

Comparative analysis: case-oriented versus variable-oriented research

Donatella della Porta

Comparative analysis holds a central place in social science research. There is a well-established view in the social sciences that it should be based on variables (see Héritier, ch. 4, and Schmitter, ch. 14). Yet much research – especially in political science, but also in some branches of sociology – is case-oriented: that is, it aims at rich descriptions of a few instances of a certain phenomenon. This chapter argues that both approaches are legitimate. Variable-oriented studies mainly aim at establishing generalized relationships between variables, while case-oriented research seeks to understand complex units. Some people would argue that case-based comparisons follow a different logic of research, while others insist that the rules are essentially the same.

The chapter starts by introducing the debate on comparative analysis, distinguishing the experimental, statistical and ‘comparative’ methods. We then single out two main strategies of research, presenting their origins in the methodological reflections by Durkheim and Weber, and focusing on the assumptions that are linked to the variable-oriented and case-oriented approaches, respectively. Advantages and disadvantages of each will be discussed on the basis of illustrations from social science works on democratization, political violence and political participation, looking at examples of large-*N* statistical research designs and contrasting them with small-*N* comparisons, especially in the tradition of historical sociology. The chapter also discusses recent attempts to bridge the gap between the two approaches, in particular with qualitative comparative analysis (QCA) and recent reflections on the case-oriented strategy. Conditions that might influence the choice of one logic or the other include environmental conditions (such as stages in a research cycle or types of data available) and researchers’ epistemological preferences as to approach and methodological skills. We then look at strategies

I am grateful to Marco Giugni, Michael Keating, Leonardo Morlino, Philippe Schmitter, Pascal Vennesson and Claudius Wagemann for helpful comments on previous versions of this chapter.

for comparative analysis, addressing some of the main methodological choices: the relevant unit of analysis; the number of cases; the trade-off between most-similar and most-different designs; and ways in which to address the time dimension.

One or two logics: the debate on comparative politics

Comparative analysis

In sociology and political science, there has been a strong tendency to consider all social sciences as following one and the same logic (see della Porta and Keating, chs. 1 and 2; King, Keohane and Verba 1994). Validity depends on following a set of rules of scientific inference, whose purpose is attempting to infer beyond the immediate data to something broader that is not directly observed.¹ If we accept this, the differences among research projects will refer to the matrix of data: large-*N* (statistical) research designs cover many cases, while small-*N* (comparative) studies only a few. From this perspective, larger-*N* projects are considered stronger in providing valid and significant inferences. In the 1970s, in sociology and political science, this position was supported by the influential works of Neil Smelser and Arendt Lijphart, both dealing with comparative analysis.

Comparative analysis responds to this need for broadening the territorial scope and depth of political information (Lasswell 1968). It has often been understood as that branch of political science concerned with comparing nations (Verba 1991). Yet, the debate on the comparative approach has played an important role in the more general methodological discussion across the social sciences.

The field of comparative politics boomed in the 1960s, in line with the acknowledgement of an 'accelerated interdependence of the world arena' (Lasswell 1968: 3). Comparative political scientists extended their range of interest from Western democracies to second- and third-world countries, shifting their concerns from formal institutions to the political process. At first, theories of development dominated the field, with a strong emphasis on global comparison and the normative aim of bringing Western-style economic and political development to underdeveloped countries. The Vietnam War brought to light the dramatic effects of such interventions, justified with the purpose of helping developing countries. In the 1970s, assumptions about a unique pattern of development in political and economic life based on

Western experience were criticized, and the developmental approach was attacked for its 'Cold War origins and overtones' (Wiarda 1991: 21). With the renewal of attention to cross-national comparison in the 1980s, the hopes for global theories were abandoned, together with the developmental approach, leaving space for various *middle-range theories* (that is, those that are meant to hold only for specific societies) in various subfields of the discipline. Comparative politics has indeed been described as aiming at developing concepts and generalizations at a middle level, between what is always true for all societies and what is true of only one society at a single point in time (Bendix 1963: 532).

Reflection on the specificity of the comparative approach remained central, however. It was defined at the outset as one of the scientific methods available to control hypotheses on the relations between two or more variables, keeping constant (or parametrizing) all potentially disturbing elements. The empirical control of hypotheses requires a distinction between conditions treated as *parameters* (which are assumed, or made not to vary) and causal conditions treated as *operative variables* which, in a specific investigation, are instead allowed (or made) to vary in order to assess their influence (Smelser 1976). Three main approaches exist within comparative analysis: the experimental method, the statistical method and the comparative method; all perform, with declining strengths, the task of converting most of the variables into parameters in order to isolate the effects of the remaining variables. There is, as so often in the social sciences, a certain terminological confusion here. Sometimes the term 'comparative method' is used to cover all three approaches; elsewhere, it is restricted to one of them. For clarity, we will use the umbrella term 'comparative analysis' to cover all three, and 'comparative method' for the third of them.²

In the *experimental* method, conversion of variables into parameters is achieved in the creation of data. In an artificial setting, we control the effect of any changes in the values of an operative variable on the values of the other operational variables, by keeping all other potential influences stable. In an experimental situation (as used in the natural sciences and some social sciences, notably psychology), it is possible to allow for changes only in the variable on which we focus our attention by, typically, taking two identical groups and introducing a stimulus in only one of them. All of the differences between the two groups may thus be attributed to that one stimulus. In this sense, the method is very strong, offering robust criteria to choose between rival theories (Lijphart 1971). Unfortunately, only a limited number of social phenomena may be investigated via experiments.

The *statistical* method – based on mathematical elaboration of empirically relevant data (Lijphart 1971; see also Franklin, ch. 13) – approximates the experimental method by intervening after the data are created. It is already weaker than the experimental method as a means of making inferences, insofar as parametrization is obtained via the mathematical elaboration of empirical evidence (Smelser 1976: 157), typically by creating subsamples in which potentially disturbing variables are kept constant. Although the statistical method is weaker than the experimental method, it still provides good tests for eliminating rival theories. The main problem of the statistical method is the need for large samples: the higher the number of variables that potentially ‘disturb’ the measuring of a correlation coefficient, the larger the number of cases needed in order to build subsamples large enough to be statistically significant. This is not only very expensive, but also often impossible because of the limited number of homogeneous macro-units endowed with particular characteristics.

