1

Qualitative Comparative Analysis (QCA) as an Approach

Dirk Berg-Schlosser
Gisèle De Meur
Benoît Rihoux
Charles C. Ragin

Goals of This Chapter

After reading this chapter, you should be able to:

- Locate QCA as an approach and grasp its key epistemological foundations
- Understand how and why QCA is "case oriented" and how one should use QCA to engage in a dialogue between cases and theories
- Understand the specific conception of causality conveyed in QCA—*multiple conjunctural causation*—and its practical consequences
- Reflect on the usefulness of QCA to reach a certain level of generalization beyond the observed cases
- Grasp key common features of QCA techniques in terms of formalization, replication, transparency, and different types of uses
- Become accustomed to some key technical terms and use the appropriate, QCA-specific terminology

LOCATING QCA

Epistemological Foundations

To better understand QCA and its various techniques and applications, it is important to locate it both in its historical epistemological context and in its relationship vis-à-vis other methods of social scientific inquiry.[1] In its more recent developments it dates back to systematic comparative procedures as they originated in the natural sciences in the 18th and 19th centuries, as, for example, in Linnaeus' (1753) work in botany or Cuvier's (1812) studies in anatomy.

The logical foundations for this method were laid by Hume (1758) and, in particular, J. S. Mill's (1967 [1843]) "canons." Among these, the "method of agreement" and the "method of difference" are the most important. The first refers to eliminating all similarities but one: "If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree is the cause (or effect) of the given phenomenon" (p. 390). By contrast, the Method of Difference establishes the absence of a common cause or effect, even if all other circumstances are identical:

> If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon. (p. 391)

Both methods thus are concerned with the systematic matching and contrasting of cases in order to establish common causal relationships by eliminating all other possibilities. Both procedures are, however, somewhat extreme in the sense that they attempt to establish a single common cause, or its absence, by controlling all other possibilities and the entire environment.

Mill also devised a combination of the two which he called the "Joint Method of Agreement and Difference" or the "Indirect Method of Difference," which consists of a double application of the Method of Agreement:

> If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance, the circumstance in which alone the two sets of instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon. (p. 396)

This "quasi-experimental" design is, however, as Mill himself stated, less convincing than the pure Method of Difference.

Mill's "canons" imply rather rigid "positivist" assumptions about relationships of cause and effect and the state of valid theory in any given area of research. On the whole, such relatively mechanical and deterministic relationships can be established only rarely even in the "hard" sciences. By themselves, therefore, these methods do not produce any new discoveries unless some truly relevant factors have been included. Similarly, they may not *prove* any causal relationship, because it is most often impossible (in social science at least) to test a clear and complete (preconceived) model of such links and to sufficiently "control" for other factors.[2] They constitute, however, a valuable step toward eliminating irrelevant factors and approximating causal conditions in the "real" world. In this sense they correspond to Popper's (1959) famous principle of "falsification." Or as it was expressed in another classic of this period, Mill's methods are nevertheless

> of undoubted value in the process of attaining truth. For in eliminating false hypotheses they narrow the field in which true ones may be found. And even where these methods may fail to eliminate all irrelevant circumstances, they enable us with some degree of approximation to so establish the *conditions for the occurrence* of a phenomenon, that we can say one hypothesis is logically preferable to its rivals. (Cohen & Nagel, 1934, p. 267; emphasis added).

The various techniques of QCA precisely identify and narrow down such "conditions of occurrence." As will be demonstrated and exemplified in the chapters that follow, these techniques are important tools for reducing the enormous complexity that we routinely confront in the social sciences. As Mill (1967 [1843]) himself put it,

> in politics and history . . . Plurality of Causes exists in almost boundless excess, and effects are, for the most part, inextricably interwoven with one another. To add to the embarrassment, most of the inquiries in political science relate to the production of effects of a most comprehensive description, such as the public wealth, public security, public morality, and the like: results liable to be affected directly or indirectly either in *plus* or in *minus* by nearly every fact which exists, or event which occurs, in human society. (p. 452; emphases in the original)

"Small-N" Research and "Macro-Comparative" Analysis . . . and Beyond

Initially, in the late 1980s and early 1990s, QCA was mostly developed for applications in political science (comparative politics) and historical sociology (e.g., welfare state studies). Thus, quite naturally, QCA has been initially conceived, in those social scientific disciplines, as a "macro-comparative" approach—because the specific subject matter in those disciplines necessitates empirical research at the "macro" level of entire societies, economies, states, or other complex social and cultural formations (Berg-Schlosser & Quenter, 1996).

