Chapter 24

Qualitative Comparative Analysis

Patrick A. Mello

Preprint version of a forthcoming chapter

Citation:

Mello, Patrick A. 2022. Qualitative Comparative Analysis. In: *Routledge Handbook of Foreign Policy Analysis Methods,* ed. by Patrick A. Mello and Falk Ostermann. Abingdon: Routledge (forthcoming).

Introduction

Foreign Policy Analysis (FPA) often seeks to account for outcomes that are rooted at different levels of analysis, bringing together a variety of potential causes. For instance, how to explain Russian President Putin's decision to attack Ukraine on February 24, 2022? Arguably, despite some warning signs, this egregious behavior was unexpected by international observers. Some have suggested to investigate leaders' personal characteristics to understand such policy decisions. In the case of Vladimir Putin, several psychological reasons may explain why he has become more risk-prone and intoxicated by power (Kaarbo 2022). Others point to a sweeping transformation of Putin's inner circle of advisers (Treisman 2022). What is clear, however, is that a comprehensive account requires the consideration of various factors and that there is no single explanatory variable to account for the observed outcome.

Qualitative Comparative Analysis (QCA) is a method that was designed to account for *causal complexity*, which entails the existence of multiple paths towards an outcome (*equifinality*), the combination of factors to jointly bring about an outcome (*conjunctural causation*), and asymmetry between the explanation of an outcome and its negation (*causal asymmetry*). As a comparative approach, QCA requires a certain minimum number of cases, but this can be as small as 12 or 15 cases. This means that QCA can also be applied in settings where large datasets are absent, and researchers are working primarily with qualitative data they have gathered themselves, such as from interviews or focus groups (Pagliarin, La Mendola, and Vis 2022; see also Deschaux-Dutard in this volume). This makes QCA an amenable method for FPA, where researchers are often interested in phenomena that are relatively rare (e.g., wars, military coups, economic sanctions, foreign policy failures) but where the aim is to conduct comparisons and to achieve (moderate) generalization across cases. Moreover, QCA can also work in medium to large-*N* settings, especially when researchers draw on pre-existing datasets or when the method is utilized in larger, collaborative projects.

In this chapter, I start out with a review of QCA applications in the field of FPA. Since QCA can still be considered an emerging method, at least when compared to more prevalent statistical or case-study approaches, the number of empirical applications remains limited. This is even though the method's usage has been growing dynamically, especially when taking a broader conception of FPA into account. Hence, I will investigate IR applications more broadly, also because the dividing line between the fields of IR and FPA can be rather fuzzy. The second part of the chapter introduces QCA's methodological assumptions and key terminology. To be sure, this will not suffice to replace a full textbook introduction to the method (see Schneider and Wagemann 2012; Mello 2021; Rihoux and Ragin 2009), especially for those who are entirely new to the method, but the presentation in this chapter should serve to provide an understanding of how QCA works and how its results should be interpreted. Hence the chapter should also be of use for those who simply want to understand how to read the results reported in QCA studies. The chapter proceeds with an empirical illustration to show how QCA works in practice. I close with a discussion of the method's strengths and limitations, together with recommendations on how to avoid frequently encountered mistakes and a brief look into prospective applications in FPA.

Qualitative Comparative Analysis in FPA

While QCA has been around since Charles Ragin published *The Comparative Method* (Ragin 1987) and the method can be considered firmly established in political science, sociology, and many other fields of the social sciences – the number of studies in International Relations (IR) and FPA is considerably smaller. In their recent review, Ide and Mello (2022) identify a total of 43 empirical applications of QCA published in IR journals that are indexed on the Web of Science. For FPA, the number will still be considerably smaller.[1] That said, the publication of QCA textbooks and the broader dissemination of the method, also through summer schools, training opportunities, and methods curricula at graduate schools, have led to a rapid increase in applications in recent years.

To be sure, a central question is how narrowly one defines the field of FPA. If one adopts a broad conception, then any "goal-oriented behavior of state actors across borders" could be considered as such (Brummer and Oppermann 2019, 1, own translation). A narrower definition places human agents and their decisions at the center of FPA (Hudson and Day 2019, 3). If we follow the first definition, then any study that focuses on foreign policy outcomes could be within the purview of FPA, whereas the latter definition would limit that scope to those studies that analyze foreign policy decision-making processes and individual or group decisionmakers. In this chapter, I follow the former conception of FPA, if only to allow for a broader inclusion of studies in my literature review. Arguably, some of the authors of these studies may not self-identify their articles as FPA studies, but for the purposes of this chapter they are considered as such because they engage with substantive matters that fall within the realm of the discipline.

A first cut to distinguish QCA applications in FPA can be achieved by looking at the level of analysis. As in other fields of the social sciences, a majority of FPA studies using QCA are situated at the country level, comparing across a medium number of states. Examples include democracies' involvement and non-involvement in the Iraq War of 2003 (Mello 2012, 2014), countries' participation in the multilateral coalition against the "Islamic State" (Haesebrouck 2018; Mello 2022), NATO burden sharing in Libya (Haesebrouck 2017b), the role of junior partners in coalition warfare (Schmitt 2018), the political contestation of military missions (Haesebrouck and van Immerseel 2020), the implementation of sanctions against “Arab Spring” countries in the Middle East and North Africa (Boogaerts 2018; Boogaerts and Drieskens 2020), the occurrence of unintended consequences of UN sanctions in targeted states (Meissner and Mello 2022), and the allocation of the foreign ministry to junior partners in governing coalitions of parliamentary democracies (Oppermann and Brummer 2020). Other studies operate at different levels, for instance examining the influence of ethnic identity groups on U.S. foreign policy (Rubenzer 2008), the inclusion of human rights in territorial peace agreements (Caspersen 2019), the conditions under which democratic leaders’ opted for defection from the multilateral Iraq War coalition (Mello 2020), the foreign policy behavior of Brazil in various international crises (de Sá Guimarães and de Almeida 2017), agency slack within UN organizations (Heldt et al. 2022), military intervention in Africa (Kisangani and Pickering 2022), conceptions of international order, as expressed in Australian and Chinese policy documents (van Nieuwenhuizen 2019), and even international arbitration under Hellenistic rulers in ancient times (Grynaviski and Hsieh 2015).

With its affinity towards medium-N comparisons of 20-30 cases, it is no surprise that QCA has frequently been applied on EU and NATO member states. In substantive terms, it is apparent that many QCA studies in FPA focus on matters of international security, including decisions on military interventions, burden sharing within alliances and among coalition partners, international sanctions in response to human rights violations, and decision-making on arms control agreements. On the latter, for example, Böller and Müller (2018) examine inter-branch dynamics between the U.S. Congress and the President in relation to military interventions and Böller (2021) investigates U.S. decisions on international arms control treaties. In sum, a review of empirical applications of QCA shows that the method has gained ground in FPA research. While QCA is still in its infancy in FPA, recent years indicate a dynamic growth in the number publications, even more so if one takes a look at IR studies at-large (Ide and Mello 2022). One particular advantage of QCA, besides its ability to account for causal complexity, is that the method can be flexibly applied and tailored towards the specific needs of a research project. It is up to researchers to define their “cases” and to craft a comparative research design on that basis. That said, one major limitation is the availability of comparative data. I will return to this point in the final section of this chapter.