The term *comparative method* is used, rather confusingly, for an approach within comparative analysis that provides an alternative to the statistical method. When the number of cases is too low for statistical manipulation, the investigator approximates it ‘though without the same degree of confidence – by systematic comparative illustration’ (Smelser 1976: 157). The comparative method supplements with *logical reasoning* the lack of a sufficient number of cases for systematic tests via partial correlations. For scholars like Smelser and Lijphart, the logic of the comparative method is identical to that of the other methods, ‘in that it attempts to develop explanations by the systematic manipulation of parameters and operative variables’ (Smelser 1976: 158). Like the other methods, it aims at establishing general, empirical relations between two variables and controlling them by keeping all other variables constant (Lijphart 1971). In this sense, the comparative method adopts the same logic as the statistical method, adapting it to those situations in which we deal with complex phenomena without the large number of cases necessary for a statistical analysis: the famous situations of ‘many variables, small N ’ (Lijphart 1971: 686). Timothy McKeown (2004) suggests that the belief that there is a single quantitative logic to all empirical social scientific research reflects the idea that all empirical research faces the same problems of causal inference as quantitative research does. This implies assumptions such as the existence of a clear distinction between the formation and the testing of hypotheses, the search for simplicity (if not parsimony) in theory, and the pre-allocation of each case within a class of cases. In a variable-oriented research design, the lower the number of cases, the fewer should be the explanatory variables, since degree-of-freedom problems would make the research design indeterminate.

The logic is, however, the same: ‘the comparative method resembles the statistical method in all respects except one. The crucial difference is that the number of cases it deals with is too small to permit systematic control by means of partial correlation’ (Lijphart 1971: 684). Conversely, ‘[a]s soon as the number of units becomes large enough to permit the use of statistical techniques, the line between the two is crossed’ (Smelser 1976: 161).

Dealing with a small number of cases – usually between two and twenty – the comparative method is a preferred strategy for political and social scientists when they investigate institutions or other macropolitical phenomena. In fact, the comparative method is considered the only choice for controlling hypotheses that apply to large units that are too few for statistical analysis. Although in this approach the quality of control of the relationship between variables is low, it is often the *only* scientific method available for the study of macrodimensional, interdimensional and institutional processes (Eisenstadt 1968).

Case-oriented versus variable-oriented: diverse tools, shared standards?

This assimilation of statistical and comparative methods into ‘one and the same logic’ did not, however, remain unchallenged. Some scholars, while agreeing on the search for shared standards, warned about the need to keep in mind the methodological implications of the use of diverse tools (Brady and Collier 2004).

Indeed, the divide between those analysing a large number of cases on a few characteristics and those studying a few cases in depth (that is, looking at a large number of dimensions, usually within a historical perspective) has been growing with the specialization of the social sciences. Given this plurality of approaches, the insistence on a single logic by King, Keohane and Verba (1994) has been criticized for ignoring the differences among the many objectives social scientists might pursue on the basis of their ontological beliefs about ‘the extent to which different “truths” are accessible to human observers, the level of abstraction at which “truths” are to be formulated, and the extent to which these “truths” can be generalized across contexts’ (Sil 2004: 314; see also della Porta and Keating, ch. 2).³

Yet in many research designs, the choice of the comparative method is not just a second-best one imposed by the availability of data; rather, it is justified by its capacity to go beyond descriptive statistical measures, towards an in-depth understanding of historical processes and individual motivations. Ragin and Zaret (1983) suggested two decades ago that there are *two* different logics in comparative politics (or social sciences in general), often addressed

Table 11.1. Durkheim versus Weber: the ‘logics’

	Durkheim	Weber
Aim at . . .	Generalization: search for trans-historical, permanent causes (different from historical contingencies)	Complexity: search for limited generalizations about historical divergence and concrete knowledge about specific processes
Relying upon (Mill’s) mode of . . .	Concomitant variation as logic of analysis	Methods of agreement and differences
Instrument of analysis is . . .	Statistical correlation, regression	Narrative
Understanding explanation as . . .	Explanation as functional proposition about patterns of relations among abstract variables; singling out (external) causes	Explanation as genetic (combinatory) understanding of historical diversity; singling out (internal) reasons
Through the construction of . . .	Social species (discrete types of society) – as intermediate between the confused multitude of societies and the single although ideal concept of humanity	Ideal types (hypothetical models developed as aids for explanations: enable generalization about historical divergence)

by contrasting Durkheim’s and Weber’s research approaches. I suggest in the following that it is indeed useful to rehearse the debate, not to challenge the need for *shared standards*, but in consideration of the specificity of *diverse tools* when prescribing methodological standards. The differing research ‘logics’ linked to Durkheim and Weber have been compared on various dimensions (Table 11.1; see also della Porta and Keating, ch. 2).

First of all, many scholars have pointed at the *different aims* present in a scientific enterprise. In statistical comparison, we aim at building law-like propositions. For Durkheim, sociology as a science must favour generalizations over details: ‘Sociological explanation consists exclusively in establishing relationships of causality, that a phenomenon must be joined to its cause, or, on the contrary, a cause to its useful effects’ (Durkheim 1982: 147).⁴ As we are going to see in what follows, survey-based research on political participation

is aimed at singling out the average effects of some variables (such as level of education, or interest in politics) upon the use of different forms of collective action. As Mahoney and Goertz (2006) recently put it, in this logic of research the aim is to estimate the average effects of independent variables, that is to investigate the 'effects-of-causes'. In historical comparison, *à la* Weber, the aim is the in-depth understanding of a context (or the searches for the 'causes-of-effects', *ibid.*). The case-oriented strategy focuses upon a relatively small number of cases, analysed with attention to each case as an interpretable whole (Ragin 2000: 22), seeking to understand a complex unity rather than establish relationships between variables. Studies oriented at understanding the reasons for the strength of, say, the Italian Communist Party or nationalist political violence in Ireland are illustrations of this type of approach.