In the present world, but also if we include relevant historical cases, the maximum number of such cases is of necessity quite limited, as for example the current 200 or so independent countries worldwide, 50 states in the United States, or 27 EU member countries. In fact, for many meaningful macro-level comparisons, the number of cases with useful and comparable data is even more limited—for example, the set of OECD countries, sub-Saharan countries, or a given set of European regions that have received, say, some structural funds for economic development. This is why QCA is still widely seen as a "small-N" approach. Out of this fact arises the characteristic "small-N–many variables" dilemma for this type of research (see, e.g., Lijphart, 1971, 1975; see also Chapter 2 on how to address this problem).

In a more general way, QCA techniques can be located in a two-dimensional matrix listing numbers of variables and numbers of cases in relation to other supplementary or neighboring approaches (Figure 1.1).

The realm of QCA techniques—that is, the "comparative method" in the more narrow sense of the term—thus has to be distinguished, in particular, from the "statistical method," which proceeds on the basis of a large number of cases, drawn on a random basis if possible, and a relatively small number of variables. Both methods have their respective strengths and weaknesses (for extensive discussions, see Brady & Collier, 2004; King, Keohane, & Verba, 1994; see also p. 170), but rather than merely adopting insights from large-scale quantitative inquiries or simply increasing the number of cases as much as possible, QCA follows a different path with several distinct emphases, as will be demonstrated in the next sections.

As QCA techniques and their applications have been developing, this positioning of QCA as a "small-N" and "macro-comparative" approach needs to be nuanced somewhat, in at least two respects. On the one hand, technically speaking, the "small-N" zone is now usually associated with a really low number of cases—say, between 2 cases (this is a "very small-N," but it does enable some form of binary comparison) and around 10 to 15 cases. Beyond this—say, between 10 and 15 and 50 and 100 cases—one finds oneself rather in an "intermediate-N" situation, which is still quite a small number of cases relative to the requirements of most quantitative (read: statistical) techniques. Besides, as shall be discussed later (see p. 174), QCA techniques have been fruitfully applied in "large-N" research designs as well. On the other hand, an increasing number of scholars, in fields such as organizational sociology, management studies, and education studies, among others, have begun to apply QCA techniques at other levels, notably at the "meso" level (the level of organizations, social networks, collective actors, etc.) or even, more recently, at the "micro" level (small groups or individuals) (see p. 173).

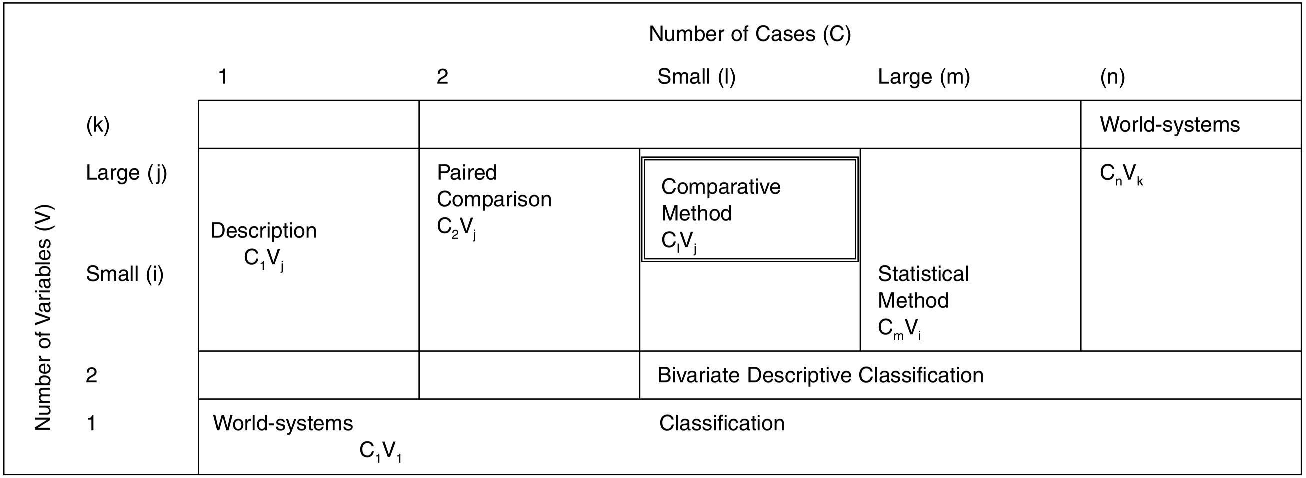

Figure 1.1 Comparative Analysis: A Typology

Source: Adapted from “Die Vergleichende Methode in der Politikwissenschaft,” by F. H. Aarebrot and P. H. Bakka, in *Vergleichende Politikwissenschaft: Ein Einführendes Studienhandbuch* (4th ed.), p. 65, by D. Berg-Schlosser and F. Müller-Rommel (Eds.), 2003, Wiesbaden, Germany: VS-Verlag.

