The potential of consecutive qualitative comparative analysis as a systematic strategy for configurational theorizing

Krupnik, S.; Szczucka, A.; Wozniak, M.; Pattyn, V.E.

Article

The potential of consecutive qualitative comparative analysis as a systematic strategy for configurational theorizing

Evaluation
1–17

DOI: 10.1177/13563890231200292
journals.sagepub.com/home/evi
Sage

Seweryn Krupnik
Jagiellonian University, Poland

Anna Szczucka
Jagiellonian University, Poland

Monika Woźniak
Jagiellonian University, Poland

Valérie Pattyn
Leiden University, The Netherlands

Abstract

Qualitative comparative analysis is gradually becoming more established in the evaluation field. The purpose of this article is to highlight the potential for evaluation research of engaging in consecutive rounds of this analysis. This is possible when approaching qualitative comparative analysis as a systematic strategy for configurational theorizing. To substantiate this potential, we present two evaluation studies on Research and Development subsidies for companies in Poland. Compared with the results of the first study, the findings of the subsequent consecutive qualitative comparative analysis studies were much more nuanced and helped in developing a full-fledged configurational program theory. In addition to elaborating on the strengths of a consecutive qualitative comparative analysis approach and the relevance of configurational program theories for evaluators, this article shares the main lessons learned in overcoming challenges common to such designs. Thus, concrete guidance is offered to researchers and evaluators who are willing to take configurational theorizing seriously.

Corresponding author:
Seweryn Krupnik, Institute of Sociology, Center for Evaluation and Analysis of Public Policies, Jagiellonian University, Grodzka St. 52, 30-962 Kraków, Poland.
Email: seweryn.krupnik@uj.edu.pl

Keywords
configurational theorizing, knowledge cumulation, qualitative comparative analysis, research and development subsidies

Introduction

Qualitative comparative analysis (QCA) allows for a systematic analysis of the set-theoretic relationship between conditions and a particular outcome (effect or impact) on the basis of a comparison of cases. The approach and its related techniques were introduced to the evaluation community in previous publications (Befani, 2020; Kahwati and Kane, 2020; Pattyn et al., 2019; Stern et al., 2012). Notably, readers of this journal have had an opportunity to learn about QCA applications in the context of evaluation studies (Befani, 2013; Blackman et al., 2013; Krupnik and Koniewski, 2022; Verweij and Gerrits, 2013) or in the context of studies examining evaluation practice itself (Balthasar, 2006; Pattyn, 2014). In addition, within the adjacent field of policy analysis, there are numerous applications of set-theoretic methods (see Rihoux et al., 2011 for an overview, or Ragin and Fiss, 2017; Thomann, 2020) which hold relevance for evaluators, and evaluators interested in applying QCA in their studies have access to a growing number of examples, evaluation-specific manuals (Befani, 2016), and standards of good practice (Rubinson et al., 2019; Thomann et al., 2022; Wagemann and Schneider, 2015). Lemire et al. (2020) describe its growing number of applications as illustrative for one of the important trends in evaluation studies: understanding *why* programs work. Yet, while QCA is continuously advancing, existing applications are not always consistent with best practice and are not always aligned with QCA as an "approach" that relies on configurational thinking (Thomann et al., 2022).

The present study aims to add to the understanding of how to apply QCA in the evaluation field. Most—if not almost all—QCA evaluations limit the application to a single QCA study, in which the research process is concluded after one "QCA research cycle." Thus, knowledge about program theories that underpin policy interventions has likely not been fully exploited. Fine-tuning robust program theories is an iterative process that involves ongoing reciprocal input between evaluators' expert knowledge, social science theory, stakeholders' opinions, and results of empirical studies (Coryn et al., 2011). Moreover, lessons drawing in public policy often proceed in non-linear and iterative ways, in which existing policy beliefs and assumptions are updated after multiple feedback loops, and after accumulating enough knowledge (Jenkins-Smith and Sabatier, 1999).

The purpose of this article, then, is to highlight the benefits of engaging in consecutive rounds of QCA. Consecutive QCA refers to the application of QCA in a series of separate studies carried out at different moments in time. The studies can have different samples and specific research objectives. Collectively, however, they contribute to an overarching research goal. In such a stepwise design, key parameters of the subsequent QCA, such as model building and case selection, are informed by the results of the former round. We show that consecutive QCA can serve cumulative knowledge-building about the evaluand, more so than when engaging in a single "standard" QCA cycle. Conducting consecutive QCA adds an additional layer to the iterative approach characterizing each individual QCA study. In simple terms, consecutive QCA involves multiple interconnected QCA studies, with each study also being conducted iteratively.

Consecutive QCA is particularly well-suited when approached as a systematic strategy for configurational theorizing, rather than solely as a research technique. In evaluation settings, this entails the necessity of framing program theories (or logic models) as configurations of conditions that lead to an outcome. Program theories depict relationships among program elements (*program inputs, activities, outputs, and* outcomes*)* and provide causal explanations for why a given program may work (Astbury and Leeuw, 2010; Rogers et al., 2000). Theory-driven evaluation is commonly understood as an "evaluation strategy or approach that explicitly integrates and uses stakeholder, social science, some combination of, or other types of theories in conceptualizing, designing, conducting, interpreting, and applying an evaluation" (Coryn et al., 2011: 201). Program theories are not rigid, but are refined based on new evidence. Despite the growing emphasis on theory building and refinement in evaluation scholarship (Turnbull, 2002; Turner et al., 2018; Vellema et al., 2013), the framing of theories in configurational terms that elucidate, for instance, how the amalgamation of different program outputs can foster the desired outcomes in a particular context remains limited. Nonetheless, one can draw inspiration from the wide array of configurational theories available in management studies (Fiss et al., 2013).