A related issue is the *logical tools* used for the explanation. Referring to John Stuart Mill's work (1843), methodologists have observed that the variable-oriented and case-oriented approaches use different logical 'canons'. While statistical analyses are based on the search for *concomitant variations* (that is, looking at whether independent and dependent variables vary together, with regression as the main instrument for measuring causal inference), comparative analyses use the methods of similarities and differences. In the Durkheimian approach, concomitant variation is considered 'the supreme instrument for sociological research' (Durkheim 1982: 153). Statistical techniques based on a probabilistic logic allow for generalizations, even when the explanation is not valid for each single case. According to the method of *agreement*, if two or more instances of a phenomenon under investigation have only one of several possible causal circumstances in common, the cause of the phenomenon is the one circumstance that is present in all the analysed instances (Ragin 1987: 36). In this sense, we proceed by looking for invariant patterns, eliminating as potential causes all variables on which the units have different values. Mill's method of *difference* assumes that when two or more cases have different values on a certain phenomenon we want to explain, we have to look for the one circumstance on which they differ. Although the determinism of the search for necessary causes has been criticized as unrealistic for the social sciences (Lieberson 1994), the search for necessary conditions has been considered of substantive relevance for social theory (Goertz 2003). Focusing on a small-*N*, case-oriented comparison usually points at similarities and differences through dense *narratives*, with a large number of characteristics being taken into account, often together with their interaction within long-lasting processes.

There is also, however, a deeper difference between variable-oriented and case-oriented research, and this refers to the very concept of *explanation*. Neil

Smelser (1976: 204) admits the differences in the ‘modes of comprehension’ with an ‘ideographic–nomothetic dilemma’. When looking at aggregated cases, the researcher is typically interested in the variables that affect one another causally. When focusing on individual cases, however, she might aim at an understanding of a complex unit, by grasping the relations among its constituent parts. He or she is not looking for a causal explanation, but rather, in Smelser’s words, ‘the operation may be more akin to an “appreciative” or “esthetic” act, an effort to understand the principles by which the parts consistently fit together’. While Smelser seemed to consider the second type of knowledge as somewhat residual, more balanced assessments developed later on. Recently, Ferejohn (2004: 150) has distinguished external, more or less causal explanations, and internal, or deliberative, explanations. *External* explanations present agents doing things because of some configuration of causal influence, while *internal* explanation identifies reasons for an action. Thus, ‘An action is explained internally as an outcome of a deliberative process in which the agent is assumed to act for reason . . . To “explain” in this sense is to “justify”’ (Ferejohn 2004: 152; see also Pizzorno, ch. 9).⁵ Statistical analysis on large-*N* cases of typical instances of political violence try to assess the contextual conditions that facilitate their development. Recently, analysis of the distribution of car burnings (as indicators of the intensity of urban riots) per municipality has been oriented to explain the French urban riots in autumn 2005 on the basis of some characteristics of the areas in which riots were more prevalent (that is, more cars were burned). In this way, characteristics such as spatial segregation, the level of poverty or rates of unemployment have been identified as causes for rioting (Lagrange and Oberti 2006). In a different perspective, ethnographic research has identified the motive of the rioters, that is the justification of their actions, in the development injustice frames (Auyero 2007).

Various *heuristic devices* are developed for working towards these different aims. In Durkheim’s work, inductive reasoning on empirical data aims at reconstructing the different social species – which he locates between ‘the confused multitude of historical society and the unique, although ideal, concept of humanity’ (Durkheim 1982: 109). The properties of a social species influence the course of the social phenomena developing within them, since ‘the causes of social phenomena are internal to the society’ (*ibid.*: 114). The search for permanent causes implies a focus upon explanations that point at patterns of relations among abstract variables that are trans-historical in nature (*ibid.*: 739). Since concomitant variation is usually oriented to the search for permanent causes (Ragin and Zaret 1983: 737), there is no space for plural causation: an effect cannot have different causes in different contexts.

In this approach, understanding the recourse to political violence in given systems would imply, for instance, finding the correlation coefficients of various indicators of potential contextual preconditions (such as the degree of democratization and per capita income) with indicators of the spread of political violence (such as the number of people wounded/killed in political events or for political reasons and the amount of material damage during protest events).

In a case-oriented approach, by contrast, an in-depth knowledge of a small number of cases provides the basis for generalizations that are temporarily limited to the cases studied and whose wider relevance should be controlled through further research. Macro-units (such as countries) are therefore considered as unique and complex social configurations (Skocpol and Somers 1980), even though concepts are built that transcend the validity of individual cases (see Goldthorpe 2000: ch. 3). In qualitative, historical comparison based on a case strategy, explanations are genetic (i.e. based upon the reconstruction of the origins of a certain event), and generalizations are historically concrete (Ragin and Zaret 1983: 740). Theorization and generalization, in this tradition, are provided not by statistical regularities but by *ideal types*. These are abstract models, with an internal logic, against which real, complex cases can be measured. An ideal type, Weber (1949: 90) explains, 'is no "hypothesis" but it offers guidance to the construction of hypotheses. It is not a description of the reality but it aims to give unambiguous means of expression to such a description'; it is an 'idea', a 'unified ideal construct', 'abstracted out of certain features' and keeping the 'essential features' (*ibid.*: 91). This analytical construct is 'ideal' in the sense that it allows singling out relationships which 'our imagination accepts as plausibly motivated and "objectively possible"' (*ibid.*: 91–2). It is oriented to facilitate the empirical analysis, without reflecting either an ethical imperative, or a historical reality. As Ragin and Zaret (1983: 731–2) noted, ideal types enable limited generalization about historical divergence, pointing to different patterns of process and structure in history. Such generalizations go beyond the uniqueness of historical events, although without approaching the degree of generality of natural scientific laws.

In this approach, understanding political violence would imply in-depth description of the contexts in which violence developed, locating the specific process of evolution of violent political actors in their broader environment. The existence of several different paths to the same outcome is largely accepted and anticipated; for instance, similar degrees of political violence in different countries or times might well be produced by different causes (or

combinations of causes). The presence of unemployment could be very important to explaining violence in some historical contexts, as it is combined with other elements (for instance, presence of armed militia); but it may be totally irrelevant in others.

Summarizing, we can distinguish a case-oriented approach from a variable-oriented one on the basis of different concepts of ‘understanding’: related either to generalizable knowledge of relations among variables (aiming at generalization), or to dense knowledge of cases. Some comparativists use case-oriented strategies in order to understand or interpret specific cases because of their intrinsic value; many, however, also have a causal-analytic purpose (Ragin 1987). A valuable feature of the case-oriented approach is the development of an extensive dialogue between the researcher’s ideas and the data in an examination of each case as a complex set of relationships, which allows causal complexity to be addressed.