KEY FEATURES AND ASSUMPTIONS OF QCA

In some respects, it can be said that QCA techniques strive to meet advantages of both the "qualitative" (case-oriented) and "quantitative" (variable-oriented) techniques. This was indeed the main ambition expressed when the first technique—initially known as QCA and now referred to as csQCA—was developed in the late 1980s and presented as a "synthetic strategy" to "integrate the best features of the case-oriented approach with the best features of the variable-oriented approach" (Ragin, 1987, p. 84). Indeed, as we explain below, csQCA and the other QCA techniques do combine distinctive strengths of both approaches (Rihoux, 2003, 2006, 2008a, 2008b), but altogether they are more clearly located on the side of "case-oriented" methods (Rihoux & Lobe, 2009). QCA techniques allow the systematic comparison of cases, with the help of formal tools and with a specific conception of cases. This is where configurations come in.

Cases and Theory

Techniques of configurational comparative analysis (CCA) are "case oriented" in the sense that they deal with a limited number of complex cases in a "configurational" way (see p. xix). This means that each individual case is considered as a complex combination of properties, a specific "whole" that should not be lost or obscured in the course of the analysis—this is a *holistic* perspective. The cases dealt with are (or should be) well known rather than anonymous, as, for example, individuals are at the micro level in large-scale survey research. Rather than being a drawback, this can become a considerable advantage that enables the researcher to go back to these cases or consult historians, country experts, and others to clarify further aspects of cases or to check and improve the relevant data.

In the process of configurational comparative analysis, the researcher engages in a dialogue between cases and relevant theories. Indeed, the choice of the variables (conditions and outcome) for the analysis must be theoretically informed. In this sense, there is a deductive aspect to QCA; however, QCA techniques can also be used more inductively, gaining insights from case knowledge in order to identify the key "ingredients" to be considered (Rihoux, 2003, 2006; Rihoux & Lobe, 2009). Further, a key richness of QCA techniques is that they use a formal language (Boolean or set-theoretic; see Chapters 3 to 5) that can be very easily translated into a theoretical discourse (and vice versa); indeed, theoretical discourse is set-theoretic by nature (Ragin, 2000, 2008) and QCA techniques enable a rich dialogue with theory (Befani, Ledermann, & Sager, 2006).

With regards to theories, QCA is best located in the more general area of "medium range" theorizing in social research (Merton, 1968; Mjøset, 2001) and thus has to be distinguished from contemporary "grand" and potentially universal social theories such as those of Habermas, Bourdieu, Luhmann or Giddens, which remain highly speculative and are actually not designed to be empirically testable.[3] In this sense, QCA is, again, more modest and context sensitive in the tradition of "grounded" approaches that are historical, qualitative, and empirically differentiated (Glaser & Strauss, 1967; discussed by Mjøset, 2003).

Seen in this perspective, QCA can lay the groundwork and be extended to even more demanding types of analyses—for example, taking into account the temporal dimension and the various "paths," "critical junctures," and overall dynamics that can be found in systematic comparative historical studies (see, e.g., Pierson, 2004; see also pp. 161–163, 173). Similarly, it can be applied conjointly at several levels of analysis—for example, linking the meso and macro levels when social cleavages and party systems are linked in empirical democratic theory. Furthermore, the specific "conditions of occurrence" identified by QCA can be integrated in the general model of social explanation as explicated in Coleman's "bathtub" (1990) and further elaborated by Esser (1993). In substance, Coleman formalizes the articulation between macro-level and micro-level changes in a given society, where societal change at the macro level or group-level change at the meso level are actually grounded in changes at the level of the constituent individuals (i.e., the micro level). In this way, structure- and actor-related aspects can be combined and brought into a more general historical or "medium-range theorizing" perspective.

