Comparing Societies: Qualitative Methods

Julian Go

As scholars in the 1970s took up the task of crafting rigorous methods for comparative social science, they all confronted similar problems. One major problem was that cross-national comparisons have a fairly small number of cases to draw from. There are only so many countries in the world (less than 200). In addition, qualitative research demands in-depth knowledge of cases which limits the possible number of cases even further. This problem became known as the “many variables, small-N” problem (Lijphart 1971, 686). As the number of cases decreases relative to the number of possible explanatory variables, the ability to infer causal relationships and adjudicate competing explanations falls perilously (see also Lieberson 1991). For this reason Lijphart (1971) famously suggested that comparative studies should be “the first stage of research, in which hypotheses are formulated” and that statistical studies based upon a larger-N should be the final goal (Lijphart 1971, 685).

So how to improve a small-N comparative study? One proposition was to increase the number of observations by adding more countries or by extending the time frame of the study to include historical cases. Another was to decrease the number of explanatory variables by either combining variables into a single scale or by a careful strategic selection of cases. This latter approach most closely approximates Mill’s “method of difference”. It has been called the strategy of “comparable cases” (Lijphart 1975), “most similar systems” design (Przeworski and Teune 1970) or “controlled comparison” (Eggan 1954). The goal is to select cases that minimize variation or, in statistical parlance, to control for “concomitant variation”. That is, rather than select randomly (as statistical analyzes might), researchers restrict the cases according to a similarity principle: the cases for the investigation are chosen because they are as similar as possible. This decreases the number of explanatory variables, thereby “controlling” extraneous variance, and in turn facilitates a specification of the cause (Faure 1994, 310).
These two strategies—the first of extending the cases, the second of minimizing variables—are both useful ways of improving the validity of causal inference in small-N studies. Various studies in comparative politics and comparative sociology after Lijphart and Smelser employed them implicitly or explicitly. The “comparable cases” strategy came to dominate the field of comparative politics. Lijphart saw it as definitive of the comparative method (DeFelice 1986, 417–18).

The comparative method can now be defined as the method of testing hypothesized empirical relationships among variables on the basis of the same logic that guides the statistical method, but in which the cases are selected in such a way as to maximize the variance of the independent variables and to minimize the variance of the control variables. (Lijphart 1975, 164)

But while these early methods were widely influential and probably helped to further legitimate qualitative cross-societal comparisons, they also carried limitations.

The first strategy of adding more cases runs into two problems. One is that the new observations must be independent from the others and from each other lest they fall prey to “Galton’s problem,” in which researchers err by treating a set of interrelated cases as if they were independent instances of a particular pattern, rather than just a single related set. This may be difficult when adding observations from backwards in time and also when adding countries who are adjacent to each other or tied together economically (King, Keohane and Verba 1994, 222). The other problem with expanding the number of the cases is the “conceptual stretching” problem (Sartori 1970) (Munck 2004, 113). This is the problem of extending concepts from one set of cases to another where they do not fit. If a researcher wants to examine political participation in different societies, and the researcher adds more countries to the study, they have to be sure that what counts as political participation is the same in the new cases (Przeworski and Teune 1970, 92–103). The problem of conceptual validity is especially dangerous for studies that aim to expand cases by historical extension as Lijphart suggested. If the researcher wants to add 16th-century England as a case, they have to somehow ensure that participation in that century means the same as in the 21st century. The problem is also cross-cultural: in some national contexts political participation may largely mean voting, but in other contexts it might mean protesting (Collier and Mahoney 1993, 845).

The other strategy (of adopting Mill’s method of difference) also carries problems. One is that the findings will likely be “overdetermined”: any variable that differs across the cases could be logically said to be the cause, and there may be an infinite number of such variables (Guy 1998, 36–39). The other problem is that it cannot detect interaction effects. It assumes that the controlled factors being held constant across cases have no impact when in fact a constant factor may be interacting with another variable to produce the effect. It assumes that the controlled factors being held constant across cases have no impact when in fact a constant factor may be interacting with another variable to produce the effect. Leiberson (1991) gives an example: assume two drunk drivers are speeding past a red light, but only one of them faces a car entering from the right and is therefore hit leading to an accident. In this scenario, using the method of difference would lead to the conclusion that the “cause” of the accident was the car entering from the right rather than drunk driving over the speed limit through red lights (Leiberson 1991, 313–14).

Because of these problems, some scholars have preferred to employ Mill’s method of agreement. Here the outcome is constant across the cases and the cause is found by looking at which explanatory variables are also constant. There is no attempt to “control” for explanatory variables through the selection of cases. In choosing cases, one only selects the cases that share the outcome. Skocpol’s (1979) analysis of revolutions has been taken to be exemplary of this method. Seeking an explanation for social revolutions, she chooses cases where revolutions occurred. She ends up with revolutions in China, France, and Russia and then asks what they all shared which might have led to the revolutions (Skocpol 1979). Przeworski and Teune (1970) called this method the “most different systems” design and contended that it was superior to Mill’s method of difference because it facilitates a Popperian elimination of rival explanations (see also Collier 1996, 73).

Critics register two main objections to this method. The first is that, like the method of difference, it cannot detect overdetermined outcomes, i.e. multiple causes or conjunctural causation (Lieberson 1991, 314–15) (Ragin 1987, 38). The other main objection is that, by selecting on the dependent variable, researchers are “truncating” and will get biased estimates of causal effects (it leads to the overrepresentation of positive cases that, in regression studies, bias results) (Collier and Mahoney