**Using Cases in Political Science: Reflections on Keith Dowding’s ‘The Philosophy and Methods of Political Science’**

Author D

**Introduction**

It is not unheard of for entries in methodological symposiums such as this one to begin with some forced pleasantries and proceed quite rapidly to scathing ad hominem attacks. This one will not. Keith Dowding’s *Philosophy and Methods in Political Science* is quite simply a book that is full of learning. The book is clearly written by someone fairly uniquely placed to bring together a coherent and incisive account of empirical, conceptual and normative work. Threading these worlds together are many reflections that loosely make up a rather good ethnography of the discipline of political science itself. The book is peppered with both discerning and amusing commentaries on some of the undesirable cultures within political science, evoking sometimes schadenfreude and sometimes discomfiture in the political scientist reading it. This is a book for students of political science that goes a long way to explaining what political science is, as well as what it should be. To that end it is a book that students should read early and often.

Decorated veterans of the great methodological wars in the discipline who come to this book expecting (perhaps hoping) for chagrin are, on my reading at least, likely to be a little disappointed. Dowding’s account of explanation in political science makes more claims to unity than hegemony. Much of the book provides essential ground-clearing for debates that have tended to become contrived over time. The book crystallises the Popperian account of falsifiability, prompting a transparent discussion of proper criteria for theory-shifting and consideration of outliers. Rather than dwell unduly on a zero-sum account of what causation is, or searching for a definitive champion among forms of explanation, Dowding instead asks when causal analysis is useful and where different forms of analysis can be usefully applied to provide explanation. He argues effectively that a single case study is not ever better than large-N analysis in demonstrating causality because of the specification problem, but that the case study has other values for explanation.

This short response draws out some thoughts prompted by my reading of these debates. In particular I want to press further on considerations about the N of cases and so-called qualitative and quantitative evidence in political research. I accept Dowding’s critique of the claims of causal process-tracing as those claims appear in some methodological textbooks, but I do not think that such an idealised case-study exists in real-life political science. Choosing a case is an integral part of performing a case study and case studies in reality are always situated within research designs that are attentive to a case’s relationship to others. And, I think as soon as we admit to this and focus our attentions away from an abstract ideal type of case-study, the critique becomes less than ideal also. Dowding’s work in my view could be complimented by a further elaboration of its pragmatic implications for how we consider and favour certain approaches to sampling and types of data in our field. In widening the debate, I ask why systematic reviews and meta-analyses are conspicuous by their absence in political science. Finally, I return to a discussion of what this means for disciplinary culture and development. Dowding entreats all political scientists to have the courage of their convictions; to “live their lives as though they believe the ontological and epistemological claims they profess” (p.2). I argue that we need to do more to recognise when identity conflicts of political scientists contribute negatively to cumulative advancement of disciplinary knowledge.

**Qualitative and quantitative research: It’s not the size of your N…**

First; a little ground-clearing of my own. If I had a pound for every time I have heard a colleague declare that the characterisation of a qualitative and quantitative divide in political research is ‘overblown’, ‘unhelpful’ or ‘no longer useful’ I would be a rich man. But if I had a penny for every time such a colleague subsequently used the terms in earnest I’d be richer still. Unfortunately, it is not a vice I am invulnerable to. Moreover, I do not blame any author trained in social research for leaning on the terms as heuristic devices out of habit. Dowding’s work prompts us to think harder about why disagreements founded on such concepts might be problematic. The impact of the terms on the discipline means we are unlikely to rid ourselves of them any time soon (although this is no reason not to try, if we agree on their detriment). We should at least strive when we apply the terms to be very conscious of their implications.