In practical terms, theory plays an important role at crucial stages in the application of QCA techniques. First, "upstream," when the model has to be elaborated, theory points at useful conditions to be included in the model and helps to operationalize them (how to measure their intensity, which thresholds to use, etc.). Theories also guide the selection of cases, in the attempt to include both the important or typical cases and the more paradoxical or contrary ones (see also p. 20). QCA indeed tends to give explanations without dismissing "exceptions" or "outliers." These nonconforming cases, on the contrary, often shed a special light on the understanding of specific processes.

Second, during the analysis, theoretical knowledge, as well as a deep knowledge of the empirical field, will help researchers make decisions regarding several practical QCA operations such as the operationalization of variables and the treatment of the so-called contradictory configurations—in short, cases that display the same values on the *condition* variables but lead to different *outcomes* (much more on this p. 44). Theoretical knowledge is also important in a key step of the analysis in QCA: the inclusion of non-observed cases, the so-called logical remainders (see p. 59). Third and not least, "downstream," after the analysis, theory will help the researcher sort the different (otherwise logically equivalent—i.e., equally parsimonious) solutions and justify any reasoned preferences among them.

Causality, Complexity, and Parsimony

QCA techniques allow for "conjunctural causation" across observed cases. This means that different constellations of factors may lead to the same result, as, for example, different "paths" in democratization research (e.g., Berg-Schlosser, 1998; Collier, 1999) or different social forces leading to the emergence of welfare states in Western Europe (e.g., Alber, 1982; Esping-Andersen, 1990). More precisely, QCA develops a conception of causality that leaves room for complexity, referred to as "multiple conjunctural causation."

Box 1.1
"Multiple Conjunctural Causation" in a Nutshell

It is a conception of causality according to which:

1. Most often, it is a combination of causally relevant conditions that generates the outcome (AB → Y).
2. Several different combinations of conditions may produce the same outcome (AB + CD → Y, + indicates a Boolean *or*[4]).
3. Depending on the context, a given outcome may result from a condition when it is present and also when it is absent (AB → Y but also aC → Y). In this example, [A] combined with [B] produces the occurrence[5] of the outcome, but its absence [a] combined with [C] also produces the outcome.

In other words: Different causal "paths"—each path being relevant, in a distinct way—may lead to the same outcome (De Meur & Rihoux, 2002). The term "multiple" refers to the number of paths, while the term "conjunctural" conveys the notion that each path consists of a combination of conditions. Thus multiple conjunctural causation contains the notion of *equifinality,* which simply means that different paths can lead to the same outcome. It should be noted that this runs completely against key assumptions on which mainstream statistical techniques rest—for example, additivity, meaning that a given factor is assumed to have the same incremental effect on the outcome across all cases, regardless of the values of other causally relevant conditions (Schneider & Wagemann, 2007, forthcoming).

Since it views causality as context and conjuncture specific, QCA rejects any form of permanent causality (Ragin, 1987). This is in line with the earlier works of J. S. Mill (see p. 2). Bottom line: By using QCA, the researcher is urged not to specify a single causal model that best fits the data, as one usually does with statistical techniques, but instead to determine the number and character of the different causal models that exist among comparable cases (Ragin, 1987).

It thus goes beyond the (often superficial or misleading) means, correlations, and regressions—computed across all cases at the same time—which average out the respective constellations and ignore specific, distinct patterns and "outliers" (see, e.g., also Berg-Schlosser & Cronqvist, 2005; Berg-Schlosser & Quenter, 1996; Ragin, 2006a).[6] Actually, with QCA, if a given combination of conditions "explains" only one single case, it is not a priori considered as less relevant or less important than another combination of conditions that would account for, say, 10 or 15 cases—because each case matters in most applications of QCA (see pp. 23, 155). In this sense, QCA moves away, quite radically, from simplistic, probabilistic causal reasoning (De Meur & Rihoux, 2002); in its case-orientedness, it is more geared toward diversity (Ragin, 2006a).

Thus QCA broadens the usual frame in the analysis of causality, by relaxing several common assumptions. First, "additivity" is no longer assumed: This means that the idea that each single cause has its own separate, independent impact on the outcome is abandoned and replaced by the assumption that "conjunctural causation" is at work, meaning that several causes can be simultaneously present (or be combined, somehow), constituting a "causal combination," for the outcome to occur. Second, a given causal combination may not be the only route to a specific result; other combinations also may be able to produce it. Third, the uniformity of causal effects is not assumed; on the contrary, a given condition may, combined with different others, sometimes act in favor of the outcome, and sometimes, differently combined, act against it. Fourth, causality is not assumed to be symmetrical—rather, causal asymmetry is assumed, meaning that the presence and the absence of the outcome, respectively, may require different explanations.