To substantiate the potential of consecutive QCA and configurational theorizing in an evaluation context, we present two evaluation studies conducted by authors of the article on research and development (R&D) subsidies in Poland. The aim of each study was to unravel the configurations of conditions leading to companies' successful use of such subsidies. Learning from consecutive QCA models helped in developing a full-fledged configurational program theory.

In the first section of the article, we position our argument in the context of literature related to QCA and consecutive theorizing. In the second section, we briefly introduce our case study, which leads to our explanation of the potential of consecutive QCA and configurational theorizing in the third section. We conclude the article with a summary of the lessons learned from our case study.

Configurational theorizing and consecutive qualitative comparative analysis

QCA is set-theoretic, that is, conditions and outcome are depicted as sets and cases have assigned membership scores to sets (Schneider and Wagemann, 2012). Complex patterns between conditions and outcome can thus be modeled on this basis. Complexity, moreover, can manifest in different ways. First, configurations of conditions can lead to an outcome (i.e. "conjunctural causation," as opposed to one condition being sufficient for its occurrence). Second, several configurations of conditions can lead to an outcome (i.e. "equifinality"). Third, if the presence of a particular combination of conditions is relevant for the outcome, its absence is not necessarily relevant to the absence of the outcome (i.e. "causal asymmetry"). The relationships between conditions and outcome are described with Boolean algebra. For example, if we analyze conditions B, E, and H, along with an outcome S, the result might be $B \times E + E \times H \rightarrow S$. In this example, B could represent big companies, E represents companies that have had experience with receiving similar public support, H represents companies belonging to the high-tech sector, and S represents companies successfully taking advantage of public support. The results indicate that two different configurations—either the co-occurrence of B and E *or* the co-occurrence E and H—lead to the outcome. However, this result does not tell us anything about configurations of conditions leading to the absence of S. Therefore, the conditions leading to the absence of S should be analyzed separately.

Any QCA application goes beyond the so-called "analytical moment" in which researchers resort to software and algorithms to arrive at Boolean solution linking conditions and outcome (Ragin and Rubinson, 2009; Thomann and Ege, 2020). A full QCA research cycle equally entails "the processes before and after the analysis of the data, such as the (re-)collection of data, (re-)definition of the case selection criteria, or (re-)specification of concepts, often based on preliminary insights gained through QCA-based data analysis" (Schneider and Wagemann, 2012: 11). *Before* the analytical moment, the outcome and conditions must be conceptualized and operationalized. Even if these steps are also present in other social research approaches, they must be conducted in accordance with the premises of the set-theoretic approach described above. Moreover, *after* the analytical moment, the researcher is expected to relate the results to cases, and often, more in-depth analysis of specific cases is required. All these steps are essential in any QCA research cycle.

Irrespective of the stages through which QCA studies typically proceed, several research purposes can be served. The most basic distinction is between inductive and more deductive purposes (Thomann and Maggetti, 2017). Berg-Schlosser and De Meur (2009) differentiate between five types of usage: summarizing data, checking data coherence, checking hypotheses or existing theories, conducting quick tests of conjectures, and developing new theoretical arguments. In this article, we focus on the last option: the potential of QCA for theory development in the context of evaluation studies. Illustrative evaluation examples of theory development include areas of skills development (Álamos-Concha et al., 2020), and developmental (Befani, 2016) and environmental interventions (Befani and Sager, 2006).

We apply "configurational theorizing" approach to theory development. This implies a focus on "understanding how or why multiple attributes combine into distinct configurations to explain a phenomenon, while also recognizing that complex causal explanations may involve more than one configuration of attributes leading to the outcome of interest" (Furnari et al., 2021: 779). Thus, it is closely aligned with the general principles of QCA presented above. Configurational theorizing is based on a distinct approach to causality (Sterntl et al., 2012) and to protocols for conducting research (Fiss et al., 2013). To be clear, while configurational theorizing is associated with QCA, it also allows the use of other set-theoretic methods (e.g. coincidence analysis). Its very name makes its crucial distinctive features intuitive to grasp by less methodology-oriented members of the evaluation community and evaluation stakeholders.

As mentioned in the introduction, theorizing typically occurs iteratively in an evaluation context: Program theories are built on diverse sources of input (i.e. literature reviews, secondary data, and stakeholders) (Coryn et al., 2011) and knowledge must be summarized across multiple interventions (Befani and Sager, 2006). More than in other research settings, therefore, an iterative approach to configurational theorizing is essential; this is, however, far from easy. Furnari et al. (2021) illustrate that the cumulation of knowledge involves many feedback loops between the stages. In this article, we show that consecutive QCA (i.e. applying QCA at different stages of the evaluation process) may be instrumental to such a purpose. In the following sections, we illustrate this with a real evaluation case.

Case study: Evaluation of Polish R&D support schemes for companies

The case that serves as an illustration for consecutive QCA concerns an evaluation of the biggest Polish R&D support scheme for companies. The scheme is run by the National Centre for

Scheme	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Measure 1.1.1. OPSG				Call	Projects			Model 1			
Measure 1.4 OPIE	Call		Projects							Model 2a	Model 2b

Table 1. Timeline of investigated programs, supported projects, and analyses.

Research and Development, the executive agency of the Ministry of Regional Development in Poland. The main goal of the scheme, which was implemented in 2012 and targets companies involved in industrial research or development, is to foster inventions or new industrial designs and to promote implementation of breakthrough innovations. The goal is reflected in the scheme's desired outcome at the company level: the introduction of the supported innovations to the Polish market. This outcome is operationalized as the income that companies earned as a result of R&D activities. Single companies, a consortium of companies, and scientific units can apply to the scheme and receive funding for specific projects.