Methodological assumptions and key terminology

The comparative approach of QCA rests on set theory and Boolean logic. This means that relationships are framed in the language of necessary and sufficient conditions. A *sufficient condition* is a condition that always leads to the outcome whereas a *necessary condition* is a condition without which the outcome does not occur. Lately, set theory has made strong inroads into the social sciences, serving as the foundation for new frameworks of analysis (Mahoney 2021; Goertz and Mahoney 2012; Schneider and Wagemann 2012). Beyond this, QCA involves certain methodological assumptions and specific terminology, which will be introduced in this section. I should highlight that, for reasons of space, I will set aside debates about *ontology* – for instance, whether and under which conditions QCA can be a suitable method for a critical realist framework (e.g., Gerrits and Verweij 2014; Rutten 2020). I will also just touch upon recent exchanges about set theory and different theories of causation (Haesebrouck and Thomann 2021, Ch. 4; Haesebrouck 2022; Mello 2021).

Causal complexity

A distinct strength of QCA is its ability to account for *causal complexity*. This overarching concept entails three components: conjunctural causation, equifinality, and causal asymmetry. *Conjunctural causation* means that two (or more) conditions may jointly bring about an outcome, but not individually in the absence of the other condition(s). For example, it may require both a leader with certain political preferences *and* a window of opportunity to implement an otherwise contentious policy. *Equifinality* means that the same outcome may be reached through multiple, different pathways. For instance, while many EU member states eventually decided to support Ukraine with arms deliveries in its self-defense against Russian aggression, the foreign policy decision-making processes through which this outcome was reached were markedly different across the EU. Finally, *causal asymmetry* indicates that the explanation for an outcome can usually not be mirrored to account for the non-outcome but requires a separate analysis.

The notion of causal complexity resonates with many phenomena that are of interest to FPA research. It also departs from commonly held assumptions, especially in quantitative research.[2] As such, conjunctural causation challenges the utility of exploring net effects of individual variables without examining their interaction, while equifinality calls into question the assumption that similar outcomes must be rooted in similar causes. Finally, causal asymmetry highlights the fact that symmetrical causation is an assumption that may also prove false. This departs from linearity assumptions frequently held in statistical analyses (Mello 2021, 69-70).

Calibration

QCA works with calibrated data. *Calibration* means that the researcher defines a target set (e.g. "aggressive foreign policy stance") and determines empirical anchors in the raw data (which may be quantitative or qualitative) to reflect a certain degree of membership in the target set. To take the example of an aggressive foreign policy stance, we may say that certain government statements or actions are considered to indicate set membership. Hence, these would be coded accordingly. This also allows for degrees of membership, which would translate into a *fuzzy set* that can take on any values from 0 to 1 (as opposed to *crisp sets* that are binary in nature). It is considered good practice to use adjectives for set labels because this indicates the qualitative direction (e.g. "*high* unemployment", "*low* aggressiveness", and the like). This reflects the fact that sets are always directed towards a qualitative state, which also distinguishes calibrated data from mere numerical scales where –lacking additional information– it is not possible to say whether a score is high or low in the given context.

Calibration can be conducted manually by assigning scores in a spreadsheet or – the more viable option for larger data sets – through an R function (for more on *software*, see below). In both cases, the researcher needs to decide about calibration criteria to be applied consistently across all the included cases. It is important to realize that calibration should not be exercised mechanically, as in transforming the raw data into calibrated data merely based on descriptive statistics like the mean, minimum, or maximum scores on a given condition. Instead, researchers should anchor their calibration decisions in substantive considerations and external criteria derived from their research area. For example, to distinguish between being poor and not poor, we may refer to the UN poverty line as an external criterion. The number could then be applied to the data as the cross-over (fuzzy score of 0.5) between being rather inside than outside the set *poor country*. Below, I describe examples in more detail, which may help to grasp how calibration works in practice. [3]

Consistency and coverage

How to assess whether an individual condition or a combination of conditions are necessary and/or sufficient for an outcome? Similar to statistical analyses (see Oktay in this volume), QCA uses metrics to describe the fit between the empirical data and certain relationships. In the case of QCA, this is about set-theoretic relationships of necessity and/or sufficiency. The primary measure of fit, *consistency* describes the extent to which an empirical relationship between a condition or a combination of conditions and the outcome approximates set-theoretic necessity and/or sufficiency. The secondary measure of fit, *coverage* describes the empirical importance or the relevance of a condition or combination of conditions (Ragin 2008; Mello 2021). These measures of fit are part of the computation the software conducts. They allow an easy assessment of whether the data fits a set-theoretic relationship. As rules of thumb, truth table rows should have a minimum of 0.75 consistency to be included in the minimization procedure (more on this below). For necessary conditions, the conventional benchmark is 0.90 consistency. As for coverage, there is no firm rule, but the aim should always be to account for at least more than half the observations (the more the better). This means that coverage should usually be in the range of 0.60 and above.

Truth tables

The core of QCA is entailed in the *truth table analysis*, which is about identifying combinations of conditions (*configurations*) that are sufficient for the outcome of interest. The *truth table* displays all logically possible configurations for a given research design and the number of conditions the researcher chose to include. The size of the truth table is calculated as 2^k where k represents the included number of conditions. For instance, a study with four conditions would yield a truth table that comprises 16 rows (2^4), whereas five conditions would result in 32 truth table rows. Apart from displaying the different configurations, the truth table also links these to the included empirical cases and indicates whether and how consistently the rows are associated with the outcome. If all cases that share certain characteristics (indicated by their assignment to the same truth table row) also show the outcome, then this row would yield a high consistency (consistency can range from 0 to 1). Conversely, low consistency indicates that a considerable number of cases in the respective row do not show the outcome.[4] By convention, the minimum threshold for truth table rows to be considered "sufficient enough" to be included in the ensuing minimization is 0.75 consistency (Mello 2021).

Truth tables also provide information about configurations that are not filled with empirical cases (the phenomenon of *limited diversity*). These "empty rows" are *logical remainders* that can be incorporated as counterfactuals during the set-theoretic analysis. For example, in a study on U.S. presidents all combinations involving a female president would be empty because historically there has not (yet) been a case with this configuration. However, researchers could still explore the potential of such a configuration and what outcome it would be associated with. And even if the aims were more limited or a researcher would not want to engage in counterfactual reasoning, it is valuable to know which configurations are filled with empirical cases and which ones are not. The truth table makes this information immediately visible.