Too often the labels qualitative and quantitative can be used as a catch-all to contrast approaches to scholarship, without any insight into whether it is (a particular combination of) the numbers of cases, approaches to causation, useful methods, or something else that is the at the nub of the conceptual delineation. Dowding’s book is written with exceptional clarity but there is some unconscious slippage into using the designations qualitative and quantitative as vague signifiers in it too. Qualitative and quantitative are conceptually slippery terms because they have, it seems to me, no clear negation. I struggle to imagine what a study with no quantity or no quality (in the sense of having no distinguishable attributes rather than ‘not being very good’) would look like. The heuristic value of ‘qualitative’ and ‘quantitative’ as signifiers of ideal-typic approaches is probably not worth the confusion encouraged by their everyday use. As concepts they have high normative content. Even accepting for the sake of argument that we can distinguish instances of more quantitative political research that consider more cases from more qualitative work that says more about some case(s), the claim that the relationship between the two is more or less linear and inverse is a value judgement masquerading as empirics. Such a designation connotes all kinds of assumptions about approaches to ‘casing’, what counts as an ‘attribute’, and what the purpose of research is. The use of small-N/large-N is preferable for the discussion of explanation on which Dowding focuses to which I now turn. Small-N/Large-N are better signifiers of different approaches but I think there is still not as much of a difference between large-N research and small-N research as is assumed by many philosophical accounts.

**So what about numbers of cases?**

Dowding takes his aim at those authors who he accuses of over-claiming the contributions that a single-case study can make to causal analysis. He argues that single cases, acting alone as it were, cannot establish causation because of the specification problem. Instead he suggests a ‘diagnostic’ use for case studies; “once potential mechanisms have been established (or plausibly identified) by large-N studies, careful qualitative work can suggest which of several potential mechanisms might have been operative in a particular case” (p. 152). He suggests that these diagnostic explanations are vital, as they give us a narration of a causal mechanism that we can be more satisfied with psychologically than the ‘distal-correlational’ generalization that a given X predicts Y with a probability p under conditions W (p.146-7).

Dowding is also quite sceptical of the merits of a ‘crucial case study’ method. The idea that a single case could confirm or disconfirm a hypothesis by virtue of its set-up is problematic. A case might uncover something that changes our opinions but it cannot “shift the odds”, that a theory or its rival is true “that far on its own,” (p.116). Dowding carefully argues that a single case can only ever represent one data point in generalizations; and therefore adding extra causes (or ‘bumps’) to a mechanistic explanation within a case does not affect the type-explanation any more than any other case could have. Of course a case-study might suggest rival hypotheses to be explored on their merits, but if I have it right, Dowding suggests that the single case cannot bear an extra-special weight on general explanations because of a discovery at another level of granularity, at the point of that discovery.

I am in agreement with Dowding for the most part. What he provides is a crystal clear explanation as to how one can over-egg the value of particular types of evidence. Dowding’s arguments might hold for an idealised single case study, but I want to press further on the issue of the weight of evidence from particular cases and the role of research design in small-N analysis by posing two questions: I want to ask 1) how much real case-studies align with this idealised version and 2) where along the small-N/large-N continuum the token and type explanations converge and diverge.

**Sampling – not just a lonely case**

It is interesting that Dowding has little enough to say about case-selection strategies (or sampling) as a result of his analysis. Arendt Lijphart seminally provided a typology of six different case-study approaches (1970). The implication of Dowding’s analysis is that this typology ought to be reduced by at least one or possibly folded into just two categories – diagnostic and conjectural. To pump my own intuitions (something Dowding has taught me to do more often!), I am left wondering if, when a student approaches me for guidance on whether a particular case is worth studying over another, I should ask them ‘are you motivated to diagnose the mechanism at work in the case or build a hypothesis from the story of the case?’, as Dowding suggests I might do; or; whether I will fall back into thinking about the extent to which the case is deviant, typical, crucial, exhibits maximum variation etc., as I might have done before reading this book. I am tempted still towards the latter.

All research involves casing – deciding what cases to study and what those cases are cases of (Ragin and Becker 1992). We distinguish the case and the universe of cases to which it pertains. This process can be more or less iterative. A researcher often will not know what her case(s) are cases of until she writes up an explanation. How we get there matters less than how we got there, in that it is how we communicate our research design at the end that bears scrutiny – if it is done transparently and truthfully any research can be assessed on merits. The process of reflexively imposing, conceptualising or operationalising[[1]](#footnote-1) an understanding of an instance of a phenomenon on the data one aims to collect (and actually collects) never happens in a vacuum – there is always information that influences this process, and information that is systematically discarded.