This support scheme was implemented within different programs, of which the two most current ones were analyzed and subjected to an evaluation. These constitute the focus of this article. The first program was Measure 1.4 in Operational Program Innovative Economy (OPIE), which was implemented in the programming period 2007–2014 (Ministry of Regional Development in Poland, 2012). The subsequent program was Measure 1.1.1 in European Union (EU) Operational Program Smart Growth (OPSG), which was implemented in the programming period 2015–2021 (Ministry of Development Funds and Regional Policy, 2022). 447 million EUR and 2.7 billion EUR were allocated for the programs, respectively.

For both evaluation studies, QCA served as the main method for investigating the conditions under which the R&D subsidies were successful. Altogether, across both studies, three QCA models were created: one (Model 1) for Measure 1.1.1 OPSG and two models (2a and 2b) for 1.4 OPIE. The conceptualization of the outcome was identical across the models. Both evaluations investigated beneficiaries (companies or consortia) whose projects were advanced enough at the time of our research. That is, we included only beneficiaries who started their project (Model 1) or ended it (Models 2a and 2b) 3 years before the respective studies took place (Table 1).

The first model involved 36 cases and was built using monitoring data from the agency implementing the subsidies. The second study (Models 2a and 2b) started with an extensive literature review, with the aim of creating a longlist of factors deemed potential influencers of the outcome.[1] On the basis of our in-depth knowledge of the program in question, we narrowed down the list of factors (conditions) to the most relevant ones for our particular evaluation setting. These were included in the second QCA model (2a), which was built using data for 89 cases. We requested an interview with each of the companies represented in Model 2a. The 34 companies that agreed to an interview were included in the third QCA model (2b). This model included additional data about conditions we identified earlier in the literature review. The QCA analysis was conducted with R (Dușa, 2018; Oana and Schneider, 2018) and fsQCA software (Ragin, 2017). The details of the analyses leading to all three models are described in Table 2 and the Supplementary Material.

Showcasing knowledge cumulation through consecutive QCA studies

In this section, we illustrate the potential of consecutive QCA studies and configurational theorizing for knowledge cumulation. We distinguish among four stages, which are common in a configurational theorizing process. An influential paper by Furnari et al. (2021) inspired our approach. They insightfully link every stage to different heuristics, which researchers can instrumentally use in the process of cumulative knowledge-building. Several of these heuristics proved highly relevant to our QCA study—we mention them in parentheses when presenting the stages. Identifying relevant conditions constitutes the first stage, called *scoping*. Evaluators may have to contend with the challenge of too many conditions that are potentially relevant to the outcome, a situation common in case-based research. Two heuristics help overcome this challenge: first, the evaluator starts with one key attribute before identifying other important ones ("Complexify from an anchor"), and then aggregates the attributes according to their conceptual similarity ("Simplify to higher order constructs"). At the second stage of theorizing, *linking*, evaluators identify configurations of conditions that they think may lead to an investigated outcome. During this task it is important to consider alternative paths leading to an outcome ("Think equifinally"), while also considering how the absence of some conditions can be an important element of configurations ("Think about absence"*)*. Finally, at the third stage, *naming*, evaluators label the configurations. The labels should be easily understood by the audience ("Articulate with simplicity") and should enable holistic interpretation across configurations ("Capture the whole").

To fully account for the nature of a QCA process, we deem it relevant to consider a fourth stage of knowledge cumulation, which complements the stages distinguished by configurational theorists. In QCA, conditions and outcomes are translated into sets, a procedure known as "calibration." Consecutive QCA offers much potential for precise cumulative learning about the calibration, and about the operationalization of conditions and outcomes, generally. We, therefore, include it as a separate aspect through which knowledge accumulation can occur in consecutive QCA studies. Table 3 provides an overview of the main insights collected during the evaluation process.

In what follows, we discuss how we proceeded from one model to another (1, 2a, and 2b), after providing a concise description of the knowledge base on which we could rely before starting the study.

Before the evaluation

The effectiveness of R&D subsidies has been studied since the early 1980s. The most cited research includes David et al. (2000), Clausen (2009), and Czarnitzki and Lopes-Bento (2013). The abundance of relevant articles has made it possible to conduct systematic reviews and meta-evaluations (Cunningham et al., 2016; Dimos and Pugh, 2016; Petrin, 2018). As these reviews and meta-evaluations clearly show, generating income from introducing an R&D project to the market is treated as a crucial indicator of success. These studies suggest that the final results can be explained mostly by a mix of companies' capabilities and orientations (as measured by size, experience, and financial situation), the subsidies' features (e.g. grant size), the projects' characteristics (e.g. innovativeness of R&D work), and external factors (e.g. a specific sector). The crucial role of company characteristics led some researchers to write

Table 2. Details of the models.