Boolean minimization

Once a truth table has been constructed, the researcher needs to determine a consistency threshold for truth table rows to be included in the *Boolean minimization* that constitutes the next step in the QCA analysis. As mentioned, the conventional threshold is a minimum of 0.75 consistency, below which rows should not be included. Higher thresholds are quite common, but this ultimately depends on the nature of the empirical data. Rows that meet this threshold will undergo a software-based comparison based on the rules of Boolean algebra. In simple terms, this means that all the included truth table configurations are compared to each other, and redundant elements are eliminated to reach a more concise solution. For example, in a simple research design with the conditions A and B we may come across two configurations that turn out to be sufficient for an outcome Y. The first comprises the presence of both conditions, while the second entails the absence of the second condition. Using Boolean operators (AND = ·, OR = +, NOT = ~), we can express this as such:

(1) $A \cdot B + A \cdot {\sim}B \longrightarrow Y$

In verbal terms, this means that *the presence of both A and B or the presence of A and the absence of B are sufficient for the outcome Y*. Hence, we can say that once A is present it does not matter whether it combines with the presence or with the absence of B – because both configurations are sufficient for the outcome. Therefore, we can apply the Boolean minimization rule and delete redundant elements B and ~B, respectively. This leaves us with:

(2) $A \longrightarrow Y$

In practice, Boolean minimization is more complicated because it will typically involve numerous conditions and configurations. There is also a consecutive step that allows for further minimization using *prime implicants* (see Mello 2021, Ch. 7). However, for our purposes, it must suffice that the truth table analysis applies the rules of Boolean algebra to logically simplify the sufficient configurations identified in the truth table. It is also important to underline that Boolean minimization merely deletes redundant information. If the first statement is true, then –by logical necessity– the second statement will also be true. But the second statement is more parsimonious.

Solution terms

QCA results are summarized in *solution terms*. This can take the form of statement (1) or (2) shown above but often solution terms are more complex, involving various combination of conditions (conjunctions) that are linked by the Boolean operator OR. What is more, QCA entails three different types of solution terms, depending on the treatment of the logical remainders. The *conservative solution* (also called *complex solution*) works solely with the empirical rows. The *parsimonious solution* allows the algorithm to include logical remainders in the minimization, provided their use yields a simpler result. As the name implies, this solution is often more parsimonious because more Boolean comparisons can be made. The downside is that the parsimonious solution may rest on *implausible counterfactuals* (Mello 2021, 205). For example, we may know that a certain combination of our conditions is impossible in the social world. Hence, it would be erroneous to assume that the respective logical remainders would show a certain outcome if it existed (because it simply cannot exist and therefore cannot show the outcome). To be fair, such problems appear less often than one may think at first glance. But it is crucial to be aware of the potentiality of implausible counterfactuals and to assess this when working with the parsimonious solution. Finally, the *intermediate solution* enables the researcher to customize the treatment of logical remainders, to avoid working with implausible counterfactuals. For instance, the intermediate solution allows a researcher to exclude a certain logical remainder row in the truth table because it may be deemed implausible as counterfactual. Likewise, researchers can introduce “directional expectations” about which qualitative states of their conditions are assumed to be associated with the outcome.

Software

The analytical part of QCA is *software*-based, with a variety of different solutions on offer.[5] Depending on the subfield, the most popular software may still be fs/QCA (Ragin and Davey 2017) and TOSMANA (Cronqvist 2019), both of which are “click-and-point” programs that are easy to navigate also for beginners. While these two programs remain in use and may be a good entry point to explore how QCA works, a host of advanced functions for QCA are available within the R environment, where the package “QCA” (Duşa 2019) covers the analytical core of QCA and the complementary “SetMethods” package provides additional functions (Oana and Schneider 2018; see also Oana, Schneider, and Thomann 2021). Because these packages run under R, users can benefit from the vast opportunities that R provides. This means that researchers can easily combine QCA with statistical tests or conduct visualizations of their results with other packages such as the versatile "ggplot2" (Wickham 2016). That said, the entry barrier to conducting QCA in R can be quite steep, especially if one has not worked with R before.[6]

Empirical illustration

This section provides an illustration of how QCA works in practice, drawing on a published study on military coalition defection during the Iraq War (Mello 2020). In that article, I sought to answer the question why some of the democratic coalition partners of the United States decided at various points between March 2003 and December 2008 to withdraw their country's forces, whereas others stayed until the end of the multinational mission. Overall, 29 democracies were militarily involved in the Iraq War, and these comprised 51 leaders who were in charge during their country's deployment. Out of these, 18 leaders announced their country's unilateral withdrawal before the end of the mission (Mello 2020, 56-58), while 33 leaders continued the military involvement throughout their tenure.

In my research design, I built an integrative theoretical framework that included factors that had been prominently suggested in the literature, such as arguments about electoral incentives and leadership changes (e.g., Tago 2009; Pilster, Böhmelt, and Tago 2013). Prior studies suggested that changes in an unpopular foreign policy, such as the Iraq War deployment, became more likely when a new leader came to power or when an incumbent faced an election. Apart from testing these general expectations on the specific case of the Iraq War, I also examined civilian and military casualties, the relative size of a country's military deployment (in relation to its military capabilities), and the political partisanship of the respective government (left-right placement on a one-dimensional scale). While all these factors were expected to be of relevance for democratic leaders' withdrawal decisions, I expected *combinations* of these conditions to be jointly sufficient for the outcome. For instance, one hypothesis was that "A change in the political leadership combined with leftist partisanship is a sufficient condition for early withdrawal from coalition operations" (Mello 2020, 50). This reflects the assumption that a substantive change in a country's foreign policy requires not just a window of opportunity (a newly elected leader) but also preferences that resonate with it (political partisanship). Likewise, it was also expected that multiple paths could lead towards the outcome, for example when leaders faced casualties and/or had made a small coalition commitment to begin with. These theoretical expectations resonated with QCA's emphasis on *causal complexity*, as discussed above.

The data collection for this study involved both qualitative and quantitative sources. In general, QCA can be used with all types of data, as long as a complete assignment of scores to cases is feasible, in the sense that there is no missing data.[7] The coding of the outcome early withdrawal was based on *qualitative* data, such as leaders' press statements, parliamentary speeches, and news coverage. This involved a labor-intensive cross-checking and verification of sources for the 51 leaders, to identify whether any announcements had been made, on which date these had been issued, and whether the statements had actually been official declarations that initiated the country's withdrawal from Iraq. For the explanatory conditions, I drew on *quantitative* information in existing databases, such as the Chapel Hill Expert Survey, the Rand Database of Worldwide Terrorism Incidents, and data from the Iraq Casualties Project, among other data sources (Mello 2020, 58-59).

Table 24.1 displays the uncalibrated *raw data* on a selection of 14 out of 51 leaders (abbreviated for reasons of space). The table indicates that some leaders made withdrawal announcements that did not qualify as early withdrawal announcements because the statements were either indeterminate or involved partial withdrawals (dates set in italics). The table further includes information on the date of the next election. Decisions made within two months of the next election were coded as instances of an *upcoming election* (as a robustness test, I also checked for a 6-month period, which lead to substantively similar results). The left-right indicator of a governments' placement in political space runs from 0 (extreme left) to 10 (extreme right). Data on the size of a country's deployment and military spending was used to inform the calculation of relative indicators of *fatalities* and *commitment*. Each of these had to be linked to the tenure of the respective leader (if there had been a withdrawal announcement, then the fatalities also had to have happened before that statement was issued).