As I will argue in what follows (see Table 11.2), these differences affect the research design. In particular, the characteristic of comparison as a method that respects the historical specificity of the units under analysis is contrasted with the sort of ‘anonymity’ of the cases belonging to a statistical sample. Variable-based projects tend to follow (or mimic) statistical rules: a high N is considered as preferable; in particular, the logic of variable-based research design implies that with a small number of cases, we can cope with only a small number of variables. Explanation is understood as measuring the different variables’ contributions to causing a certain phenomenon (how the dependent variable covaries with each independent variable). The assumption of homogeneity of the units of analysis (see Héritier, ch. 4) is made at the beginning of the research. Here, ‘generality is given precedence over complexity’, and therefore ‘the wider the population, the better’ (Ragin 1987: 54). Time is used mainly for increasing the number of cases by building subunits, through periodization or as points of observation within longitudinal studies.

In contrast, case-based logic tends to explore diversity (and deviant cases) by thick description of one or a small number of cases, often contrasted on several dimensions. This means that a few cases are analysed based on a large number of characteristics. Explanations are narrative accounts with limited interest in generalization. The degree to which the cases selected do belong to the same category, and therefore are comparable, is assessed in the course of the research itself (Ragin 1994). The method is not very sensitive to the frequency distribution, and a single case can cast doubt on a cause–effect relationship established on the basis of many observations (Ragin 1987). Time is especially useful here in order to build narratives of processes.

Table 11.2. Research design in variable-based versus case-based comparisons

	Variable-based	Case-based
Cases as	Anonymous (transformed into variables)	Names with capitals (complex units)
Concepts	Predetermined and operationalized	Constructed during the research
Independence of cases	Assumes cases that are independent from each other	Addresses systematic process analysis
Number of cases	Increase <i>N</i> whenever possible	Keep <i>N</i> low
Number of variables	Reduce the number of variables in order to avoid undetermined research design (degrees of freedom problem)	Increase number of variables in order to make the description thicker (full accounts; case knowledge)
Case selection	Tend to select randomly or on the independent variable	Tend to select paradigmatic cases
Diversity as . . .	Parametrization – search for generalization in area studies or subsystem research project	Understanding through differences – exploring diversity
Use of time	Periodization	Processes and temporal sequences; eventful temporality

We shall discuss these elements in more detail in what follows.

Definition of case and case selection

All of these differences in the research logics (or tools) must be taken into account when dealing with the steps of a research design, an important one being the selection of cases.

What is a case?

First of all, the process of defining cases (*casing*) is different. In variable-oriented research, the homogeneity of the units of analysis is stated at the very beginning,

when defining the population of cases, considered as empirically given (Ragin 2000). In case-oriented research, cases tend not to be determined at the beginning of a research project – instead, ‘they often coalesce in the course of the research through a systematic dialogue of ideas and evidence’ (Ragin 2004: 127). In this process of casing, singling out the degree of homogeneity of the cases (by answering the question ‘What is this a case of?’) is part and parcel of the research process, which often ends with the construction of types and the allocation of cases to them.

This difference is linked with the different function and timing of conceptualization (see Mair, ch. 10, and Kratochwil, ch. 5): concepts are predefined and then operationalized at the onset of the research in a variable-oriented design; and constructed (in their sociological meaning) in the course of the research in a case-oriented design. Additionally, it reflects differences in the consideration of the unit of analysis: in variable-oriented approaches, statistical procedures decompose the original cases into values on variables, while in case-oriented approaches they maintain their unitary character; that is, even when variables are mentioned, the single cases are still approached as complex units (Corbetta 2003: 18; see also della Porta and Keating, ch. 1). In variable-oriented approaches, the cases become anonymous; in case-oriented ones, they are complex units, given capitalized labels.

The number of cases

The two ‘logics’ also have different implications for the *number of cases*. As noted, comparison by variable tends to privilege large *N*: ‘because the comparative method must be considered the weaker method, it is usually advisable to shift to the statistical method if sufficient cases are available for investigation’ (Lijphart 1975: 165). In a similar vein, Giovanni Sartori (1971: 8, emphasis added) agrees that ‘comparison is a control method of generalizations, provisions or laws in the form of “if . . . then”, that may be used in *cases in which stronger methods are not available*’.

The issue of the number of cases is dealt with in variable-oriented research designs by some specific rules oriented to address the issue of the degree of freedom (see Franklin, ch. 13). The number of cases should vary according to the number of variables included in a research design: the larger the number, the more likely that regression coefficients are statistically significant. Indeterminate research designs – with a smaller number of cases than required by the number of operational variables – are defined as designs from which ‘virtually nothing can be learned about the causal hypotheses’, since the

researcher has ‘more inferences to make than implications observed’ (King, Keohane and Verba 1994: 118–19). An increase in the number of variables would require an increase in the number of cases or, if this is not feasible, a refocusing of the study on the effects of particular explanatory variables rather than on the causes of a particular set of effects.

Within this approach, case studies are considered useful mainly for the falsification of hypotheses or their specification through the analysis of *deviant* cases. In Lijphart’s (1975) view, case study stands apart from other methods in that it cannot produce empirical generalizations, nor be used to test hypotheses.⁶ The case study is ‘a system for questioning, not for answering’ (Stretton 1969: 247), and its context-dependent, ideographic knowledge is considered less useful for social sciences than the general knowledge derived from large-*N*, variable-oriented studies.