Box 1.2
Causal Relations in QCA: Assumptions That Are *Not* Taken Onboard

It is crucial to bear in mind that QCA does *not* take onboard some basic assumptions that lie at the heart of the mainstream statistical approach (and thus underlie most statistical techniques). In QCA:

- Permanent causality is *not* assumed.
- Uniformity of causal effects is *not* assumed.
- Unit homogeneity is *not* assumed.
- Additivity is *not* assumed.
- Causal symmetry is *not* assumed.

Note that other core mainstream statistical assumptions, such as linearity, and so on, are not taken onboard either.

Of course, QCA techniques do not guarantee the final grasp of the "true" causal grounds of a given phenomenon because the issue of causality is a much more complex matter (see, e.g., Abell, 2004; Gerring, 2005; Mahoney, 2004). Besides, the conclusions of any empirical analysis (QCA or any other) are totally dependent on the choice of "ingredients" put under the microscope, including the condition variables as they have been operationalized as well as the selection of cases. Yet, if several competing theories try to explain the same result, QCA techniques will quickly disqualify the theories that are unable to discriminate correctly between cases with and without the outcome under study. This will be indicated by the presence of so-called contradictory configurations (see p. 44).

Among the remaining theories, those that best satisfy the "parsimony principle" (Occam's "razor") will emerge. The parsimony principle, successively reinvented and reinforced through the centuries, can be translated into the commonsense adage: "Why make complicated when one can make simple?" Or, as Einstein put it in his famous dictum: One should express things "as simply as possible, but no simpler." To sum up: QCA techniques strive to achieve some form of "short" (parsimonious) explanation of a certain phenomenon of interest, while still providing appropriate allowance for causal complexity.

Naturally, the search for causal regularities implies the acceptance of the postulate that there are indeed underlying causal regularities in human and social phenomena, even if sketchy (Ragin, 1987, p. 19; Skocpol, 1984, pp. 374–375; Zelditch, 1971). In QCA, as discussed in Box 1.3, two key regularities are framed in terms of necessity ("necessary [combinations of] conditions") and sufficiency ("sufficient [combinations of] conditions"). In fact, any empirical scientific process—even those within the "hard" sciences—is based on this postulate. The opposing postulate, that of an "unstructured chaos" of phenomena, would preclude the search for explanations as well as for meaning.

Box 1.3
Necessity and Sufficiency Back in the Picture[7]

Note that the key concepts of *necessity* and *sufficiency* (as defined on p. xix) are very much in line with the multiple conjunctural view on causation. Indeed, a given path toward an outcome usually consists in a combination of conditions that is *sufficient* (a sufficient *combination* or "intersection" of conditions) to produce that outcome. However, this path is not always *necessary*, as some other alternative paths (with different conditions, at least partly) could very well produce the same outcome. Let us pursue the example used on p. xix and consider three possible conditions leading to the outcome "building a democratic state":

- Condition A: Hold regular competitive elections.
- Condition B: Ensure comprehensive civil liberties.
- Condition C: Ensure independence of political decision-makers vis-à-vis the military leadership.

In this example, there could be two paths leading to the outcome of interest:

- Path 1: the combination of A and B
- Path 2: the combination of A and C

This can be translated as follows:

- Path 1 is the first *sufficient combination of conditions* leading to the outcome.
- Path 2 is the second *sufficient combination of conditions* leading to the outcome.

Neither of these two paths, considered separately, is both sufficient *and* necessary (as there is always an alternate path leading to the outcome). Note, finally, that one condition (A: Hold regular competitive elections) is present in both paths. Hence, we can say that:

- A is a *necessary condition* for the outcome to occur (because it is always present when the outcome occurs).
- However, it is *not* a sufficient condition, because condition A alone does not produce the outcome—it needs to be combined with either condition B or C.

Modest Generalization

Generalization is an important part of any empirical scientific endeavor. The goal of research is not limited to description, as exhaustive as possible, of some corpus of observations. The search for "explicit connections" (Ragin & Rihoux, 2004a) or "specific connections" (see Rihoux, 2008b) takes an important part in the process of understanding. Explicit connections give a formal shape to observed regularities that occur in the data set, and this allows for further investigations, as they are dissected to elaborate an "explanation"—an attempt to describe the mechanism at work. They also give a predictive tool, providing assertions on the behavior of new, not yet observed cases and therefore offering an opportunity to test the model and go a step further.