Table 24.1: Democratic leaders, conditions, and outcome (raw data)

Country/leader (selection, 14 out of 51)	Withdrawal announcement	Next election	Left-right	Fatalities		Commitment	Deployment		Military expenditure	
				per depl.	Nomina	Depl./ME	Relative	Troops	Relative	Mil. USD
Australia (J.W. Howard)	*17-Apr-03*	9-Oct-04	7.46	0.19%	2	0.85	2.35	1,048	2.78	9,927
Australia (K. Rudd)	30-Nov-07	21-Aug-10	3.90	0.00%	0	0.62	2.98	1,330	4.81	17,185
Bulgaria (S. Sakskoburggotski)	31-Mar-05	25-Jun-05	5.62	2.06%	10	6.81	1.09	485	0.16	569
Czech Republic (M. Topolánek)	*8-Oct-07*	29-May-10	7.40	0.00%	0	0.36	0.25	110	0.69	2,450
Denmark (A.F. Rasmussen)	21-Feb-07	13-Nov-07	7.28	0.83%	4	1.16	1.08	480	0.93	3,333
Estonia (A. Ansip)	-	6-Mar-11	5.34	0.00%	0	0.75	0.09	40	0.12	444
Italy (S. Berlusconi)	*19-Jan-06*	9-Apr-06	7.33	1.00%	24	0.64	5.38	2,400	8.47	30,242
Italy (R. Prodi)	7-Jun-06	13-Apr-08	3.49	0.00%	0	0.31	2.91	1,300	9.35	33,408
Netherlands (J.P. Balkenende)	21-Oct-04	22-Nov-06	6.29	0.23%	3	1.24	2.89	1,288	2.34	8,356
Poland (D. Tusk)	*23-Nov-07*	9-Oct-11	5.94	0.00%	0	0.84	2.02	900	2.4	8,589
Portugal (J.M.D. Barroso)	-	20-Feb-05	6.50	0.00%	0	0.31	0.27	120	0.87	3,110
Slovakia (R. Fico)	18-Oct-06	10-Mar-12	4.51	1.18%	1	0.73	0.19	85	0.26	911
Spain (J.M. Aznar)	-	14-Mar-04	7.60	0.83%	10	0.75	2.71	1,208	3.61	12,881
Spain (J.L.R. Zapatero)	19-Apr-04	12-Apr-08	3.70	0.08%	1	0.68	2.91	1,300	4.27	15,262

Note: Italics indicate partial or indeterminate withdrawal announcements. Left-Right scores range from 0 (extreme left) to 10 (extreme right).

As mentioned above, QCA requires calibrated data. This meant I had to transform the raw data listed in Table 24.1 into calibrated set-membership values. This can be done in a qualitative fashion, where the researcher defines calibration criteria for the target concepts and then assigns the respective scores by hand.[8] More commonly, the "direct method of calibration" is used, which is a software-run procedure that transforms numerical raw data into decimal-score fuzzy-set values based on a logarithmic function (Ragin 2008). For example, the measure of governments' political partisanship is based, among others, on data from the Chapel Hill Expert Survey. This data runs from 0 to 10, where low scores indicate leftist partisanship. For the calibration procedure, the empirical anchors were set at 3.75 (full membership), 5 (crossover), and 6.25 (full non-membership). Hence, raw data scores of 3.75 and lower were coded as being "fully inside" the set *leftist partisanship*. For example, the leftist multi-party coalition of Prime Minister Romano Prodi received as left-right score of 3.49. This translated into a fuzzy score of 0.97 for leftist partisanship (almost fully in). Here, it should be highlighted that small differences in the decimal scores in QCA should not be overinterpreted. What is more important are the qualitative differences between cases with fuzzy scores above 0.5 (these are rather inside the set) and those below. Thus, it is also of no utility to report more than two decimals in QCA applications. Table 24.2 lists the calibrated data. The conditions upcoming elections and leadership change were coded in binary fashion as *crisp sets*. The outcome also largely followed a binary coding, but fuzzy sets allowed for further differentiation (e.g., when withdrawal announcements included caveats or no clear timeline was given, resulting in scores just above 0 and below 1, respectively). Finally, the conditions leftist partisanship, fatalities, and low commitment were calibrated with the direct method, resulting in fine-grained fuzzy scores.[9]

Table 24.2: Democratic leaders, conditions, and outcome (calibrated data)

Country/leader (selection, 14 out of 51)	Outcome	Explanatory conditions				
	Early withdrawal	Upcoming elections	Leadership change	Leftist partisanship	Fatalities	Low commitment
Australia (J.W. Howard)	0.10	1.00	0.00	0.00	0.75	0.71
Australia (K. Rudd)	0.90	0.00	1.00	0.93	0.00	0.90
Bulgaria (S. Sakskoburggotski)	1.00	0.00	0.00	0.19	1.00	0.00
Czech Republic (M. Topolánek)	0.20	0.00	1.00	0.00	0.00	0.98
Denmark (A.F. Rasmussen)	0.90	0.00	0.00	0.00	0.99	0.28
Estonia (A. Ansip)	0.00	0.00	1.00	0.31	0.00	0.81
Italy (S. Berlusconi)	0.30	1.00	0.00	0.00	1.00	0.89
Italy (R. Prodi)	1.00	0.00	1.00	0.97	0.00	0.98
Netherlands (J.P. Balkenende)	1.00	0.00	0.00	0.05	0.79	0.20
Poland (D. Tusk)	0.20	0.00	1.00	0.10	0.00	0.72
Portugal (J.M.D. Barroso)	0.00	1.00	0.00	0.03	0.00	0.98
Slovakia (R. Fico)	1.00	0.00	1.00	0.76	1.00	0.83
Spain (J.M. Aznar)	0.00	1.00	0.00	0.00	0.99	0.81
Spain (J.L.R. Zapatero)	1.00	0.00	1.00	0.96	0.61	0.87

Note: All scores reflect fuzzy sets, ranging from 1 (full set membership) to 0 (full set non-membership).

The first step in the QCA analysis proper involves testing for *necessary conditions*. In my analysis, none of the conditions turned out to be formally necessary for the outcome. However, the absence of upcoming elections turned out to be a condition that could be considered "almost necessary" for early withdrawal because its test for necessity comes close to the benchmark of 0.90 consistency. This prompted one of the peer reviewers of the article to request an additional Chi-square test of association, which is not normally part of QCA but in this research setting made sense to also explore the relationship between upcoming elections and early withdrawal from a statistical perspective. This test showed a statistically significant difference in early withdrawal decisions among leaders who faced elections and those who did not. While 31 leaders faced elections during their tenure and their country's Iraq deployment, only two of these initiated a withdrawal in a two-month period before elections (Dominican President Mejia and Portuguese Prime Minister Santana Lopes).

The second step in the analysis is the construction of the *truth table*. Since the analysis entails five conditions, this results in a truth table with 32 rows (configurations of conditions). Of these, 25 rows were populated with empirical cases. Table 3 shows an abbreviated truth table with 6 rows and 14 cases. While it must be kept in mind that this merely presents a part of the complete truth table, it should suffice to illustrate the principles of QCA truth tables (for the complete truth table, see Mello (2020, 63) and the accompanying online supplement. We can see that the top three rows consistently lead towards the outcome (with consistency levels above 0.89) and that the bottom three rows feature low consistencies. Hence, only the top three rows are used for the ensuing Boolean minimization (in addition to other rows not shown in this abbreviated table). On the left side of the truth table, we can see single letters that indicate the five conditions leadership change (L), upcoming elections (E), leftist partisanship (P), low commitment (C), and fatalities (F). Each case is assigned to the row that best describes its configuration of conditions. For example, Spain (Zapatero) is located in the second row, which reflects a new leader, no upcoming elections, leftist partisanship, low coalition commitment, and fatalities. This configuration is shared by Slovakia (Fico) and associated with a perfect consistency of 1, which indicates that membership in this row is a sufficient condition for membership in the outcome.