Model details		Conditions used in models							
Studied program	Data	Beneficiaries' capabilities and orientation				Cooperation during the project	Product characteristics	Implementation conditions	
Model 1									
Measure 1.1.1 OPSG	Agency data from applications and monitoring reports	Microenterprise (MIC)			Experienced in the implementation of innovations (EXP_S)				
Model 2a									
Measure 1.4 OPIE	Agency data from applications and monitoring reports	Large company (BIG)	On the market for 5 years or longer (OLD5)		Experienced in the implementation of innovations (IMP3)	Subcontractor in the project (SUB)	Technological readiness on a level higher than 3 (TRL)	The company has similar products in its portfolio (USU)	The solution is present in the market (PCT)
Model 2b									
Measure 1.4 OPIE	Agency data from applications and monitoring reports Interview data	Macrocondition—Access to inspiring resources and knowledge transfer (RES), based on: *Linked to another enterprise (AUT)* *Large company (BIG)* *Knowledge transfer through informal cooperation with R&D unit before the project (INF)* *RES = (AUT + BIG) × INF*		Macrocondition—Strategically oriented (STR), based on: *Strategic importance of the project for the company (STR2)* *Strategic importance of the revenue from this project (REV)* *STR = STR2 × REV*	Macrocondition—Experienced in R&D (EXP), based on: *Experienced in R&D projects (IMP)* *Experienced in R&D commercialization (COM)* *Experienced in similar projects (USU)* EXP = IMP × COM × USU	Macrocondition—Formal cooperation (FORM), based on: *The R&D public units as subcontractors in R&D stage (SUB)* *Other companies' involvement in commercialization stage (FCOM)* FORM = SUB + FCOM		Macrocondition—Expanding existing business (BAU), based on: *Introducing similar products to the market (SIM) through the same sales channels (CAN)* BAU = SIM × CAN	No radical changes in a market's context (CON)

Table 3. Scoping, linking, naming, and calibrating in consecutive QCA.

	Before the evaluation	Model 1	Model 2a	Model 2b
Identifying relevant conditions (scoping)	Longlist of conditions related to: (a) companies having an earlier advantage (b) features of the project (c) external factors	Five conditions included in the analysis Omitted conditions in the model (low consistency)	Larger number of conditions identified in the literature review and included in the model	Navigating among many conditions and alternative models: Creating macroconditions *(H: Simplify to higher order constructs*)* Categorizing conditions into triggering and auxiliary (*H: Complexify from an anchor*) Decisions based on familiarity with cases
Proposing configurations of conditions (linking)	Conjunction of (a) more capabilities and resources (b) low-risk projects (c) friendly circumstances **knowledge gap: configurations leading to heterogeneity**	Being big and having experience (obtaining sales revenues from new product in previous years) leading to an outcome	Seven configurations, including companies without clear advantage *(H: Think equifinally)*; limited interpretability	Seven interpretable configurations showing diverse advantages companies may use to achieve success Absence of experience linked to formal cooperation (*H: Think about absence*)
Labeling configurations (naming)	Matthew effect *(H: Articulate with simplicity)*	Companies with advantage *(H: Articulate with simplicity)*	Configurations cannot be labeled because not enough familiarity with cases; therefore, limited interpretability	Motivated learners and Pragmatic entrepreneurs *(H: Capture the whole)*
Operationalizing conditions and outcome (calibrating)	Income as crucial indicator of success	Outcome should be measured in a longer time frame (at least 3 years after the project)	Limitations of measuring the outcome based only on monitoring data	Outcome and conditions saturated by diverse data sources

Note: QCA = qualitative comparative analysis.
*The use of heuristics (H) is indicated with italics.

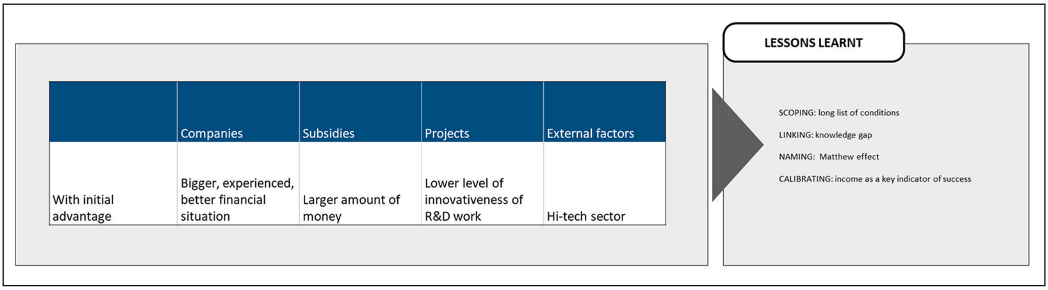

Figure 1. Model before the evaluation.

about the cumulative advantage for companies receiving the subsidies. That is, to use one of the heuristics—"Articulate with simplicity"—companies with advantage achieve success (Figure 1). This process is known in the social sciences as the Matthew effect (Merton, 1968). Broadly defined in the context of R&D subsidies, it refers to granting subsidies to companies who are already more competitive. The beneficiaries take successful advantage of subsidies and are even more competitive after participating in the program (Krupnik, 2012). Although the Matthew effect is acknowledged in evaluation and among policy-makers, its interpretation from a public policy perspective remains ambiguous and the process itself is under-researched (Antonelli and Crespi, 2013; Pereira and Suarez, 2017).

Most of the studies mentioned above indicate that, on average, R&D subsidies lead to higher sales of innovative products. However, an increase has not been observed for each beneficiary (Petrin, 2018). This heterogeneity reflects the claim of realistic evaluation that there is no public intervention that works for everyone, and that public interventions may generate different effects depending on the contextual factors in which they are implemented (Pawson and Tilley, 1997). It is precisely the inconclusiveness of many studies on subsidies that has led to calls to consider alternative research designs. Cunningham et al. (2016) and others, for instance, underlined the importance of going beyond the average impact for all beneficiaries and investigating the effects for specific groups of beneficiaries. Similar conclusions were drawn by Mouque (2012: 12), who suggested "the use of other forms of evaluation (notably observational 'theory-based' methods such as case studies) to shed light on the results." Our article addresses this call. Furthermore, our study explicitly built on existing literature but took it a step further by engaging in QCA research. QCA, as mentioned, provides the potential to account for contextual factors that affect how R&D subsidies generate effects (or a lack thereof) within firms, while also enabling modest generalization.