Case-oriented researchers, on the other hand, oppose the suggestion that increasing the number of cases produces ‘better-determined’ research designs. They stress, first of all, the methodological losses involved, especially in cross-national comparison, with increasing the *N*. First, an increase in the number of cases normally brings about an increase in the number of *third variables* – that is, of variables external to the hypothesis we want to control – thus reducing the reliability of our inference or imposing a further increase in *N* (on this point, see Morlino 1990: 387–8). Especially in cross-national research projects, including new countries augments the problem of concept-stretching (Munck 2004; Mair, ch. 10) as well as of the reliability and comparability of measures and indicators used to translate national experiences into comparable operational categories (Mair 1996). Working with many countries or long historical periods, in a field in which few reliable and comparative ‘hard data’ are available, increases the risk of building on insufficiently deep knowledge of each single country. More generally, some scepticism has been voiced about the ability, even in large-*N*, non-experimental research designs, to have enough observations to adjudicate among rival explanations. In this sense, the differences between experimental designs and statistical ones have been noted – so much so that some have found it ‘problematic to suggest that any observational study can ever be “determinate”’ (Collier, Brady and Seawright 2004b: 236). While experiments are indeed capable of keeping the third variable constant, causation can only be inferred in observational studies if the researcher imposes ‘several restrictive assumptions, which may be difficult to test or even to defend’ (Collier, Seawright and Munck 2004: 48). Indeterminacy can also derive from multicollinearity, where two or more independent variables move together.

Some scholars emphasize the contribution of interpretative work, and of other qualitative approaches, to goals that a regression-oriented framework addresses less successfully – including concept formation and fine-grained description (Brady and Collier 2004; Collier, Seawright and Munck 2004). Case-oriented studies are said to be stronger in these two tasks, as well as in research programmes oriented towards understanding the cognitive protocols that capture the actors' definition of the situation (McKeown 2004: 153). They are also considered particularly effective in identifying causal processes and therefore in developing theories. So 'seen in this light, the test of a hypothesis – the central theoretical activity from the standpoint of conventional quantitative research – is but one phase in a long, involved process of making sense of new phenomena' (McKeown 2004: 167).

Recent debates on case studies and small-*N* comparison have challenged the idea that – as Dietrich Rueschemeyer (2003: 305) put it – exploring the impact of a large number of relevant factors and conditions in only a few cases does not help in learning anything that is theoretically relevant. Case studies are praised for their detailed knowledge of processes (at different moments, or 'data points' in Rueschemeyer's definition), considered as particularly useful for the discovery of social mechanisms (see Héritier, ch. 4). In this sense, a case study goes beyond a single observation, and confronting analytical propositions with many data points can be useful not only for theory building but also for theory testing.⁷

The selection of cases

The choice of number of cases is linked with that of the types of cases. In variable-oriented designs, methods of sample selection are usually constrained by statistical rules. Random samples (or stratified ones) are preferred when the main aim is to randomize unwanted sources of variation (Smelser 1976: 211). As King, Keohane and Verba stated, 'if we have to abandon randomness, *as it is usually the case in political science research*, we must do it with caution' (1994: 124, emphasis added). So they accept that random sampling is only one of the possible ways of selecting cases, with some obvious advantages but difficult preconditions of applicability. Not only in qualitative research, but also in much quantitative research, random selection might not be feasible because the universe of cases is not clearly specified. Even when feasible, it is not always the best strategy, given the risk of missing important cases. In these situations, they suggest selecting observations that would ensure variation in the explanatory variable and the control variables. King, Keohane and

Verba, in fact, follow a long tradition of insisting that we should never sample on the dependent variable. It may be tempting, in this way, to search for observations that fit our theory, but selecting only cases with the same value (or a limited range of values) on the dependent variable would prevent us making any causal (statistical) inference about the relationships between the dependent and independent variables. This is because cases with different values on the dependent variable could, for all we know, be correlated with the same independent variable. For example, we could take a group of cities that had experienced riots and find that they all had high levels of unemployment. Yet it is possible that other cities, which had not experienced riots and which we therefore did not consider, also had high unemployment; hence unemployment cannot be the critical variable.

Case-oriented research follows a different strategy of case selection. Selection of cases for small-*N* research is, in this perspective, not to be evaluated on the basis of the classical rules oriented to avoiding selection biases in statistical (especially regression) analysis. In particular, selecting on the dependent variable is a quite common and legitimate practice. Case-oriented researchers may intentionally select cases that differ relatively little from each other with respect to the outcome under investigation (Ragin 2004), focusing in particular on positive cases, that is cases where a phenomenon (such as revolution) is present. There are analytic gains to be derived from an in-depth analysis of positive cases of a phenomenon like revolution, especially when little is known about it (Collier, Seawright and Munck 2004: 48), or from the higher capacity to evaluate the impact of a main causal variable by focusing on cases with high scores on both the dependent and the independent variables (Collier, Mahoney and Seawright 2004: 102). Typically, research on peasant revolts or revolutions or anti-WTO riots focuses on cases in which those phenomena developed, without taking into account the entire range of variation in outcome. In contrast to variable-oriented analysis, the selection of cases in case-oriented research requires an appreciation of their relevance for a specific set of hypotheses. Additionally, some cases are considered as more substantially important and non-conforming cases are evaluated in detail. So, a theory of revolutions that is unable to account for the 1789 revolution in France would be highly problematic.

For research following both strategies, criteria for case selection have been suggested. Smelser (1976: 174) has listed five criteria that may guide our choices: units of analysis 'must be [1] appropriate to the kind of theoretical problem posed by the investigator . . . [2] relevant to the phenomenon being studied . . . [3] empirically invariant with respect to their classificatory crite-

... [4] reflect the degree of availability of data referring to this unit ... and [5] decisions to select and classify units of analysis should be based on standardized and repeatable procedures'. All selection of cases implies, however, trade-offs among what Gerring (2001) called:

- *plenitude*, referring to the number of cases: the larger the number of cases used to posit a causal relation, the higher the confidence in the results; additionally, large samples help in specifying propositions;
- *boundedness*, referring to the range of generalizability and therefore the inclusion of relevant cases, but also the exclusion of irrelevant ones (*ibid.*: 172);
- *comparability*, referring to the similarity among cases on some relevant dimensions;⁸
- *independence*, referring to the autonomy of units: if a unit is strictly linked to another, one risks studying the same unit twice;
- *representativeness*, referring to the capacity of the sample to reflect the properties of the entire population;
- *variation*, referring to the range of values registered on relevant variables;
- *analytical utility*, with reference to the theory to test, or the scientific approach chosen;
- *replicability*, referring to the possibility of replicating the study.