Table 24.3: Truth table for the outcome early withdrawal (abbreviated)

Conditions					Outcome			
L	E	P	C	F	W	*N*	Consistency	Leaders
0	0	0	0	1	1	3	1.00	Bulgaria (Sakskoburggotski), Netherlands (Balkenende), Denmark (A.F. Rasmussen)
1	0	1	1	1	1	2	1.00	Spain (Zapatero), Slovakia (Fico)
1	0	1	1	0	1	2	0.89	Italy (Prodi), Australia (Rudd)
0	1	0	1	1	0	3	0.25	Italy (Berlusconi), Australia (Howard), Spain (Aznar)
1	0	0	1	0	0	3	0.22	Estonia (Ansip), Poland (Tusk), Czech Republic (Topolánek)
0	1	0	1	0	0	1	0.17	Portugal (Barroso)

Note: L = leadership change, E = upcoming elections, P = leftist partisanship, C = low commitment, F = fatalities, W = early withdrawal.

In the final step of the analysis, the truth table is minimized to gain a solution term for the outcome. It is generally recommended to derive all three solution terms (Schneider and Wagemann 2010), but researchers are –in principle– free to choose which of the solution terms they want to interpret and emphasize in their analysis.[10] This can vary, depending on the research aims and the level of complexity in a given study (especially the number of conditions included). Most often, the conservative solution will be too complex for meaningful interpretation. For my application, I chose the intermediate solution because I wanted to gain a more concise solution but also wanted to exclude some implausible counterfactuals (that is why I did not use the parsimonious solution).

Table 24.4 shows the solution paths towards early withdrawal.[11] There are different ways to summarize QCA solutions (cf. Rubinson 2019). The format displayed in Table 24.4 has the advantages that all conditions can be spelled out on the left side (rather than using acronyms and Boolean expressions), that all relevant measures of fit and all the cases are shown, and that the black and crossed-out circles (for the presence and absence of the respective condition) are more intuitive to read than Boolean notation. Hence, we can see that Path 1 entails the configuration of a new leftist leader without upcoming elections. This is shared by five leaders, as can be seen in the lower area ("covered cases"). *Uniquely covered cases* are those that are only covered by one single path of the solution. One downside of this notation is that the cases are represented by acronyms (to save space). Alternatively, one could of course also spell out the country names and leaders for a stand-alone table.

Table 24.4 further displays equifinality because four paths consistently lead towards the outcome. Each of these contains a different configuration of conditions, which also means that the paths and their underlying logic should be discussed separately in QCA study. The measures of fit below the paths show that all are highly consistent (scores close to 1) but their coverage varies quite a bit. Raw coverage indicates how much of the empirical evidence is accounted for, while unique coverage reflects only that part that is covered solely by the respective path. Hence, we can see that Path 2 has the highest raw coverage, but Path 1 has the highest unique coverage. The bottom of the table reports the overall solution consistency (0.96) and coverage (0.76). This indicates that the solution is highly consistent, but it does not cover all cases, some of which may be accounted for with alternative explanations. We will return to this in the next step.

Table 24.4: Solution for the outcome early withdrawal

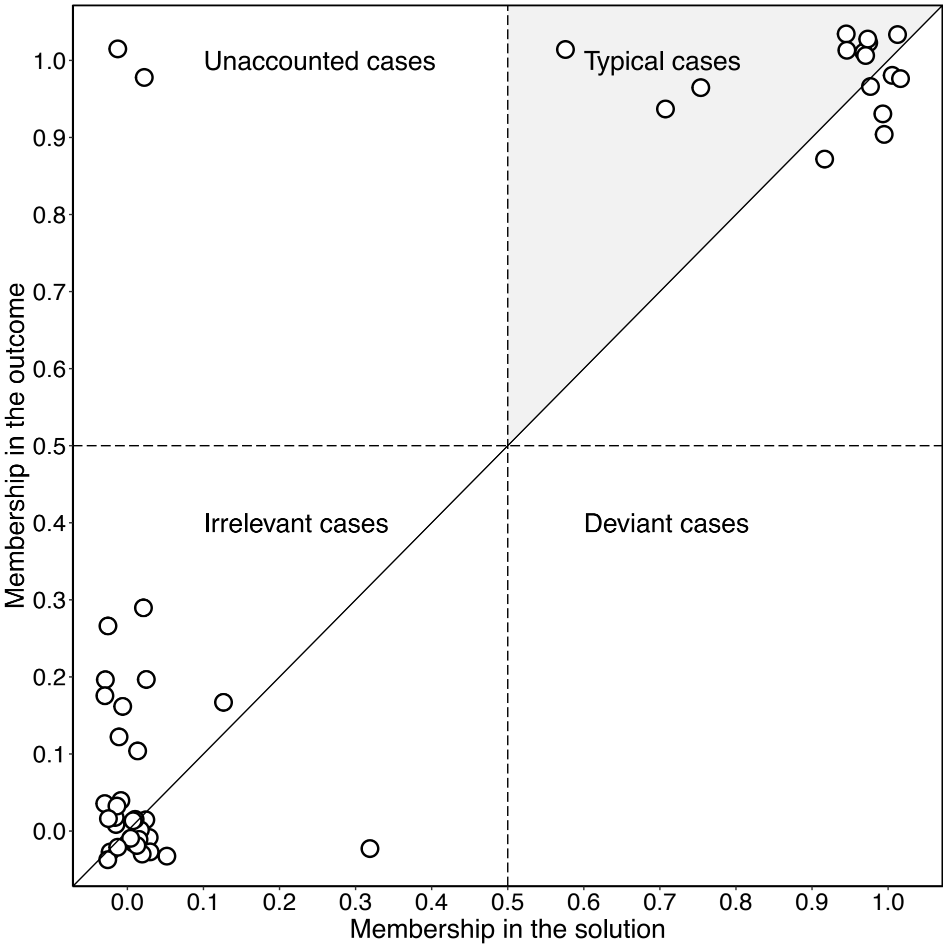
	Solution paths			
	1	2	3	4
Leadership change	●		⊗	⊗
Upcoming elections	⊗	⊗	⊗	⊗
Leftist partisanship	●		⊗	
Low commitment				●
Fatalities		●		
Consistency	0.92	0.99	0.98	0.98
Raw coverage	0.23	0.41	0.39	0.23
Unique coverage	0.15	0.10	0.10	0.04
Covered cases / uniquely covered cases (bold)	**AU2** ES2 **HU2** **IT2** SK2	BG1 DK1 ES2 JP1 **LV2** NL1 PH1 **RO2** SK2	BG1 DK1 **HN1** JP1 **NI1** NL1 NO1 PH1	JP1 NO1 **NZ1** PH1
Solution consistency	0.96			
Solution coverage	0.76			

Note: Black circles indicate the presence of a condition, crossed-out circles its absence.