Model 1

Model 1 indeed confirmed the presence of a Matthew effect among the investigated beneficiaries (Figure 2). Companies that were bigger than microenterprises and experienced in introducing new products to the market achieved success.

However, the model also had some limitations. While the interviews following the QCA analytical moment helped us identify how our model could be improved, the limitations concerned both the outcome and the included conditions. With respect to the outcome, we realized that income from sales of the R&D activities' effects was measured much too early. For many

	MIC	EXP_S	N
Not being microenterprise, obtaining sales revenues in previous years	~	■	20

LESSONS LEARNT

SCOPING: omitted conditions

LINKING: being bigger than microenterprise and experienced leads to quick outcome + the need to identify other paths

NAMING: companies with advantage

CALIBRATING: outcome measured too early

Figure 2. Model 1.
Note: Condition may be present (black), absent ("~") or not included in the pathway; N = 36, Consistency = 0.75, Coverage = 0.83.

	OLD5	BIG	USU	IMP3	PCT	SUB	TRL	N
Old, big, experienced companies working on similar innovative solutions with subcontractors	■	■	■	■	~	■		7
Old not big companies working on new not innovative solutions with high TRL with subcontractors	■	~	~		■	■	■	4
Old not big inexperienced companies working on similar solutions high TRL with subcontractors	■	~	~	~		■	■	4
Old, big, experienced companies working on similar existing high TRL solutions with subcontractors	■	■	■	■		■	■	8
Not big inexperienced companies working on similar existing high TRL solutions with subcontractors		~	■	~	■	■	■	7
Not old not big inexperienced companies working on new innovative low TRL solutions with subcontractors	~	~	~	~	~	■	~	3
Old not big inexperienced companies working on similar innovative high TRL solutions without subcontractors	■	~	■	~	~	~	■	5

LESSONS LEARNT

SCOPING: more difference-making conditions in the model

LINKING: companies without clear advantage included

NAMING: limited interpretability

CALIBRATING: qualitative data needed for better understanding

Figure 3. Model 2a.
Note: Condition may be present (black), absent ("~") or not included in the pathway; N = 82, Consistency = 0.90, Coverage = 0.49. The relatively low coverage was a result of including in the model only configurations of conditions that were represented by at least two companies.

companies, the outcome had not yet been observed but was likely to be present in the near future. As for the conditions, we learned that factors such as the company's orientation toward the subsidy (i.e. strategic importance of the project or prospective revenue for the company), company's knowledge of the target market, and contextual factors (e.g. market disruptions) were also relevant for cases not explained by the model. Notably, some companies that did not have an identified advantage (i.e. inexperienced microenterprises) did achieve success. In other words, even if the model was correct, it did not paint the whole picture and it was worth developing it further in the next step of the evaluation; this was also necessary to account for the cases that were not (yet) explained.

Model 2a

With Model 2a, these limitations could be partially overcome (Figure 3). Some paths showed companies having a previous advantage (e.g. bigger, older, and more experienced) but there were also groups of companies for which these conditions were absent. This observation was aligned with the heuristic, "Think equifinally" (i.e. about multiple configurations that may lead to an outcome).

While the model was built on secondary data, the interpretability and understandability of its paths were limited. Without enough familiarity with the cases, it was difficult to make causal claims about the configurations of conditions that turned out to be sufficient for success. For example, having a subcontractor seemed to play an important role; however, the nature and extent of this condition's role in the configurations remained unclear. It was also unclear whether the presence of this condition was an artifact obscuring another important condition that was not included in the model. Thus, despite the advances and progression in nuance made in Model 2a compared with that in Model 1, the picture remained relatively blurred and merited more follow-up investigation.

Model 2b

Model 2b combined the strengths of its predecessors. It was saturated with data from diverse sources and it included the whole spectrum of conditions that were found to be relevant in the previous models. The additional conditions introduced the challenge of having too many conditions, however, and thus the risk of violating a good ratio between cases and outcomes.

As often happens, many alternative models were possible that could have explained the outcome and would have had satisfying parameters. Configurational theorizing was very instrumental at this point. First, macroconditions—conditions that are merged with more conditions—were created (heuristic: "Simplify to higher order constructs"). Second, conditions were categorized (heuristic: "Complexify from an anchor") into causal triggers and auxiliary conditions. Three company conditions were conceived as causal triggers: access to inspiring resources, that is, being a large company or linked to another enterprise and being involved in informal cooperation with other R&D units or interactions with scientists (RES), strategic importance of the project or the revenue from this project (STR), and expanding existing business (BAU). Three other conditions served as auxiliary: experience in R&D (EXP), formal cooperation (FORM), and no radical changes in context (CON). Their configurations explained the success of the R&D subsidy, that is, the company received income from their introduction of the subsidized project to the market. Third, one of the heuristics related to the linking stage—"think about absence"—was inspirational in that it alerted us to investigate how the absence of a condition is connected to the presence of another condition. In Model 2b, the absence of EXP is clearly connected with FORM. Companies that did not have much experience but engaged in formal cooperation were successful. Therefore, experience and formal cooperation seemed to serve as alternative conditions for success.[2]

The seven paths leading to the outcome were aggregated into two groups of beneficiaries at the naming stage of configurational theorizing (heuristic: "Capture the whole"). *Motivated learners* were the companies that did not have enough experience but based their success either on their strategy or on inspiration from formal cooperation with academic institutions. They were aware of their deficits and made up for them by engaging in formal cooperation with other partners and learning through the process. Alternatively, *pragmatic entrepreneurs* took advantage of their relationships with existing clients and their experience with comparable products. They usually had the necessary R&D experience.