Without the ambition to generalize, in the search for explanations, research would produce only tautologies and descriptions. This is not to say that more interpretive or "thick" descriptive work is devoid of value—indeed such work can yield very useful insights to grasp phenomena, to understand their deeper mechanisms, to gain an understanding of complex cases (Gerring, 2006; Ragin & Becker, 1992). But it is crucial to recognize the importance of producing new conjectures and to take the risk of confronting them with new data.

The degree of maturity and robustness of a generalization will strongly depend on the quality of the empirical data set constructed by the researcher, and it will generally be a long and hard job to produce it, with many trials and errors, new questionings, and assessments. Contrary to popular myth, those readjustments should not be considered opportunistic manipulations of data; they are necessary steps in their elaboration as researchers increase their substantive and theoretical knowledge—this is why QCA techniques, as shall be demonstrated in further chapters, are iterative by nature.

A good index of the quality of research results could be precisely their ability to withstand refutation when confronted with new cases. In this respect, we should remember that a theory maximizes its robustness when it avoids individualizing explanations—that is, when it avoids providing a specific "explanation" for each specific case (it is then only an accumulation of "descriptions," and not an "explanation"). Only generalization makes it possible to achieve more succinct explanations—such as in the example in the previous section, where condition A is identified as a key regularity (a necessary condition). This again speaks strongly in favor of parsimony.

Yet, the efforts toward generalization that have a reasonable chance of success must stay inside the initial "homogeneity space," within which the empirical data set is contained (see p. 20.). Nothing supports the idea that conditions not included in the analysis would *not* affect the results of the analysis. Hence a well-executed QCA should go beyond plain description and consider "modest generalizations": QCA results may be used in support of "limited historical generalization" (Ragin, 1987, p. 31). More specifically, from a systematic comparison of comparable cases, it is possible to formulate propositions that we can then apply, with appropriate caution, to other similar cases—that is, cases that share a reasonable number of characteristics with those that were the subject of the QCA. Note that this view on generalization is much more modest than statistical inference, which allows very broad generalizations (from a sample of, say, 1,000 respondents to a population of millions of individuals).

Data, Replicability, and Transparency

As mentioned above, QCA techniques require that each case be broken down into a series of features: a certain number of condition variables and an outcome variable. For instance, if we consider athletes as cases, if the outcome is the ability to throw a discus beyond 60 meters, then some conditions could be being tall (versus not tall), being fast (versus slow), being muscular (versus thin), and so on. Then we could measure these attributes for each "case" (athlete): Case 1 could be tall, fast, and muscular; Case 2 not tall, fast, and thin; and so on.

This means that, as with statistical analyses, QCA techniques allow one to develop an *analytical* strategy. However, this segmentation into variables does not affect the perception of each case as a whole. The aim here is to allow for major concerns of both quantitative (defining variables) and qualitative (keeping in touch with the *holistic* perspective) approaches. Having done so, one will be able to compare cases as "whole units," each one of these being defined as a combination of features (i.e., as a *configuration,* as defined p. xix).

In this analytical process, QCA techniques enable us to take into account both "qualitative" and "quantitative" phenomena. When the first QCA technique (csQCA) was developed, Ragin, and several other scholars, used the "qualitative" label to refer to phenomena that vary by kind rather than by degree and also to stress the importance of considering cases as specific and complex configurations, or combinations, of features (Ragin, Berg-Schlosser, & De Meur, 1996, p. 749). Incidentally, we should mention here that it is perfectly possible to work with "subjective" or "qualitative" data using QCA. The only practical requirement is to be able to transform these data into categories or numbers. For instance, if our cases are political parties after a given election and a condition is defined as "perception of electoral defeat by the party activists," the nature of the data could be very subjective (i.e., based on discussions with party activists, assessment of the "tone" in the party press, a feeling about the atmosphere in the post-election party congress, using participant observation methodology, etc.). For any given party, one could still allocate a numerical score of "1" ("yes, or mostly yes") or "0" ("no, or mostly no") on this condition.