We make judgement calls about the relevant scope of a particular population, and what new things we can expect to learn by studying a number of instances of that population. Case studies, then more often have an implicit comparative element because they are seen to be members of a class which is related to other classes through relatively common notions of theory and expertise. They are part of a larger-N collaborative and cumulative programme of comparative research and their designers know it.

Dowding is aware of this. He considers that as pattern-seekers if we try to characterise a case as completely unique we are admitting that we think we can learn nothing from historical cases (p. 97), and that research questions and cumulation of research are always dependent on knowledge of background conditions (p. 98; 146). Therefore what is missing for me from his account is a clear treatment of what information-oriented sampling (Flyvbjerg 2006), brings to bear on how evidence affects the odds that a hypothesis is true relative to its rival.

Each sampling decision in research is a mixture of information-oriented and random or formulaic sampling. Even completely randomly generated samples require arguments about quality of a frame, and machine learning requires some initial direction. Good sampling/case-selection then is the art of reducing sampling error. The problem of sampling has not as much to do with numbers of cases studied (though this is a consideration), as it has to do with availability of information to make reasonable judgements about how samples relate to populations and what populations look like. Dowding tells us that “One problem for token narrative accounts is that the decision process of what to leave out is less systematic than when data are collected according to a specific scheme. Hence the probability of bias is much greater in qualitative than quantitative analysis,” (p.59). This is not my, nor indeed others, experience of small-N and large-N research (Flyvbjerg 2006). Those who engage in small-N research necessarily spend much time thinking about what kind of an example of the class of cases a case is. They focus some of their data-gathering activity on the case itself, but part of their analysis (their iterative sampling) carefully considers arguments about what the prior expectations for the case are; the distribution of known cases; reasonable assumptions about what kinds of cases are likely to occur, and the context of the case and what conditions scope the population. It is in-depth analysis of a case that confirms degrees to which it is a member of a class and exhibits particular variation; as well as what that variation means for explaining outcomes.

None of this is strictlyinconsistent I think with Dowding’s claims. He is clear on the role mechanisms play in signifying the need to sometimes engage in conceptual analysis and add/remove conditions to models in order to accommodate empirical findings. However, what is not clear is the degree to which he will allow that certain cases are more likely to yield results that are beneficial to corroborating or falsifying rival hypotheses than others. Without all the know-how I have alluded to each case is only one data point, but with work on casing and populations, cases become extreme, maximum variant, deviant, or most-likely/ least-likely (the latter more “relaxed variants of the crucial case” as Rohlfing puts it (2012: 84)), etc. Effective case selection strategies for small-N research activate more cases beyond those being studied.

**The use of case-study data in comparative political science**

The point of dialogue between smaller- and larger-N research perhaps becomes clearer when we move to research that explicitly compares a handful or more cases. To my mind the vast majority of political science still takes the form of small-N comparisons (sometimes ‘the comparative method’), or at least incorporates the logics of small-N reasoning at some level of abstraction. There is sub-disciplinary variation in approaches of course, but the state remains the foremost unit of political organisation. Therefore, even where large-scale individual-level data collection and analysis is carried out, it often takes the form of within-case analysis that is formally or informally nested within a smaller-N comparison of units of political organisation with carefully imposed scope conditions e.g. we can compare the effects of immigration on public support for right-wing parties within western democracies.

Again because Dowding’s argument focuses on the reappropriation of token, narrative accounts as beneficial for diagnostic explanation, he has less to say in scrutiny of small-N comparative research/medium-N research; or of large-N research’s claims and contributions to explanation. Where does prediction take over from diagnosis? Should Small-N comparativists comparing, say, four countries claim to be making four diagnoses, and separately adding four data points to a very modest empirical generalisation? While there is nothing egregious in such an account, my worry is that the presentation draws too sharp a distinction between token and type explanation and closes doors on dynamic research design. When small-N research is well-conceived it motivates variables in interesting ways across cases that create a dialogue between token and type explanations. Such a dialogue can alert researchers to specific problems of redundant determinants in token accounts and determinants that are absent from type accounts. These insights come about because researchers are comparing token accounts with many determinants and type accounts with fewer in more or less the same breadth.