The final model (Figure 4) goes beyond the previously explained Matthew effect. Even if the initial advantage plays an important role, it may be built on the diverse foundations of RES, STR, and BAU. R&D experience may serve as an additional advantage;[3] however, it

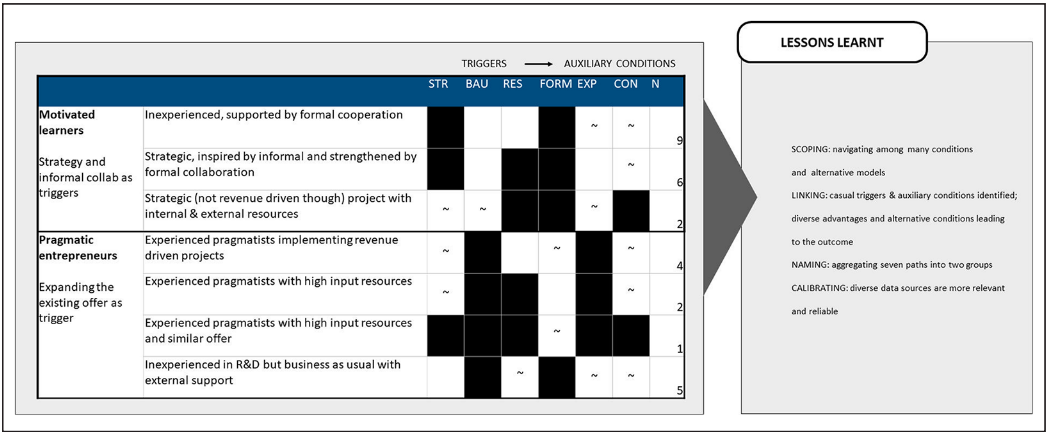

Figure 4. Model 2b.
Note: Condition may be present (black), absent ("~") or not included in the pathway; N = 35, Consistency = 0.96, Coverage = 1.

may be effectively substituted by formal cooperation with other partners who have the necessary resources or capabilities.

Importantly, the process was far from linear, even when additional knowledge was accumulated and fine-tuned at each stage. It is difficult to convincingly argue that Model 2a was unambiguously better than Model 1, but it was decidedly an important step in the overall process.

Concluding discussion

This article illustrates the potential of configurational theorizing and consecutive QCA for cumulative knowledge-building and, particularly, for theory development. Compared with the results of the first study, the final findings were much more nuanced, and helped us develop a more fine-grained configurational program theory. Moreover, conducting QCA in a consecutive manner helped us deal with the methodological challenges typically present during the process: both handling multiple conditions and not having all the necessary data at some stages of the research. The final model proved much better than its predecessors, especially given that the insights gathered along the way related to the calibration of the outcome and conditions.

Our article makes two contributions to the literature on the use of QCA in evaluation studies. First, it demonstrates how *configurational theorizing* may be instrumental in program theory development. As demonstrated by our case example, configurational theorizing and the associated relevant heuristics (Furnari al., 2021) provided a useful instrument for guiding our approach, and other evaluators may wish to adopt the same strategy. In addition, we believe that configurational theorizing has potential for the evaluation community at large, irrespective of the methods being used. It can, for example, be a useful vehicle for articulating program theories at the first stages of the evaluation process such as during workshops with stakeholders.

Second, the article shows how *consecutive QCA* is remarkably compatible with configurational theorizing and has the potential to help evaluators substantially advance program theories when confronted with program interventions that are characterized by conjunctural complexity. Consecutive QCA lends itself particularly well to any series of project or program evaluations, that are conducted at different moments in time, for which the results of the preceding QCA can inform the next one. Together, as we explain in our definition of consecutive QCA, the series of studies should have an overarching research goal, but can have different samples and specific research objectives.

Configurational theorizing is as demanding as any other theory development process. Researchers are faced with many consequential decisions. However, its configurational language, existing good practices, and heuristics make the decisions transparent and replicable. Consecutive QCA is also challenging; it self-evidently requires more resources (finances, data, and time) than when a study is restricted to one QCA cycle. Thus, evaluation commissioners should carefully weigh the benefits and costs prior to implementing such designs.

In addition to addressing the purpose of overall knowledge cumulation, we highlight other reasons for engaging in consecutive QCA. First, two QCAs may be conducted for different purposes. For example, researchers may analyze a larger sample of cases and choose some of these for more in-depth analysis. Then, QCA may be conducted on the smaller sample of more closely investigated cases (Models 2a and 2b in this article). The second model will have more conditions, added on the basis of the in-depth analysis. Different research objectives may also be related to different stages of the policy and evaluation cycle. Researchers may want to conduct an exploratory QCA analysis in a mid-term evaluation (Model 1) and then engage in fully fledged theory development in an ex-post evaluation (Models 2a and 2b). Second, after conducting an exploratory QCA study and reporting it to stakeholders, researchers may learn that the calibration of conditions or the outcome should be altered. That was the case after Model 1 was built. The lessons learned were used in the next QCA study. Finally, researchers may want to conduct robustness checks of their models for different cases. A robustness check is important whenever researchers want to make generalizations, which is often the case in evaluation studies. Even if QCA results cannot be generalized in a statistical sense, some modest generalization claims are feasible under certain conditions. It is our hope that future work will evaluate our model with other R&D subsidies.