In addition to summarizing the QCA results as in Table 24.4, it is often a good idea to visualize the position of each case through a scatter plot (also known as XY plot). Figure 24.1 displays such a visualization for the 51 cases entailed in the study used for the empirical illustration. This was created with the R package “ggplot2” (Wickham 2016). The x-axis shows cases’ fuzzy-set membership in the solution and the y-axis shows membership in the outcome (early withdrawal). The plot also includes a diagonal line. This line separates cases with values that are equal to or higher for the outcome than for the solution ($Y \geq X$) from those that are equal to or lower for the outcome than for the solution ($Y \leq X$). Cases with equal values for the solution and the outcome are situated exactly on the diagonal line. In set-theoretic terms, a perfect sufficient condition (or perfectly sufficient solution, which may entail any number of configurations of conditions) is indicated if all cases are on or above the diagonal line. We can see that most cases in Figure 24.1 are located above the diagonal, but some cases are just below it, which means that these cases subtract from the overall solution consistency.[12] However, apart from the numerical estimate of consistency, it is also important to check for qualitative differences in cases’ positions.

Qualitatively speaking, four types of cases exist, separated by the dashed lines that divide the plot into four squares. Cases in the bottom left hold neither membership in the solution nor in the outcome. Hence, these can be largely considered *irrelevant cases* because they do not show the outcome of interest, nor do they feature configurations of conditions that are associated with the outcome. Conversely, cases in the top right hold membership both in the solution and the outcome. These are *typical cases* for the relationship under study. The gray shaded area further defines cases that hold membership in the solution and the outcome, and which are also consistent with a set-theoretic relationship of sufficiency. However, the fact that several cases are just below the line should not be a reason for concern. It simply means that their membership in the solution (slightly) exceeds their membership in the outcome. Qualitatively, these should be treated in same way as other cases in the top right corner. We can also see that there are two cases in the top left corner. These *unaccounted cases* are not part of the solution, but they still show the outcome we are interested in. This is not unusual because there can always be cases that require an alternative explanation. That said, if the number of unaccounted cases grows too large, then this should prompt the researcher to reconsider her research design (and possibly include anther condition or reconceptualize one of the existing ones). Finally, the bottom right corner shows cases that hold membership in the solution, but which do not show the expected outcome. Hence, these are considered *deviant cases*. These pose a problem because they directly challenge the assumed relationship between membership in the solution and the outcome. Empirically, this study entailed no deviant cases but two unaccounted cases, which are discussed in the study itself.

Figure 24.1: Fuzzy-set membership in solution and outcome



Conclusion

While still being considered an emerging method in some corners, QCA has seen a vast increase in empirical applications in IR (Ide and Mello 2022) and also in the narrower field of FPA. It is likely that this trend will continue, not least due to a growing diversity in methodological approaches, moves towards methodological pluralism and analytical eclecticism (Sil and Katzenstein 2010), and increased training opportunities across a broader spectrum of methods at summer schools, workshops, and graduate schools.[13]

That said, what are the method's strengths and limitations and how can new users avoid frequently made mistakes? The main asset of QCA is its ability to account for *causal complexity* – the identification of multiple paths towards the same outcome and combinations of conditions that jointly bring about the outcome. Another asset is the *rigorous analytical framework* which has been strengthened over the years of QCA's development and now presents a robust comparative approach to examine medium to large numbers of cases. This also shows in the increased functionality of the respective R packages, which now allow for a broad range of customization, depending on one's research aims (Dușa 2019; Oana, Schneider, and Thomann 2021). Additionally, robustness tests are become more and more standardized (Oana and Schneider 2021). For peer-reviewed publications it is now more or less expected to conduct robustness tests, especially when the QCA study worked predominantly with quantitative data sets where calibration decisions can be difficult to justify.[14] Finally, it must be highlighted that despite its rigorous structure, QCA allows for *ample flexibility* in tailoring the analysis to the requirements of an individual research project. As such, the method can be used with qualitative and quantitative types of data, in small to large settings and with various forms of analytical steps and approaches (Mello 2021, 2017; Schneider and Wagemann 2012; Rihoux and Ragin 2009; Pagliarin, La Mendola, and Vis 2022).

But the method is not without limitations. The first limitation is empirical: to conduct a QCA the researcher needs to have a certain number of cases and data on these cases. For new research phenomena or topics where data can be sensitive and hard to obtain (e.g., many issues in security policy), this can pose a hurdle that effectively undermines a comparison and may lead researchers to opt for a single case study or a paired comparison (Mello 2017). The second limitation relates to the analysis itself, which is sensitive to changes in the calibration of conditions and the selection of cases (Skaaning 2011; Oana and Schneider 2021). This means that researchers ought to pay close attention to their criteria for calibration and case selection. Analytical decisions must be justified and where alternatives are feasible, robustness tests should be conducted. Finally, even though QCA provides a rigorous comparative framework, its comparisons are inherently static. If researchers are interested in processes of decision-making and temporal aspects of policy, then these elements must either be incorporated at the research design level (Pagliarin and Gerrits 2020), or in combination with other methods such as *Process Tracing* (see van Meegdenburg in this volume). Indeed, multi-method research has been hailed as the new gold standard in the social sciences, even though there are quite different conceptions of what multimethod research entails (cf. Goertz 2017; Seawright 2016).

What are frequently made mistakes in QCA? Here, we can distinguish between conceptual errors at the research design stage and those committed during the analysis and the interpretation of the set-theoretic results (see also Rubinson et al. 2019). To begin with, QCA studies should aim to explore set relations and causal complexity. Researchers should highlight why they apply QCA and to which extent set relations and causal complexity are expected. An often-overlooked issue in research design is the relationship between cases and conditions. Simply put, the number of cases should generously exceed the number of conditions to have enough empirical material for all possible combinations of conditions (but this does not mean that all combinations must be filled empirically). As a rule of thumb, four cases per condition should be included in a study and even higher ratios are recommended when a study features more than five conditions (Mello 2021, 27). This means that studies with few cases should focus their research design on a select number of conditions, rather than aiming to cover all potentially relevant factors. When it comes to the analysis, one often-encountered issue are low coverage scores (e.g., in the range of 0.5/0.6 coverage, and even lower). Frankly, a coverage around 0.5 raises serious questions about a study's selection of conditions (because it indicates that a substantial share of the phenomenon of interest may be explained by alternative factors that were not included). Another important point concerns the selection of the solution term. As indicated above, researchers can choose which of the three different solution terms they want to focus on in their interpretation. However, each requires a thorough justification, which can relate to the research aims, theoretical expectations, the extent to which there is limited diversity, and the nature of the logical remainders. Finally, QCA deals in complexity, but its results should be presented in a parsimonious way. Here, transparency and clarity are key to make it unambiguous for readers (and reviewers) to grasp how a study achieved its results and how results were interpreted (Mello 2021, 191).

What are the prospects for QCA in FPA? As summarized in the literature review section, QCA has increasingly been used in research on foreign policy, and particularly in work on international security. From my perspective, several areas invite further exploration. First, due to its emphasis on a combinatorial logic, QCA should be fruitful as a method to use in combination with *Leadership Trait Analysis* (LTA, see Brummer in this volume) or *Operational Code Analysis* (OCA, see Schafer and Walker in this volume). This could be used, for instance, to explore comparative research questions about the behavior of leaders during crises and conflict (on *Comparative Foreign Policy*, see Feng and He in this volume). To be sure, such an undertaking would require a tremendous data effort and sequential analytical steps, hence this may best be pursued collaboratively as part of a larger project. Second, due to its emphasis on theoretical integration and the combination of systemic and domestic-level variables, *neoclassical realism* appears to be a promising theoretical framework for set-theoretic applications in FPA (Meibauer 2021; Davidson 2011; Rosa and Cuppuleri 2021). At the time of writing, there is only a single study that embraces the combination of neoclassical realism and QCA (Dall'Agnol 2022), but with their inherent synergy such integrative approaches are bound to see wider application in future years.