*Systematic Cumulation of Research*

Dowding argues the key contribution of process-tracing is that it provides the mechanism that allows us to understand and accept probability-based explanations that rely on correlations of more distal data. The account Dowding gives of that relationship is enlightening and a key reason why anyone should read his book. A further unmentioned contribution of process-tracing though is the provision of robust data for large-N analysis. Process-tracing observations are not directly comparable in that they may come from different sources and exhibit different levels of measurement. In process-tracing a sequence of conditions is mapped out revealing a chain of proximate causes. For each single case then, there may be many data points, assiduously measured. The data may not be comparable for the research question at hand in the case-study, but when taken together, case-studies leave behind sets of observations that may be cumulated for comparative studies at a more abstract level. It is the recognition and demarcation of different ‘bumps’ (or recordings of the degree of presence of particular variables in the case) that create this data.

It is surprising that political scientists have not spent much time trying to consider how to harvest data from case studies for comparison, or report and record case studies in such a way that they can be more easily systematically cumulated. The wealth of data that sits on shelves and servers is enormous. Political scientists have not yet become unduly interested in systematically reviewing evidence and providing meta-analyses. They should be interested in cumulating evidence more systematically – and case-studies can help. It is my hope that changing practices in the use of data-repositories among small-N researchers, as well as the availability of methods such as qualitative comparative analysis for systematic review should see more cumulative evidence presented by political scientists in future. There is more to say then, it seems, about the contribution of various forms of case analysis to the forms of explanation with which Dowding is concerned.

**Conclusion: Cooperation in Political Science**

While I have asked for clarifications on some aspects of Dowding’s account of the use of cases in political research, I want to conclude on the same subject he does in the book – the profession of political science itself. I think part of the reason for an absence of cumulation is a preference often, when approaches are challenged (or misrepresented through rhetorical obfuscation), for subsequent retreat into sub-disciplinary silos.

Dowding’s calls to avoid -ismic ad-homenism are extremely important, because pigeon-holing of particular contributors stultifies exchange across the profession. I share some of Dowding’s frustrations in my observations of debates across sub-disciplines and cultural approaches. Too often writers are allowed to get away with attributions of ‘-isms’ or other shorthand identifications in order to dismiss a researcher or an entire group of researchers and their nuanced approaches to explanation. Of course some political scientists may identify strongly with rival arguments or approaches and this can lead to debates that are highly polemical but reveal new insights based on the merits of arguments. That is when disciplinary pluralism brings out the best in us. When critiques of arguments slip into dismissals of identities or labels, the profession suffers. Such labelling is too often used to avoid serious engagement with other approaches. There is little or no pluralism in valuing plural approaches to political science only when it suits arguments. And following Dowding, if we profess to value a plurality of approaches we should conduct our political science as though we believe those claims and keep engaging.

Political Science journals and articles collectively have some of the lowest impact factors across academic disciplines. As crude a measure as that may represent where citations are influenced by disciplinary culture, it still indicates a comparative lack of impact of much work across the profession. Ad-hominem attacks increasingly incentivise returns to disciplinary silos and the pop-up industry of new journals that are often no more than echo chambers for disaffected groups. An important feature of Dowding’s work, beyond the sage nuggets of advice on how to negotiate professional challenges, is that his account of explanation does not attempt to elevate an explanatory hegemon or defame other writers. Other political scientists would do well to follow his lead and always ask if they are contributing to cumulation of knowledge or not.

**References**

Dowding, Keith (2016) *The Philosophy and Methods of Political Science*. Basingstoke: Palgrave Macmillan.

Flyvbjerg, Bent (2006) “Five misunderstandings about case study research” *Qualitative Inquiry* 12(2): 219-245.

Lijphart, Arend (1970) “Comparative politics and the comparative method.” *American Political Science Review* 65: 682-93.

Ragin, Charles C. and Howard S. Becker (Eds.) (1992) *What is a case? Exploring the foundations of Social Inquiry* Cambridge: Cambridge University Press.

Rohlfing, Ingo (2012) *Case Studies and Causal Inference: An Integrative Framework* Basingstoke: Palgrave

1. Readers of a different ‘–ismistic’ hue might prefer to use a different form of words here but their actions in carrying out a case study I don’t think would differ much. [↑](#footnote-ref-1)