (3) it may not be possible to reduce causes to something more basic. In each of these three cases, one can find sufficient justification for case studies (and also *independently* for inferential statistics), but the usual arguments for combining the two run into logical difficulties. This is because the usual justification for multi-method designs is in fact a confusion of distinct metaphysical views about the nature of causation that are not necessarily complementary. How does one know that the mechanism connecting a cause with an effect in a particular case study is the same mechanism connecting causes to effects in all the other cases? What part of the study does the causal work, the case studies or the statistical analysis? If it is the case study then the statistical analysis should not convince us, and if it is the statistical analysis then the case study should not convince us. This epistemological dilemma arises because the problem is not merely methodological; it involves our fundamental, and most often implicit, metaphysical assumptions about the nature of the world. Let us examine these issues in turn.

That small-N is not merely an epistemological problem becomes evident when we ask under what definition of "cause" should small-N be a problem for establishing causal relations. The answer has to do with statistical theory and the Humean conception of causation that sufficiently—though not necessarily—justifies it. To understand this, let us consider the epistemological and methodological implications of this conception. In other words, given a Humean view, how would one go about discovering causal relations? Now very briefly, Humean definitions come in both deterministic and stochastic versions. Causes precede their effects, and are either necessary, sufficient, or both necessary and sufficient conditions (in the deterministic versions), or increase the conditional probability of their effects (in probabilistic versions). In both cases, every singular causal statement must be an instance of one of more general causal laws. The singular phenomenon itself need not be repeated as long as the unique phenomenon can be shown to be the result of a combination of laws that recur in other singular phenomena. Epistemologically therefore, the singular phenomenon cannot play a role in the establishment of a causal relationship since it is itself dependent on preexisting regularities that have already been established. Both the deductive nomological (D-N) scheme of explanation,proposed most clearly by Hempel and Oppenheim (1948), and Hempel's (1942) inductive-statistical (I-S) scheme follows directly from such conceptions of causation.

The epistemological problem is that of discovering regularities when many laws are instantiated simultaneously. Under ideal conditions experimentation would be the first best method (this obviously is not unique to Humean views; experimentation as a method is consistent with almost all ontologies of causation, but interpretations of experiments would differ depending on the definition of causality). One way to overcome the problem of simultaneous instantiation would be to isolate individual causes and observe their effects repeatedly to establish lawlike regularities. When we move from the experimental sciences to the social or non-experimental sciences, the goal remains the same, i.e. the discovery of regularities, but this time they have to be detected from purely observational data. This is where statistical models come in. Such models try to approximate the experimental situations described above. These models assume that the data being generated are akin to the result of a series of independent experiments or observations generated from mutually independent processes where nature manipulates the independent or explanatory variable under different background conditions or controls (again, it is also possible to give other *interpretations* to inferential statistics). Inferential statistics is also consistent with the definition of causes as generalizations; that is, the "regularity" part of the definition, or alternatively the definition of causes as "types." The latter is obviously because insofar as it informs one of average effects, generalization (over a particular population) is built into the interpretation of inferential statistics.

The link between a reductionist and regularist metaphysics on one hand and inferential statistical methods on the other should be clearer now. The impossibility of fitting case studies into this framework should also be evident. Indeed, some prior discussions in political science have clearly recognized this. For example, Sartori defended comparative case studies as a third-best method behind experiments and statistical studies (1994:16). His argument was that though it is true when it comes to drawing causal inferences, comparative case studies are inferior to either experiments or statistical control, the phenomena that most interest certain political scientists do not occur enough times to lend themselves to statistical studies. The problem with this defense is that the acceptance of the logic of statistical inference entails that a few cases cannot or should not lead us to believe that a cause exists. This is the crux of Lieberson's (1991) argument against drawing causal conclusions from a few comparative cases (also see Sekhon, 2004). Using the example of automobile accidents, Lieberson shows how fragile our conclusions can be as to the causes of accidents if we rely on only a few cases, assuming that knowledge of causes entails knowledge of regularities. The most logical conclusion in this instance would be to state that given the paucity of cases one cannot say anything about the presence or absence of causes.