References

Álamos-Concha, Priscilla, Valérie Pattyn, Benoît Rihoux, Benjamin Schalembier, Derek Beach, and Bart Cambré. 2021. "Conservative Solutions for Progress: On Solution Types when Combining QCA with In-depth Process-Tracing." *Quality & Quantity* (3 August).

Böller, Florian. 2021. "Brakeman or Booster? Presidents, Ideological Polarization, Reciprocity, and the Politics of US Arms Control." *International Politics* (7 July).

Böller, Florian, and Marcus Müller. 2018. "Unleashing the Watchdogs: Explaining Congressional Assertiveness in the Politics of US Military Interventions." *European Political Science Review* 10 (4): 637-662.

Boogaerts, Andreas. 2018. "Beyond Norms: A Configurational Analysis of the EU's Arab Spring Sanctions." *Foreign Policy Analysis* 14 (3): 408-428.

Boogaerts, Andreas, and Edith Drieskens. 2020. "Lessons from the MENA Region: A Configurational Explanation of the (In)Effectiveness of UN Security Council Sanctions between 1991 and 2014." *Mediterranean Politics* 25 (1): 71-95.

Braumoeller, Bear F. 2003. "Causal Complexity and the Study of Politics." *Political Analysis* 11 (3): 209–233.

Brummer, Klaus, and Kai Oppermann. 2019. *Aussenpolitikanalyse*. Berlin: De Gruyter.

Caspersen, Nina. 2019. "Human Rights in Territorial Peace Agreements." *Review of International Studies* 45 (4): 527-549.

Cronqvist, Lasse. 2019. *Tosmana [Version 1.61]*. University of Trier.

Dall'Agnol, Augusto C. 2022. "The Diffusion of Military Power: A Neoclassical Realist Analysis." Paper presented at the ISA Annual Convention.

Davidson, Jason W. 2011. *America's Allies and War: Kosovo, Afghanistan, and Iraq*. New York: Palgrave Macmillan.

de Sá Guimarães, Feliciano, and Maria Hermínia Tavares de Almeida. 2017. "From Middle Powers to Entrepreneurial Powers in World Politics: Brazil's Successes and Failures in International Crises." *Latin American Politics and Society* 59 (4): 26-46.

Duşa, Adrian. 2019. *QCA with R. A Comprehensive Resource*. Cham: Springer.

Gerrits, Lasse, and Stefan Verweij. 2014. "Critical Realism as a Meta-Framework for Understanding the Relationships between Complexity and Qualitative Comparative Analysis." *Journal of Critical Realism* 12 (2): 166-182.

Goertz, Gary. 2017. *Multimethod Research, Causal Mechanisms, and Case Studies*. Princeton, NJ: Princeton University Press.

Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton, NJ: Princeton University Press.

Grynaviski, Eric, and Amy Hsieh. 2015. "Hierarchy and Judicial Institutions: Arbitration and Ideology in the Hellenistic World." *International Organization* 69 (3): 697-729.

Haesebrouck, Tim. 2017a. "EU Member State Participation in Military Operations: A Configurational Comparative Analysis." *Cambridge Review of International Affairs* 30 (2-3): 137-159.

---. 2017b. "NATO Burden Sharing in Libya: A Fuzzy Set Qualitative Comparative Analysis." *Journal of Conflict Resolution* 61 (10): 2235-2261.

---. 2018. "Democratic Participation in the Air Strikes Against Islamic State: A Qualitative Comparative Analysis." *Foreign Policy Analysis* 14 (2): 254-275.

---. 2022. "Relevant, Irrelevant, or Ambiguous? Toward a New Interpretation of QCA's Solution Types." *Sociological Methods & Research*.

Haesebrouck, Tim, and Eva Thomann. 2021. "Introduction: Causation, Inferences, and Solution Types in Configurational Comparative Methods." *Quality & Quantity* (July 29).

Haesebrouck, Tim, and Anouschka van Immerseel. 2020. "When does politics stop at the water's edge? A QCA of parliamentary consensus on military deployment decisions." *European Political Science Review* 12 (3): 371-390.

Heldt, Eugénia C., Patrick A. Mello, Anna Novoselova, and Omar Ramon Serrano Oswald. 2022. "When Do International Organizations Engage in Agency Slack? A Qualitative Comparative Analysis of United Nations Institutions." *Global Studies Quarterly* (forthcoming).

Hudson, Valerie M., and Benjamin S. Day. 2019. *Foreign Policy Analysis: Classic and Contemporary Theory*. Lanham: Rowman & Littlefield.

Ide, Tobias, and Patrick A. Mello. 2022. "QCA in International Relations: A Review of Strengths, Pitfalls, and Empirical Applications." *International Studies Review* 24 (1): 1-20.

Ide, Tobias, Miguel Rodriguez Lopez, Christiane Fröhlich, and Jürgen Scheffran. 2021. "Pathways to Water Conflict During Drought in the MENA Region." *Journal of Peace Research* 58 (3): 568-582.

Kaarbo, Juliet. 2022. "What Happened to Vladimir Putin?". International Affairs Blog. https://medium.com/international-affairs-blog/what-happened-to-vladimir-putin-5755addf2a15.

Kisangani, Emizet F., and Jeffrey Pickering. 2022. *African Interventions: State Militaries, Foreign Powers, and Rebel Forces*. New York: Cambridge University Press.

Kuehn, David, and Ingo Rohlfing. 2022. "Do Quantitative and Qualitative Research Reflect two Distinct Cultures? An Empirical Analysis of 180 Articles Suggests “no”." *Sociological Methods & Research* (31 March).

Mahoney, James. 2021. *The Logic of Social Science*. Princeton, NJ: Princeton University Press.

Meibauer, Gustav. 2021. "Neorealism, neoclassical realism and the problem(s) of history." *International Relations* (29 July).

Meissner, Katharina L., and Patrick A. Mello. 2022. "The Unintended Consequences of UN Sanctions: A Qualitative Comparative Analysis." *Contemporary Security Policy* 43 (2): 243-273.

Mello, Patrick A. 2012. "Parliamentary Peace or Partisan Politics? Democracies’ Participation in the Iraq War." *Journal of International Relations and Development* 15 (3): 420–453.

---. 2014. *Democratic Participation in Armed Conflict: Military Involvement in Kosovo, Afghanistan, and Iraq*. Basingstoke: Palgrave Macmillan.

---. 2017. "Qualitative Comparative Analysis and the Study of Non-State Actors." In *Researching Non-State Actors in International Security: Theory & Practice*, edited by Andreas Kruck and Andrea Schneiker, 123-142. Oxon: Routledge.