Ideally, researchers want their consecutive QCA to be both fully coherent (i.e. have the same calibration of conditions and outcome for comparison's sake) and flexible (i.e. use the best available calibration for the research stage and data at hand). During the actual research process, however, researchers might encounter a trade-off between coherence and flexibility. We consider this to be a primary challenge when employing consecutive QCA—or at least, that was our experience. Addressing this trade-off requires a precise delineation of the objectives behind the consecutive QCA. This trade-off can only be addressed by being clear about the objectives of the consecutive QCA. On the one hand, researchers applying consecutive QCA at different stages of the research process should be more flexible in their analysis, as should those who want to go beyond the limitations discovered in their earlier study. On the other hand, coherence should be a priority during robustness checks.

The next challenge researchers applying consecutive QCA will likely encounter is how to deal with differences between models that are built along the way. If Condition A is present in the first model but is recalibrated in a second model with different cases, it may play a different role in configurations covered by the model. Researchers are expected to determine whether these differences arise from weak robustness in the earlier model, different scope conditions and contexts, or a different calibration of conditions or outcome. While researchers dealing with this challenge do not yet have clear guidelines at their disposal, they can rely on a growing literature body with various strategies to incorporate the time dimension into the analysis (Caren and Panofsky, 2005; Pagliarin and Gerrits, 2020; Ragin and Fiss, 2017; Verweij and Vis, 2021), robustness checks (Oana and Schneider, 2021), theory evaluation (Oana and Schneider, 2018), and the role of context and scope conditions (Falleti and Lynch, 2009). All these resources can help in this endeavor. Building cumulative knowledge is, thus, also at stake when learning about consecutive QCA designs themselves, and we believe our article can usefully contribute to these challenges.

Acknowledgements

The authors are grateful to the editor, as well as the two anonymous reviewers for their guidance and comments during the review process. The authors thank for the valuable suggestions and comments from the experts during 5th International QCA Paper Development Workshop, 1st Annual QCA Conference of Americas, and 10th International QCA Expert Workshop, especially to Tim Haesebrouck and Claude Rubinson. The authors express their gratitude to all their interviewees and experts from Section for Analyses and Evaluation, National Centre for Research and Development in Poland for their cooperation. The authors also thank Anita Harman, PhD, from Edanz (www.edanz.com/ac) for editing a draft of this manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This publication has been funded by Polish National Science Center (project 2019/35/B/HS5/04238), and from the budget under the program "Excellence Initiative—Research University" at the Jagiellonian University.

ORCID iDs

Seweryn Krupnik https://orcid.org/0000-0003-2486-0702
Valérie Pattyn https://orcid.org/0000-0001-7844-8858

Supplemental material

Supplemental material for this article is available online.

Notes

1. The literature review included 69 papers that investigated the effects of R&D subsidies on R&D expenditure and income from innovations' sale on the firm- or plant-level data (published between 2010 and 2020 in reviewed English-language journals available in the *Web of Science's* core collection).
2. This insight also helps us understand the role of condition SUB in the Model 2a, which was discussed above. The condition "Subcontractor in the project" evolved into "Formal cooperation." Thanks to the insights gathered in the interviews, the condition was more precisely defined and became more easily interpretable.
3. Experience in R&D was categorized as an auxiliary condition and not a causal trigger on the basis of the insights gathered during the interviews.

References

Álamos-Concha P, Cambré B, Pattyn V, et al. (2020) What drives training transfer effectiveness and how does this transfer work? Antwerp. Available at: https://publicaties.vlaanderen.be/view-file/40312

Antonelli C and Crespi F (2013) The "Matthew effect" in R and D public subsidies: The Italian evidence. *Technological Forecasting and Social Change* 80(8): 1523–34.

Astbury B and Leeuw FL (2010) Unpacking Black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation* 31(3): 363–81.

Balthasar A (2006) The effects of institutional design on the utilization of evaluation: Evidenced using qualitative comparative analysis (QCA). *Evaluation* 12(3): 353–71.

Befani B and Sager F (2006) QCA as a tool for realistic evaluations. In: Rihoux B and Grimm H (eds) *Innovative Comparative Methods for Policy Analysis*. New York: Springer, 263–284.

Befani B (2013) Between complexity and generalization: Addressing evaluation challenges with QCA. *Evaluation* 19(3): 269–83.

Befani B (2016) Pathways to change: Evaluating development interventions with qualitative comparative analysis (QCA). Rapport till Expertgruppen För Biståndsanalys (EBA). Available at: http://eba.se/wp-content/uploads/2016/07/QCA_BarbaraBefani-201605.pdf

Befani B (2020) Diagnostic evaluation and Bayesian updating: Practical solutions to common problems. *Evaluation* 26(4): 499–515.

Berg-Schlosser D and De Meur G (2009) Comparative research design: Case and variable selection. In: Rihoux B and Ragin CC (eds) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Los Angeles, CA: SAGE, 19–32.

Blackman T, Wistow J and Byrne D (2013) Using qualitative comparative analysis to understand complex policy problems. *Evaluation* 19(2): 126–40.

Caren N and Panofsky A (2005) TQCA: A technique for adding temporality to qualitative comparative analysis. *Sociological Methods & Research* 34(2): 147–72.

Clausen TH (2009) Do subsidies have positive impacts on R&D and innovation activities at the firm level? *Structural Change and Economic Dynamics* 20(4): 239–53.

Coryn CLS, Noakes LA, Westine CD, et al. (2011) A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation* 32(2): 199–226.

Cunningham P, Gök A and Larédo P (2016) The impact of direct support to R&D and innovation in firms. In: Edler J, Cunningham P, Abdullah G, et al. (eds) *Handbook of Innovation Policy Impact*. Cheltenham: Edward Elgar, 54–107.