Lieberson's critique applies equally to solutions to the problem that urge us to somehow increase the number of cases by, among other things, performing "within case analyses" by looking at multiple implications or consequences of a particular theory or causal statement within the same case (Campbell 1975: 184–189). But if we assume that regularities are most basic and knowledge of causes entails knowledge of regularities, it is difficult to count multiple implications as an augmentation of the number of cases. For at a given level of analysis, each implication of any causal statement must be considered separately. It is for this reason that statistical models require each observation to be independent. And multiple implications of the same causal statement or theory can never be considered independent from each other. There is a rebuttal to Lieberson's argument, but as we shall shortly see, it makes sense only within decidedly non-Humean ontology of causation. Within the Humean ontology, Lieberson's position is very convincing indeed.

Again, early discussions seemed to have conceded this. Still case studies were defended variously as "a first stage of research, in which hypotheses are carefully formulated," (Lijphart 1971: 685), or as explications of particular cases for their own sake in light of theory, as in Verba's "disciplined configurative approach," (1967: 114–115) among others. In such an approach the researcher seeks to explain the event with the help of established regularities and general causal statements. It is important to note here that though disciplined configurative explanations rest on general laws, the explanation itself does very little to strengthen or weaken the validity of the said laws (Lijphart 1971: 692). Yet these concessions are sometimes accompanied with arguments that cannot easily be reconciled with the former. Thus, for example, Lijphart's subsequent assertion that such studies can be considered "crucial experiments" if values on the variables are extreme is difficult to reconcile with his statement quoted above. Why should extreme values on variables in one case cause us to reexamine our prior theory, especially since the latter could be based on a large number of cases? The same applies for 'deviant case' analyses. As Mckeown (1999) has also observed in a slightly different context, a single additional case can never, by this logic, lead us to weaken an original proposition that is, in Lijphart's own words, "solidly based on a large number of cases" (1971: 692). The problem is that some of Lijphart's epistemological points about the contributions of case studies make sense only when decoupled from his ontological orientation which seems to underlie the bulk of his other points.

Another popular defense of case studies—that such studies are better at handling determinism (Gerring 2004: 347, Munck 1998: 33)—is based on conflation of ontology with epistemology. It is perfectly consistent to have a deterministic and Humean view of causality—indeed, the original Humean view was in fact deterministic and some philosophers have argued that "Hume and indeterminism don't mix" (Dupre and Cartwright 1988)—and still claim at the epistemological level that statistical inference is the best way to establish this causality. As Laplace observed a long time ago, an (ontologically) deterministic relation can appear to be (epistemologically) stochastic because of ignorance of all relevant laws and initial conditions. It does not matter whether the view of causality is deterministic or stochastic as long as causation is reduced, and it is reduced to regularities either deterministic or stochastic. In both situations case studies can never be logically justified as the "first best" method. Similar interpretations can also be given to the use of inferential statistics in political science. As a result, criticisms such as Lieberson's against the use of Mill's methods would still be valid. What we discover from Mill's methods cannot even be considered "cause" in this sense.

The reference above was to deterministic sufficient conditions. But can deterministic necessary conditions justify case studies, as Dion (1998) has argued? Dion's argument protects case studies against the small-N criticism only under extremely restrictive conditions. The argument has more to do with the problems that classical inferential statistics faces in tackling necessary conditions than the inherent strengths of case studies. In fact it could be seen primarily as an advocacy of Bayesian statistics over classical statistics when it comes to necessary conditions.

Since Bayes' rule depends crucially on known probabilities to determine posterior probabilities, its applicability is limited to only certain kinds of systems. To be precise, it is crucial that the mechanism that generates prior probabilities is well-known, and alternative hypotheses have well-known probability outcomes or likelihoods. The prior probability is a source of great debate in both philosophy and statistics (See Sober 2002, for example). It is uncontroversial in cases of systems where there is a clear way of assigning prior probabilities. But it is slightly more controversial in cases where we can't. Then the question becomes what should the prior probabilities be based on? Should they be based on statistical regularities, "common sense," case studies, or subjective opinions? As soon as we ask these questions, we realize that we are back to the square one. Additionally, and more pertinent to the use of such statistics to defend small-N's is the fact that we would have to consider multiple hypotheses with determinate likelihoods for effective statistical control; at which point the difference in terms of sample size between classical inferential statistics and Bayesianism begins to disappear. Even this argument, as a result, cannot provide sufficient justification for case studies.