Even with these specifications, comparative social science remains a wide field with many strategies of comparison, and scholars' preferences on the number of cases have varied over time. In the 1960s, large-scale comparisons were at the core of an increasing attention to comparative politics. After the 1970s, there was a resurgence of comparisons of a small number of countries, often analysed over long periods (Collier 1990). Growing attention to interpretative social sciences stressed the relevance of 'thick descriptions' of few cases (Geertz 1973).⁹ In the early 1990s, much of the work aimed at a limited generalizability, with middle-range or even lower-level theories for which the specificities of the historical context played a crucial role (Mair 1996). More recently, the preferred number of cases has increased again, under external pressures such as the development of new statistical methodologies for multicase comparison and the enlargement of the European Union.

An intermediate strategy is offered by Charles Ragin (1987, 1994, 2000), in his *qualitative comparative analysis*. Based upon Boolean algebra, this relies upon medium-*N* comparison based upon analysis of similarities and differences in a search for necessary and sufficient conditions. It compares *configurations of causes* – that is, the effects of the contemporaneous presence/absence of a combination of factors, not of the presence or absence of

each of them. Although still following a deterministic logic, it allows for multiple causation through the analysis of several different combinations of causes.

Especially within neoinstitutional approaches, historical sociology or international relations, the use of case studies continues, however, to be considered as a main strategy in order to address complex historical phenomena (see Vennesson, ch.12, and Steinmo, ch. 7).

Similar versus different cases

Preferences vary, not only on the number of cases, but also on the right balance of similarities/differences among them. Two different strategies have been identified: the so-called *most-similar systems* design, in which we compare similar cases, and the *most-different systems* design, where we compare dissimilar ones.

Working with *similar systems* (for example, similar countries) facilitates the *ceteris paribus* rule – that is, it reduces the number of ‘disturbing’ variables to be kept under control. For Lijphart (1975), cases for comparative analysis should be selected in such a way as to maximize the variance of the independent variables but minimize the variance of the control variables. Within a most-similar systems design, we assume that factors common to the countries sampled are irrelevant in explaining some observed differences, focusing instead on the variables that are different. If we want to explain why left-wing terrorism spread in the 1970s in Italy, but not in France, we would mention neither the presence of a Communist party nor of a pluralistic system of industrial relations, since these were present in both.

In many fields of sociology and political science, cross-national comparisons often address countries belonging to a common geographical area (such as southern Europe or eastern Europe) and sharing historical traditions, cultural traits or economic development. The advantage is that many variables are ‘parametrized’: if we have more or less the same degree of economic development, similar culture and the like, we can consider these characteristics as constant and check for the influence of other factors. In area studies, the relative similarity of situations enables an appreciation of the marginal difference and its causes (Dogan and Pelassy 1990: 134).

A disadvantage, however, is that, in comparing similar systems, we cannot go beyond so-called middle-range theories – theories that apply only in a restricted area. An additional problem is that comparison of similar cases still

leaves open a risk of overdetermination (Przeworski and Teune 1970), where many variables may intervene, and we cannot control for their influence. The contexts of the compared situations are never similar enough to permit considering as null the influence of the environment; accordingly, the researcher will never be able to exclude from her conclusions the contextual variables that she could not keep constant (Dogan and Passy 1990).

By *maximizing the differences* among the cases, we may instead generalize beyond a restricted area, although at the cost of an increase in the number of independent variables to be kept under control. As Przeworski and Teune (1970: 35) have suggested, in the most-different systems design, the choice is in fact to sample *different* countries in order to 'identify those independent variables, observed within systems, that do not violate the assumption of the homogeneity of the total population'. A *most-different systems* design allows for checking if a correlation holds true no matter in which country. This type of analysis focuses on a lower level than the systemic one – most often at the level of individual actors (Przeworski and Teune 1970). This relies on the assumption that individuals will act the same way faced with the same stimulus; hence researchers look for general statements that are universally true.¹⁰ The research strategy that may produce them is based on random samples of the world population, regardless of the social systems to which individuals, groups or subsystems belong. So social science theories should aim not at explaining phenomena as accurately as possible in their specific historical circumstances, but rather at explaining phenomena wherever and whenever they occur (Przeworski and Teune 1970).

In privileging variables referring to individual over systemic variables, Przeworski and Teune (1970: 7) admit that social science based on this kind of assumption, would be 'a priori a-historical'. So research on individual political participation has sampled individuals from different countries with the aim of finding common patterns – for instance, the research of Verba, Nie and Kim (1978) on the impact of social inequalities on political participation in seven countries of the 'first', 'second' and 'third' worlds. Recent research with large numbers of countries searches for a common explanation of individual behaviour beyond historical specificities in different countries (Norris 2002). In their *Dynamics of Contention*, Doug McAdam, Sidney Tarrow and Charles Tilly (2001) apply a most-different strategy design to paired comparisons, not to look for correlation between variables, but to identify common mechanisms. The analysis of 'most-different' countries and historical periods aims to depart from the common foundational tradition by using paired comparisons 'not to maximize resemblance or even to pinpoint differences among whole countries,

but to discover whether similar mechanisms and processes drive changes in divergent periods, places and regimes' (McAdam, Tarrow and Tilly 2001: 82).

Useful for investigating some micro-dynamics of participation, the most-different systems design does have shortcomings. The most ambitious projects, aiming at explaining phenomena worldwide, risk ending up with hypotheses that explain little. As past attempts have indicated, the hope for global theories is likely to be frustrated. For instance, the search to explain development once and for all brought explanations that were too big for accurate empirical work (Verba 1991). Similarly, the relative deprivation theory, based on macro-comparison of large numbers of countries (Gurr 1971), was strongly criticized when in-depth case studies indicated that grievances are always present in a society, but they are mobilizable only when resources are available for the aggrieved groups (Oberschall 1973).