However, csQCA, as well as the other QCA techniques, is also able to consider phenomena that are "more" or "less" similar—that is, to express a degree of (dis)similarity (the differences therefore being of a quantitative nature)—as well as phenomena that differ by their nature (the differences being in this case qualitative). For instance, a "wealth versus poverty" variable could be based on some fine-grained data on yearly income of individuals—this would then be a typically "quantitative" measurement of wealth versus poverty. Alternatively, one could consider—for instance, following some official UN or OECD criteria—that the crucial analytical difference is to consider whether a given individual is "poor" or "not poor." Usually, this is done by specifying an income threshold (say, below 1,000 EUR/month, in a country such as Germany) below which one is considered "poor." This is exactly what will be done in the basic, dichotomous csQCA: switching, in the course of the analysis, from quantitative (numerical) to dichotomous (still numerical, but tapping a more qualitative distinction) variables, using substantive knowledge to guide the dichotomization. By dichotomizing the originally fine-grained (quantitative) data, our aim is to identify what can be considered a more fundamental, more basic distinction, of a more qualitative nature.

In this process, similarly to the quantitative (mainly statistical) approach, QCA offers tools that are both formalized and replicable. QCA techniques are formalized in the sense that they are based on a particular language (Boolean algebra and set theory) whose rules and well-defined solutions formalize and translate the rules of logic (formal operations and algorithms; see Chapters 3 to 5). Because these formal rules are fixed and stable, they allow replicability. Simply stated, this means that another researcher using the same data set and selecting the same options will obtain the same results (King et al., 1994, p. 26). This is a major asset of QCA techniques compared to many ad hoc or less formalized qualitative techniques. From a certain perspective, one could say that replicability provides the "scientific" character of the approach, in the sense that it eliminates vagueness and interpretation in the application of techniques (mathematics, for example, is universal as far as technique is concerned).

Another advantage of QCA techniques is their transparency. They demand that the researcher, at several points in the analysis, acts with transparency in his or her choices—selecting variables, processing them, choosing tools for the analysis, intervening during the analysis, and so on. During this process, the researcher regularly refers back to the cases with all their richness and specificity. This back-and-forth "dialogue with the cases," combined with the transparency of choices, is unquestionably a virtue of QCA techniques. What also makes this transparency possible for QCA techniques is that the formal language used by the software takes its inspiration from principles used in everyday life and, for this reason, can be more easily understood by non-specialists.

With most statistical tools, the researcher enters the data and the software finds the "solution." In contrast, QCA opens the "black box" of formalized analysis, by demanding from researchers not only that they make choices but also that they account for them. Using QCA, researchers must be engaged in the analytic process, for it is not mechanical or "push-button." For sure, such requirements should also apply to statistical work. The difference, with QCA, is that the user is more active, gets a better grip on the "mechanics" of the formal operations,[8] makes more decisions in the course of the analysis, and follows an iterative logic, with frequent "returns to the cases." Researchers may feel uncomfortable with this, but this lack of comfort is beneficial, because it compels them to use critical thought during the analysis and opens up the research to others for confirmation or falsification (Popper, 1963).

FIVE TYPES OF USES OF QCA TECHNIQUES

QCA techniques may be exploited in at least five different ways. According to their specific needs, researchers can use different features of QCA. Here, we consider only the three versions of QCA (csQCA, mvQCA, fsQCA). MSDO/MDSO is a technique geared toward one specific type of usage.

Box 1.4
Five Types of Uses of QCA Techniques

1. Summarizing data
2. Checking coherence of data
3. Checking hypotheses or existing theories
4. Quick test of conjectures
5. Developing new theoretical arguments

First, QCA techniques may be used in a straightforward manner, simply to *summarize data*, to display them in a more compact way, and to describe more synthetically the relevant empirical universe. This is thus a purely descriptive use of QCA. More specifically, this is done by means of using the software to generate a synthetic table that shows, in a straightforward way, how some cases cluster together—the so-called truth table (see Chapters 3 to 5). In this way, the researcher will be able to bring to light similarities between cases that may, at first sight, seem quite different. QCA is therefore an excellent tool for data exploration.

Second, the researcher may take advantage of QCA to *check the coherence* of the data. During the analysis, one often detects contradictory configurations—that is, cases that are identical with respects to causal conditions, but different in outcome (see p. 44). Contradictions are plainly displayed in the truth table produced by the software. Detecting contradictions, however, does not necessarily mean that researchers have failed. On the contrary, contradictions will tell them something about the cases they are studying. By seeking a solution to these contradictions, the researcher will both get a more thorough knowledge of relevant cases (through his or her "dialogue with the cases") and develop a more coherent body of evidence (see p. 48).