This brings us to the final and most popular set of justifications for both case studies and their incorporation in multi-methods research, namely, that case studies are uniquely suited to discover or enunciate what are called "causal mechanisms," which statistical studies are less able to do. However, "mechanism" is yet undefined. Further, of two possible understandings of the concept (of mechanism), one does not provide any justification for case studies, while the other—while sufficiently justifying case studies—cannot easily support their incorporation in multi-method designs.

If mechanisms are defined as, "in effect, variables that operate in sequence," (Sambanis 2004: 288), or any variation thereof, some of the same criticisms that we started with apply. The difficulty of defending case studies while holding this particular understanding of mechanisms stems from the fact that that it implies just another version of the Humean definition extended to intervening variables. It is theoretically possible to multiply the number of steps between cause and effect while remaining steadfastly Humean. Each link or mechanism in a longer chain can be represented by equations that can be construed as statements of regularity and as such the same epistemological concerns that were raised earlier about the confirmation of causal claims with case studies apply here too. Various statistical models such as path models would seem to be the natural recourse. If this is a fair representation of some definitions of causal mechanisms, then again the sufficiency of case studies cannot be defended.

More avenues open up once we abandon either reductionist or regularist (or both) understandings of the concept of "mechanism." But these latter conceptions, though equally supportive of inferential statistics independently, cannot easily accommodate the usual manner of performing multi-method research without running into logical contradictions.

"Singularist" definitions of causality hold that singular events and not regularities are more basic. The definition decouples generalizations from the definition of causality (Ducasse 1993; Salmon 1980, 1997). Epistemologically, therefore, one need not look for generalities, and the explanation of a single event or case can count as a causal explanation. Process tracing in case studies receives sufficient metaphysical justification here. But this ontology presents us with a problem. Such reductionist but singularist definitions of causality have difficulty distinguishing spurious causes from "real" causes at the definitional level. One way of overcoming this is to attach counterfactuals to singularist mechanisms. Counterfactuals, however, are very sensitive to contrast spaces. The truth condition of counterfactuals depends on the contrast space of any explanation and therefore causality also becomes context and contrast space dependent in this case. So, for instance, causes of revolutions as opposed to near-revolutions can be very different from causes of revolutions as opposed to non-revolutions, or revolution in country A as opposed to revolution in country B. Generalizations, if any, in this case are "bottom up" and change based on the relevant contrast spaces rather than "top-down" and ostensibly universal. Furthermore since singular events are more basic, there is no expectation that generalizations will necessarily emerge. But if we define contrast spaces with as much generality as possible, for instance in our example above, as all possible cases of near-revolution, and if we call answers to both kinds of questions (the limited and expanded contrasts, respectively) "cause," certain problems recur at the epistemological level in combining methods since there is no presumption that an answer to one question will have any bearing on an answer to the other. Thus though attaching counterfactuals to singular mechanisms suffices to justify case studies, they cannot justify multi-method research.

For instance Evan Lieberman's (2005) latest attempt to suggest a framework for multi-method research faces this particular problem. He writes that "a nested research design implies that scholars will pose questions in forms such as "What causes social revolutions?," while simultaneously asking questions such as 'What was the cause of social revolution in France?'" (2005: 436) For an answer to both questions to qualify as "causes" almost necessarily implies a singularist view of causation. Under a regularist view an answer to the second question cannot differ from an answer to the first, and the former has to be at least a subset of the latter. His advice is to start with a large-N analysis and then—in case of robust and satisfactory results—"test" the model with small-N analysis by choosing cases that fall within the average prediction of the large-N model (2005: 437). Why should we expect the small-N cases to be consistent with the large-N predictions? Even if they are, why should we have any confidence that the average prediction of the large-N analysis and case study research point to the same causal relationship? In the absence of robustness Lieberman advises model building and analysis of predictions that fall in the average, and also the outliers (2005: 439–440). The criteria for "robustness" and "satisfaction" must be statistical; it is therefore difficult to see why lack of robustness should motivate case studies. Indeed there are well-known remedies within inferential statistics for such problems as lack of statistical significance or any bias in a model and none of these involve looking at case studies. Note that all the questions raised here do not imply that Lieberman is wrong, but that the argument contains large gaps, owing to insufficient appreciation of the metaphysical implications of methods. Additional arguments have to be supplied to reconcile mixing of the two methods.