The definition of most-similar and most-different systems designs refers to the units of analysis, but it also has implications for the type of knowledge we seek. Very often, the most-different design is used to obtain generalizable results – that is, to look for historically invariant correlations. The most-similar design often looks to other countries for confirmation of a hypothesis developed in a single country. However, we may have other choices. Differences among dissimilar countries may be used to contrast contexts; or differences among similar countries may aim at specifying hypotheses. In historical sociology, macrohistorical analysis has been, and continues to be, pursued in different ways: looking for single or multiple forms of a phenomenon, or trying to explain one or all cases (Tilly 1984). Among the studies that privilege the search for a single form, *individualizing* comparisons deal with each case as unique, while *universalizing* comparisons identify common properties among all instances of a phenomenon (*ibid.*). Other studies identify multiple forms of a phenomenon, either to explain, in an *encompassing* way, a single instance, or to *find variations* among all cases. As Tilly admits, empirical research usually simultaneously involves different types of comparison, mixing the ideal types; but there is often an implicit or explicit preference for one design or the other. Alternatively, however, most-different systems can be chosen in order to explore deviant or paradigmatic cases. In this sense, good cases are not the most typical, but the most telling, because they help to clarify theoretical problems. In particular, qualitative analysts often select cases where the outcomes of interest occur (positive cases). This strategy, often criticized as selecting on the dependent variable, has been defended as particularly useful for singling out different paths to certain outcomes (Mahoney and Goertz 2006). Additionally, the selection of 'positive cases' can be consid-

ered as a choice oriented to finding necessary causes of some phenomena, especially rare ones such as revolutions. In this sense, it is linked to the definition of the population more than to that of a 'dependent variable'.

Besides the individual preferences of the researcher, the various research designs tend to follow a certain order in the accumulation of knowledge of a certain phenomenon. Skocpol and Somers (1980) suggested a 'research cycle' in which the comparative method is oriented towards: (a) *macro-causal analysis*, in which historical cases are compared for the purpose of making causal inferences about macro-level structures and processes; (b) *parallel demonstration* of theories, applying old theories to new cases; and (c) *contrast of contexts*, looking instead to 'bring out the unique feature of each particular case' included in the research. According to Skocpol and Somers (1980: 196), the three logics are complementary for the accumulation of knowledge:

Parallel comparative history tends to call forth Contrast-oriented arguments when the need develops to set limits to the scope or claims of an overly generalized social-scientific theory. Contrast-oriented comparative history may give rise to Macro-analytic arguments when juxtapositions of historical trajectories begin to suggest testable causal hypotheses. Finally, too, Macro-analytic comparative history can create a demand for the kind of general theorizing that precedes the construction of a Parallel comparative analysis.

Time and history in comparative politics

The definition of the units of analysis and the selection of cases also involve another strategic choice: the use of time. The historical approach is particularly relevant for case-oriented research designs that are by definition context-bound. Long-term processes are particularly important for 'internal' interpretation (what is usually called *verstehen* rather than *erklären*). Variable-oriented analysis is less in need of historical depth, aiming at general knowledge. However, especially in the field of comparative politics, the variable-oriented approach also has a particular interest in the use of time, especially in the form of periodizations that allow for the multiplication of (sub)units of analysis; so the same country in different time periods could be treated as a set of distinct cases. This has received less attention in the social sciences (Bartolini 1993: 131).

References to history do not automatically make for a *diachronic* research design – that is, for a matrix of data that involves collection for at least two points in time. For instance, Theda Skocpol's classic work on revolutions

(1979) refers to Mill's methods of agreement and differences, but without introducing time as a variable. According to Stefano Bartolini (1993: 135), hers is an example of a research design based upon the observation of cross-case *synchronic* variance: 'history is present in the unquestionable "historicity" of events located in the remote past; but there is no time in the scheme, no variance along the temporal dimension in the variables which are considered, and therefore there is no method that is specifically historical'.

Bartolini calls for a use of time through research designs that are *explicitly diachronic* – that is, based on the collection of data at several points in history. Very often, case studies analyse the development of some characteristics, in a single unit, over a certain time span; comparison is then developed between periods. Allowing for the parametrization of many variables and an in-depth historical knowledge, the *cross-time comparison within a single unit* offers many advantages for hypothesis building. Historical analyses of a single country are useful in the development of hypotheses in new fields, insofar as they are able to keep under control – or at least, have knowledge about – a vast range of independent variables that may intervene to 'disturb' the control of a hypothesis. On the basis of an in-depth analysis of a single country, as well as by taking into account the timing of some events, historical case studies may help in developing new hypotheses (see Vennesson, ch. 12).

What is true in a certain country (with a peculiar culture, social structure, model of economic development, and especially configuration of all the different variables), however, is not necessarily true in others. *Cross-national diachronic studies* tend to reach higher levels of generalization and to specify the hypotheses developed in historical case studies that aim at comparing the case of country A at time X with that of country B at times Y and Z. Within a variable-oriented strategy, while increasing the number of countries has the disadvantage of increasing the number of variables to be kept under control, expanding the time span reduces that risk, allowing an in-depth historical knowledge of the cases under analysis (though hampering the assumption of independence between cases).

The use of diachronic research designs is especially common when we expect relevant changes in some dimensions between time t and time $t + n$. In this sense, we treat time as a variable. This is done, for instance, in research on developmental processes involving an interest in steps or thresholds, crisis and transition phases, or trends and sequences. Time is central in grand theories of development, which often assume a teleological scheme 'in which the description of some "primitive" stage enables a number of factors of development to be identified, which then point to some future direction' (Bartolini

1993: 143). In Rokkan's (1970) research, the *timing* of the different processes of nation building and industrialization influenced the evolution of the main social cleavages that survive today. Similarly, Robert Dahl (1971) pointed at the different outcomes related to the precedence given in democratization processes to extending the number of rights to contestation/opposition versus the number of people who enjoy those rights. The timing of the various phases, steps and thresholds involved in the pattern of political modernization is particularly illuminating for understanding democratization in various countries. These analyses tend to share some of what historical sociologist William H. Sewell (1996) calls *teleological temporality*, which explains events through abstract transhistorical processes 'from less to more' (say, urbanization or industrialization), and *experimental temporality*, which compares different historical paths (for example, revolution versus non-revolution, democracy versus non-democracy).

Referring to Skocpol's work, Mahoney has focused attention on a strategy to assess causal inference, which he calls *narrative*, in contrast with the nominal and ordinal approaches. While the nominal strategy (using nominal variables) relies upon Mill's logic of similarities and differences and is therefore deterministic, and the ordinal strategy allows for (probabilistic) analyses of concomitant variation, the narrative strategy addresses phenomena such as revolutions as 'the product of unique, temporally ordered and sequentially unfolding events that occur within cases' (Mahoney 1999: 1164). In the narrative strategy, 'one criterion for judging a causal argument rests with the ability of an analyst to meaningfully assemble specific information concerning the histories of cases into coherent processes' (*ibid.*: 1168). In this sense, it allows one to control, at a disaggregated level, whether the posited causal mechanisms plausibly link explanatory variables with a specific outcome.