---. 2020. "Paths towards Coalition Defection: Democracies and Withdrawal from the Iraq War." *European Journal of International Security* 5 (1): 45-76.

---. 2021. *Qualitative Comparative Analysis: An Introduction to Research Design and Application*. Washington, DC: Georgetown University Press.

---. 2022. "Incentives and Constraints: A Configurational Account of European Involvement in the Anti-Daesh Coalition." *European Political Science Review* (24 February): 1-19.

Oana, Ioana-Elena, and Carsten Q. Schneider. 2018. "SetMethods: An Add-on R Package for Advanced QCA." *The R Journal* 10 (1): 507-533.

---. 2021. "A Robustness Test Protocol for Applied QCA: Theory and R Software Application." *Sociological Methods & Research* (26 August).

Oana, Ioana-Elena, Carsten Q. Schneider, and Eva Thomann. 2021. *Qualitative Comparative Analysis Using R: A Beginner’s Guide*. New York: Cambridge University Press.

Oppermann, Kai, and Klaus Brummer. 2020. "Who Gets What in Foreign Affairs? Explaining the Allocation of Foreign Ministries in Coalition Governments." *Government and Opposition* 55 (2): 241-259.

Pagliarin, Sofia, and Lasse Gerrits. 2020. "Trajectory-Based Qualitative Comparative Analysis: Accounting for Case-Based Time Dynamics." *Methodological Innovations* online first (27 September) (doi:10.1177%2F2059799120959170).

Pagliarin, Sofia, Salvatore La Mendola, and Barbara Vis. 2022. "The “qualitative” in qualitative comparative analysis (QCA): research moves, case-intimacy and face-to-face interviews." *Quality & Quantity*.

Pilster, Ulrich, Tobias Böhmelt, and Atsushi Tago. 2013. "Political Leadership Changes and the Withdrawal from Military Coalition Operations, 1946-2001." *International Studies Perspectives* 16 (4): 463-483.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.

---. 2006. "Set Relations in Social Research: Evaluating Their Consistency and Coverage." *Political Analysis* 14 (3): 291–310.

---. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Ragin, Charles, and Sean Davey. 2017. *fs/QCA [Version 3.0]*. Irvine, CA: University of California.

Rihoux, Benoît, and Charles C. Ragin, eds. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, CA: Sage.

Rosa, Paolo, and Adriana Cuppuleri. 2021. "Dangerous dyads in the post-Soviet space: explaining Russia's military escalation decisions, 1992–2010." *Italian Political Science Review/Rivista Italiana di Scienza Politica* 51 (3): 355-372.

Rubenzer, Trevor. 2008. "Ethnic Minority Interest Group Attributes and U.S. Foreign Policy Influence: A Qualitative Comparative Analysis." *Foreign Policy Analysis* 4 (2): 169-185.

Rubinson, Claude. 2019. "Presenting Qualitative Comparative Analysis: Notation, Tabular Layout, and Visualization." *Methodological Innovations* 12 (2): 1-22.

Rubinson, Claude, Lasse Gerrits, Roel Rutten, and Thomas Greckhamer. 2019. Avoiding Common Errors in QCA: A Short Guide for New Practitioners.

Rutten, Roel. 2020. "Applying and Assessing Large-N QCA: Causality and Robustness from a Critical-Realist Perspective." *Sociological Methods & Research* (26 April): 1-21.

Schmitt, Olivier. 2018. *Allies that Count: Junior Partners in Coalition Warfare*. Washington, DC: Georgetown University Press.

Schneider, Carsten Q., and Claudius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." *Comparative Sociology* 9 (3): 397–418.

---. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. New York, NY: Cambridge University Press.

Seawright, Jason. 2016. *Multi-Method Social Science: Combining Qualitative and Quantitative Tools*. Cambridge: Cambridge University Press.

Sil, Rudra, and Peter J. Katzenstein. 2010. *Beyond Paradigms: Analytic Eclecticism in the Study of World Politics*. Basingstoke: Palgrave Macmillan.

Skaaning, Svend-Erik. 2011. "Assessing the Robustness of Crisp-set and Fuzzy-set QCA Results." *Sociological Methods & Research* 40 (2): 391–408.

Tago, Atsushi. 2009. "When Are Democratic Friends Unreliable? The Unilateral Withdrawal of Troops from the 'Coalition of the Willing'." *Journal of Peace Research* 46 (2): 219–234.

Treisman, Daniel. 2022. "Putin Unbound: How Repression at Home Presaged Belligerence Abroad." *Foreign Affairs* 101 (3): 40-53.

van Nieuwenhuizen, Simone. 2019. "Australian and People's Republic of China government conceptions of the international order." *Australian Journal of International Affairs* 73 (2): 181-197.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

[1] A search on the Web of Science using the terms "qualitative comparative analysis" and "foreign policy" identifies 15 SSCI journal articles that empirically apply the method of QCA (as of May 11, 2022). The first of these was published in 2008, but the majority of the articles appeared throughout the last few years, showing an upward trajectory in publication trends.

[2] This should not imply that statistical research cannot model causally complex relationships (for an exception, see Braumoeller 2003). For an empirical examination of "quantitative" and "qualitative" cultures, see Kuehn and Rohlfing (2022).

[3] For a discussion of the mathematical transformation of raw data into calibrated data, see Ragin (2008). An illustration and further applied examples are given in Mello (2021).

[4] To be precise, the consistency of a truth table row is calculated on the basis of all cases' set-membership scores for the respective configuration. With fuzzy sets, cases may also hold partial membership in a configuration (e.g. a fuzzy score of 0.1, as in being almost entirely outside the set) but their score still affects the calculation of consistency. This explains why consistency may not be "perfect" even when all cases with membership above 0.5 (cases that are qualitatively considered to be inside the set) also show the outcome.

[5] A comprehensive list of software is maintained on the COMPASSS website (which also hosts other QCA-related resources): https://compasss.org/software/.

[6] A user-friendly introductory manual to QCA in R complements my textbook (Mello 2021). This can be accessed at: https://patrickmello.com/qca-research-design-application/.

[7] This is an important difference to some statistical approaches, where a certain extent of missing data is expected. In QCA, complete data is required to proceed with the analysis.

[8] On guidelines for using qualitative data in QCA, for instance when working with information from interviews or focus groups, see Pagliarin, La Mendola, and Vis (2022).

[9] Further details on the calibration of these conditions can be found in the online supplement to Mello (2020), available at: https://doi.org/10.7910/DVN/8UWS1R.

[10] For different perspectives on QCA solution terms and their interpretability, see, among others, Haesebrouck and Thomann (2021), Álamos-Concha et al. (2021), and Haesebrouck (2022).

[11] A side-by-side comparison of all three solutions is provided in Table S5 in the online supplement to Mello (2020).

[12] On the calculation of consistency and coverage, see Ragin (2006).

[13] Regular courses and workshops on QCA are offered at the ECPR Summer School of Methods and Techniques, the FORS Summer School in Social Science in Lugano, and as part of the newly instituted MethodsNET. The COMPASSS website and newsletter provide regular updates on courses offered (see link above).

[14] For examples on how such robustness tests can be documented, see, Haesebrouck (2017a) and Ide et al. (2021). Meissner and Mello (2022) also apply the new robustness test protocol for QCA (Oana and Schneider 2021).