Case studies also receive sufficient justification if we abandon a reductionist view of causation or causal mechanisms. This would reverse the order of priority in the relationship between regularities and causes. Instead of regularities being signifiers or definers of causes, prior knowledge of causes would restrict and inform the kinds of inferences one is able to make from statistical relations. This is also an effective rejoinder to Lieberson's criticism of the comparative method. This is part of Nancy Cartwright's argument for considering causal "capacities" as primitive. She contends that it is the arrangement of capacities in certain ways that produce regularities; "nomological machines," or "socio-economic machines" as she calls them, are particular arrangements of capacities that "in the right sort of stable (enough) environment will, with repeated operation, give rise to the kind of regular behavior that we represent in our scientific laws." (1999: 50) Capacities, further, cannot be identified by any particular manifestation. They can be compared to qualities such as kindness or tenacity that are carried by human beings. Such qualities are not identified with any one particular behavior; instead they are instantiated in multiple circumstances as different behaviors all of which have in common the fact that they are displays of kindness or tenacity (1999: 51). Socioeconomic machines are essentially fables that illuminate important aspects of how the world works, while capacities can be equated with morals of such fables. The relationship between the fable and the moral is not that of similarity but "that of the general to the more specific…[e]ach particular is a case of the general under which it falls" (1999: 39). This means inter alia that "satisfying the associated concrete description that applies on a particular occasion is what satisfying the abstract description consists of in that occasion" (Ibid). Thus any particular arrangement of capacities is also general, and in turn, every general capacity finds its manifestation only in particular arrangements. Once we understand capacities well enough (as is the case in certain natural sciences) we can further manipulate these capacities and arrange them in different ways to produce different laws. As Cartwright observes, "anything *can* cause anything else. In fact, it seems…not implausible to think that, with the right kind of nomological machine, almost anything can *necessitate* anything else" (1999: 72).

The epistemological consequences of this view urge us to treat both (most) large-N statistical studies and case studies as essentially alike in that both can be interpreted as attempting to 'guess' the arrangement of hypothesized capacities in the world. Sometimes when we know about enough capacities and other background conditions "[w]e accept laws on apparently slim experimental bases…[and] the data plus the description of the experimental set-up deductively imply the law to be established" (Cartwright 1999: 93). Case studies, both single and comparative, can therefore be considered similar to fables that substantiate morals. The fables however have to be very carefully constructed with great attention to capacities and their arrangements. They are necessarily concrete, but they are at the same time general. This is precisely why studies like John Gaventa's (1980) of one particular locality in one country are also general. Notice that domain restriction finds its best justification under this ontology. In fact if we follow this logic, restrictions of domains is imperative, since what we are describing are particular nomological machines, the very definitions of which carry the connotation of restriction. This is because as we observed earlier, it is the arrangement of particular capacities in certain orders and under certain conditions that could generate laws. But domain restriction does not mean restricted generalization. The fact that some physical laws are literally true only within the confines of the laboratory does not prevent them from also being general. This answers certain criticisms of case studies based on their domain restriction. Thus to say that domain restriction in case studies necessarily implies limited causal force is to implicitly accept an ontology that cannot justify case studies in the first place.

Even in this case, however, the usual manner of combining case studies with a large-N (inferential) statistical analysis cannot be logically supported because of the reasons pointed out earlier. On the other hand, one way of avoiding the usual contradictions in mixing methods would be to truly "triangulate" within the general framework of a case study. In other words instead of using the usual procedures of picking one case out of any sample, one could try to empirically describe the arrangement of capacities (of course, in the context of prior background knowledge of capacities) of any one case, and then examine the implications of such an arrangement using quantitative evidence. This would work because as Cartwright pointed out, it is the particular arrangement of capacities that produces regularities. But it must be a necessary preliminary to first explain why and how the arrangement of capacities came to be. This kind of suggestion is most relevant to the literature on institutions in political science and sociology, especially the ones based on single cases.

Notes

[1] Another possible but independent reason, particularly of interest to those interested the sociology of knowledge is that multimethod research, especially when used in doctoral theses signals to potential employers competency in both statistical and others kinds of research methods thus satisfying the largest possible coalition of potential employers. To reiterate, this is one possible hypothesis in need of further study, and will not be addressed further in this contribution.

[2] Though modern versions are significantly different from what Hume originally may have suggested.