Third, QCA may be used to *test hypotheses* or *existing theories.* More precisely, it enables us to corroborate or falsify these hypotheses or theories. When using QCA in this way, the researcher aims at operationalizing some theory or hypothesis, as explicitly as possible, by defining a series of conditions that should yield a particular outcome. QCA is a powerful tool for this kind of application because it allows theory testing or hypothesis testing that is both systematic and empirical. When the researcher discovers, through QCA, a large number of contradictory configurations, it may enable him or her to falsify the theory or hypothesis (a very important achievement from a Popperian perspective; see above and p. 50). Furthermore, QCA allows us to refine the hypothesis testing process, by taking into account the actual number of cases related to falsification or corroboration.

A fourth use, close to the former one, is the *quick test of any conjecture* formulated by the researcher him- or herself—that is, without testing a preexisting theory or model as a whole. This is another way of using QCA for data exploration. The researcher specifies an expression (a formula) reflecting a specific conjecture, for example, to test an ad hoc theory or part of a theory. This yields a truth table, which allows the researcher to check whether his or her conjecture was accurate—whether it is confirmed or falsified by the set of cases under analysis.

Last but not least, QCA also may be used in the process of *developing new theoretical arguments* in the form of hypotheses. By obtaining a truth table free of contradictions and then conducting QCA, the researcher obtains a reduced expression (called a "minimal formula"). This may then be interpreted through a "dialogue with the cases" to yield new theoretical arguments. In this way, QCA can be used in a more grounded manner.

One specific technique is especially relevant for this fifth use of QCA. It consists of revising by hand the reduced expression (results of truth table analysis) generated by the software. More specifically, the researcher treats these results as a conventional algebraic expression (Boolean sums of products) and factors it to highlight shared conditions, or to rearrange it in other algebraically acceptable ways so that it speaks as directly as possible to theoretical and substantive interests (see, for example, p. 58). As such, however, QCA does not yield new theories. What it may do, once performed, is to help the researcher generate some new insights, which may then be taken as a basis for a further theoretical development or for reexamination of existing theories. Only by returning to empirical cases will it be possible to evaluate whether it makes sense to highlight a particular condition.

Key Points

- The logical foundations of QCA date back to previous work by J. S. Mill, and in particular to Mill's "canons."
- QCA was initially geared toward "small-N" designs (few cases) and macro-level cases; however, the niche for QCA applications has broadened to "intermediate-N" and "larger-N" designs and also to more meso-level or micro-level cases.
- QCA techniques feature strengths of both qualitative and quantitative techniques but are still located closer to "case-oriented" techniques.
- QCA conveys a particular conception of causality: "multiple conjunctural causation." It is a nonlinear, non-additive, non-probabilistic conception that rejects any form of permanent causality and that stresses equifinality (different paths can lead to the same outcome), complex combinations of conditions, and diversity.
- It is possible to produce generalizations with QCA; however, these are only "modest" generalizations.
- QCA techniques are analytical, transparent, and replicable and can process various sorts of data, from more quantitative (numerical) to more qualitative or subjective; they require an ongoing dialogue between case-oriented knowledge and theoretical knowledge.
- QCA techniques can be exploited in several different ways.

Key Complementary Readings

Goertz (2006b), Mill (1843), Popper (1963), Ragin (1987, 2000, 2003, 2006a).

NOTES

1. For a complementary view, see Caramani (2008).
2. And also because of the "limited diversity" problem (see p. 27).
3. A note of nuance: One may attempt to empirically and systematically test some fragments of such grand theories—see for instance how Andersen (2005) succeeds in testing Luhmann's systems theory. Part of Bourdieu's own work has also consisted in deriving concrete and testable empirical propositions from his grand theoretical framework—but alas only very few of his contemporary "believers" have followed that path.
4. See p. 34 and Box 3.1: Main Conventions and Operations of Boolean Algebra.
5. Ibid.
6. It is fair to say that the broad majority of mainstream statistical work does little to deal with causal complexity. However, some suggestions have been made in the

statistical literature as to how to deal with issues such as asymmetrical causation, necessity, sufficiency, and so on. See Schneider and Wagemann, 2007, forthcoming.

7. Example suggested by Lasse Cronqvist, further elaborated by Benoît Rihoux. For more extensive discussions, see Goertz, 2006b, Schneider and Wagemann, 2007, forthcoming; Wagemann & Schneider, 2007.

8. These logical operations, as well as all their underlying assumptions, are much simpler than statistical operations. Very few users are really able to understand the deep mechanisms and assumptions at work behind the statistical operations.