While the first two strategies are useful in producing parsimonious theories by eliminating variables, the narrative method scores better on in-depth ideographic knowledge. For instance, if we want to explain why terrorism develops in some countries and not in others, we might proceed by sampling cases in which terrorism was present and others in which it was not, contrasting them on the basis of a nominal logic. If we trust statistics on terrorist events, we might instead measure the presence of terrorism in different countries and rank them in an ordinal way. However, these data are usually static: they allow us to eliminate variables that are not necessary causes or have low or statistically insignificant correlation coefficients, but not to look at the processes through which terrorism develops. This could be done instead through an in-depth narrative of one or a few cases in which terrorism

developed, with attention paid to tracing back the temporal evolution of the various steps of radicalization (della Porta 1995; see also Vennesson, ch. 12).

Sewell (1996) reflects on another way of dealing with time: using the notion of *eventful temporality* for research that recognizes the power of events in history. Events are defined as a ‘relatively rare subclass of happenings that *significantly transform structure*’; and an eventful conception of temporality takes into account the transformation of structures by events. Events produce historical changes mainly by ‘transforming the very cultural categories that shape and constrain human action. Because the causalities that operate in social relations depend at least in part on the contents and relations of cultural categories, events have the power to transform social causality’ (Sewell 1996: 263). Attention to ‘eventful temporality’ reflects the assumption that conjuncture and strategic action make transformative events possible. The conception of an ‘eventful sociology’ implies social processes that ‘are inherently contingent, discontinuous, and open-ended . . . “Structures” are constructed by social action and “society” or “social formation” or “social systems” are continuously shaped and re-shaped by the creativity and stubbornness of their human creators’ (Sewell 1996: 272). Such events as the seizure of the Bastille or, less dramatically, the ‘Battles of Seattle’ (during the contestation of the Millennium Round of the World Trade Organization) not only have a transformative impact on the lives of those who took part in them, but their symbolic relevance spreads to those not directly involved, changing routines and disrupting institutions (see Steinmo, ch. 7, and Keating, ch. 6).

For both diachronic case studies and diachronic cross-national comparisons, *periodization* is a delicate step, since in order to identify temporal variance it is first necessary to define the temporal units which determine such variance (Bartolini 1993). While spatial units are often easy to single out, as they are defined by geopolitical borders, temporal units are not. In fact, temporal variance is assessed by the observations of different time points (separated by more or less regular intervals) or of the general character of *periods* that follow one another. In order to understand how a variable has changed over time, we have to choose significant points in time – that is, to define a time 1, time 2, and so on. Various periodizations may appear as legitimate: what we need is a periodization that is significant according to our theoretical model. It must take into account the main changes in the dependent variable, but it cannot overlook the evolution of the other operative variables.¹¹ Already in a single-country design, the need to take into account variables that vary with a different timing may imply difficult choices between different periodizations. In cross-national designs, we have to deal with the additional

problem of finding comparable periodizations in various countries: similar phases may well develop in different historical periods.

Some projects locate research within a historical perspective, recognizing the value of the long time span, or *longue durée*, with its attention to structures as ‘coherent and fairly fixed series of relationships between realities and social masses’ (Braudel 1980: 31). The field of historical sociology has been particularly sensitive to this issue. Charles Tilly’s research (1986) on the change in the repertoires of collective action in the evolution of the nation-states, covering many centuries of French history, is an example of this type.

Conclusion

We have reviewed two main types of social science comparative analysis and various elaborations of these. One type focuses on large numbers of cases, regularities in behaviour and universal patterns. The other concentrates on context, complexity and difference. Some scholars argue that these follow two different logics, as outlined in Chapter 2 of this volume. Others insist that there is a single logic and that both must follow the same basic rules, albeit using different techniques and materials. The response will be obviously related to the (still vague) conceptualization of ‘logic’: Henry E. Brady and David Collier (2004) have recently underlined in the very title of their edited volume that *social inquiry* must follow *shared standards* while allowing for *diverse tools*. However, the discussion is still open in the social sciences regarding which should be the ‘shared standards’ and how much the presence of ‘diverse tools’ affects the various steps of a research design. In this chapter, we have suggested that many choices in the research design, such as those that refer to conceptualization, case selection and the very conception of explanation and inference, are indeed influenced by the (more or less ontological) preferences for a focus either on cases or on variable-oriented design. This does not, however, have an effect on the standards of empirical research, which must be kept high in both logics.

NOTES

- 1 For a similar conception of inference as the basis of the sociological enterprise, see Goldthorpe (2000, especially ch. 3).
- 2 Similarly, ‘comparative politics’ is used in different ways. Sometimes it refers to the study of countries one by one; sometimes it involves an insistence on the thematic and cross-national

study of institutions and behaviour; while at other times it refers to the use of a variable-oriented approach.

- 3 Scholars also disagree on the capacity of statistical methods to match the experimental design through mathematical manipulation of the data.
- 4 Criticizing John Stuart Mill, Durkheim (1982: 148) states that his 'alleged axiom of a plurality of causes is a negation of the principle of causality'.
- 5 Internal explanations have also been called teleological (understandable on the basis of goals); external ones are causal (or mechanistic).
- 6 In this understanding, 'a case is an entity on which only one basic observation is made, and in which the dependent and independent variables do not change during the period of observation' (Lijphart 1975).
- 7 Critics have countered the accusation that case studies are biased towards verification, stressing instead their importance for the falsification of (non-probabilistic) hypotheses (Rueschemeyer 2003).
- 8 Units are usually considered homogeneous when they respond in similar ways to similar stimuli (Gerring 2001: 176).
- 9 This trend was also helped by statistical techniques that are better suited to the analysis of comparative politics (with small N) because they reduce the impact of deviant cases and allow for simulations that artificially increase the number of cases (Collier 1990: 495).
- 10 'If all relevant factors were known, then the same multivariate statement would yield a deterministic explanation regardless of time and place' (Przeworski and Teune 1970: 7).
- 11 Periodization can be deductive, derived from theoretical assumptions, or inductive, based on empirical observations.