References

Campbell, Donald T. 1975. "'Degrees of Freedom' and the Case Study." *Comparative Political Studies* 8:2, 178–193.

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

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:2, 127–145.

Ducasse, C.J. 1993. "On the Nature and the Observability of the Causal Relation." In *Causation*. Ernest Sosa and Michael Tooley, eds. (New York: Oxford University Press), 125–136.

Dupre, John and Nancy Cartwright. 1988. "Probability and Causality: Why Hume and Indeterminism Don't Mix." *Nous* 22:4, 521–536.

Gaventa, John. 1980. *Power and Powerlessness: Quiescence and Rebellion in the Appalachian Valley.* Urbana: University of Illinois Press.

Gerring, John. 2004. "What is a Case Study and What is it Good For?" *American Political Science Review* 98:2, 341–354.

Hempel, Carl G. 1942. "The Function of General Laws in History." *Journal of Philosophy* 39, 35–48.

Hempel, Carl G. and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15, 135–175.

Hume, David. 1999 [1748]. *An Enquiry Concerning Human Understanding*. Tom L. Beauchamp, ed. (New York: Oxford University Press).

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99:3, 435–452.

Lieberson, Stanley. 1991. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70:2, 307–320.

Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65:3, 682–693.

McKeown, Timothy J. 1999. "Case Studies and the Statistical Worldview." *International Organization* 53:1, 161–190.

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

Salmon, Wesley. 1980. "Causality: Production and Propagation." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.

Salmon, Wesley. 1997. *Causality and Explanation*. Oxford: Oxford University Press.

Sambanis, Nicholas. 2004. "Using Case Studies to Expand Economic Models of Civil War." *Perspectives On Politics* 2:2, 259–279.

Sartori, Giovanni. 1994. "Compare Why and How: Comparing, Miscomparing and the Comparative Method." In *Comparing Nations: Concepts, Strategies, Substance*. Mattei Dogan and Ali Kazancigil, eds. (Oxford: Wiley-Blackwell), 14–34.

Sober, Elliott. 2002. "Bayesianism—Its Scope and Limits." In *Bayes' Theorem*. Richard Swinburne, ed. (Oxford: Oxford University Press), 21–38.

Sekhon, Jasjeet S. 2004. "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals." *Perspectives on Politics* 2:2, 281–293.

Verba, Sidney. 1967. "Some Dilemmas in Comparative Research." *World Politics* 20:1, 111–127.

Speedbumps on the Road to Multi-Method Consensus in Comparative Politics

Michael Coppedge
University of Notre Dame
coppedge.1@nd.edu

Is there a multimethod consensus in comparative politics? My short answer is: not quite. For example, recently I was updating my department's reading list on research methods for the comprehensive exam in comparative politics, and I added a chapter by Lakatos (Lakatos 1970) to it and sent it to my colleagues for feedback. One replied, "I'm especially glad to see Lakatos added!" Another replied, "What is Lakatos doing in there?" (Actually, my colleagues are unusually collegial.) But there is very little consensus on any aspect of comparative politics, so it is unrealistic to expect anything resembling consensus in our subfield (España-Nájera, Márquez, and Vasquez 2003).

My longer answer is that there is rough agreement in principle that multimethod work would be a good thing. There is also agreement that, in practice, some aspects of multimethod work are hard to pull off. But I think that there are some other challenges in multimethod work that are not yet sufficiently appreciated—speedbumps on the road to the great multimethod harmonic convergence.

On the encouraging side, we agree that we can do case studies to verify causal mechanisms or explore anomalies identified by statistical analyses or formal theories; we can do statistical analyses to test whether arguments generated by case studies or formal theories are generally true; we can develop formal theories to explain tendencies turned up by case studies or statistical analyses; and so on, with many variations (Lieberman 2005).

On the discouraging side, we can agree that it is hard for any one researcher to develop cutting-edge expertise in all three methods, and that not being on the cutting edge can be an obstacle to publishing multimethod work. One can overcome this obstacle by collaborating with those who have greater expertise in different methods, but each person tends to feel that he or she is doing more work and getting only partial credit for it. And there is some truth to that (Bennett and Braumoeller 2009). These difficulties are well known and accepted. But I think there are other obstacles to multimethod work that have not received as much attention, and yet remain serious obstacles. The first is the mismatch between concepts used by different approaches. The second is disagreement