Czarnitzki D and Lopes-Bento C (2013) Value for money? New microeconometric evidence on public R&D grants in Flanders. *Research Policy* 42(1): 76–89.

David PA, Hall BH and Toole AA (2000) Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy* 29(4–5): 497–529.

Dimos C and Pugh G (2016) The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature. *Research Policy* 45(4): 797–815.

Duşa A (2018) *QCA with R: A Comprehensive Resource*. New York: Springer.

Falleti TG and Lynch JF (2009) Context and Causal Analysis. *Comparative Political Studies* 42(9): 1143–66.

Fiss PC, Marx A and Cambre B (2013) *Configurational Theory and Methods in Organizational Research: Introduction*. Bingley: Emerald Group.

Furnari S, Crilly D, Misangyi VF, et al. (2021) Capturing causal complexity: Heuristics for configurational theorizing. *Academy of Management Review* 46(4): 778–99.

Jenkins-Smith HC and Sabatier P (1999) The advocacy coalition framework: An assessment. In: Sabatier P (ed.) *Theories of the Policy Process*. Boulder, CO: Westview Press, 117–66.

Kahwati LC and Kane HL (2020) *Qualitative Comparative Analysis in Mixed Methods Research and Evaluation*. Thousand Oaks, CA: SAGE.

Krupnik S (2012) Evaluating the social mechanism of investment subsidies using an abductive approach. *Evaluation* 18(4): 466–76.

Krupnik S and Koniewski M (2022) Choosing a qualitative comparative analysis solution in multi-method impact evaluation. *Evaluation* 28(2): 192–209.

Lemire S, Peck LR and Porowski A (2020) The growth of the evaluation tree in the policy analysis forest: Recent developments in evaluation. *Policy Studies Journal* 48(S1): S47–70.

Merton RK (1968) The Matthew effect in science: The reward and communication systems of science are considered. *Science* 159(3810): 56–63.

Ministry of Development Funds and Regional Policy (2022) Szczegółowy Opis Priorytetów Programu Operacyjnego Inteligentny Wzrost [Detailed description of operational program smart growth's priorities]. *Warsaw*. Available at: https://www.poir.gov.pl/media/113093/SZOOP_POIR_29.pdf

Ministry of Regional Development in Poland (2012) Szczegółowy Opis Priorytetów Programu Operacyjnego Innowacyjna Gospodarka [Detailed description of operational program innovative economy's priorities]. Available at: https://www.poir.gov.pl/media/92011/SZOP_wersja_20_zaakceptowana.pdf

Mouque D (2012) What are counterfactual impact evaluations teaching us about enterprise and innovation support? DG for Regional and Urban Policy. Available at: https://www.capire.org/capire-informa/scaffale/2012_02_counterfactual.pdf

Oana I-E and Schneider CQ (2018) SetMethods: An add-on R package for advanced QCA. *The R Journal* 10(1): 507–33.

Oana I-E and Schneider CQ (2021) A robustness test protocol for applied QCA: Theory and R software application. *Sociological Methods & Research*. Epub ahead of print 26 August. DOI: 10.1177/00491241211036158.

Pagliarin S and Gerrits L (2020) Trajectory-based qualitative comparative analysis: Accounting for case-based time dynamics. *Methodological Innovations* 13(3): 1–11.

Pattyn V (2014) Why organizations (do not) evaluate? Explaining evaluation activity through the lens of configurational comparative methods. *Evaluation* 20(3): 348–67.

Pattyn V, Molenveld A and Befani B (2019) Qualitative comparative analysis as an evaluation tool: Lessons from an application in development cooperation. *American Journal of Evaluation* 40(1): 55–74.

Pawson R and Tilley N (1997) *Realitistic Evaluation*. London: SAGE.

Pereira AM and Suarez D (2017) Matthew effect, capabilities and innovation policy: The Argentinean case. *Economics of Innovation and New Technology* 27(1): 62–79.

Petrin T (2018) A literature review on the impact and effectiveness of government support for R&D and innovation. Available at: http://www.isigrowth.eu/wp-content/uploads/2018/02/working_paper_2018_05.pdf

Ragin CC (2017) User's guide to fuzzy-set/qualitative comparative analysis. Manual based on fsQCA 3.0. Available at: https://www.socsci.uci.edu/~cragin/fsQCA/software.shtml

Ragin CC and Fiss PC (2017) *Intersectional Inequality: Race, Class, Test Scores, and Poverty*. Chicago, IL: The University of Chicago Press.

Ragin CC and Rubinson C (2009) The distinctiveness of comparative research. In: Landman T and Robinson N (eds) *The SAGE Handbook of Comparative Politics*. London: SAGE, 13–34.

Rihoux B, Rezsöhazy I and Bol D (2011) Qualitative comparative analysis (QCA) in public policy analysis: An extensive review. *German Policy Studies* 7(3): 9–82.

Rogers PJ, Petrosino A, Huebner TA, et al. (2000) Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation* 87: 5–13.

Rubinson C, Rutten R and Greckhamer T (2019) Avoiding common errors in QCA: A short guide for new practitioners. Available at: https://compasss.org/wp-content/uploads/2019/07/Common_Errors_in_QCA.pdf

Schneider CQ and Wagemann C (2012) *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

Stern E, Stame N, Mayne J, et al. (2012) *Broadening the range of designs and methods for impact evaluations*. DFID Working Paper 38. London: UK Department for International Development.

Thomann E (2020) Qualitative comparative analysis for comparative policy analysis. In: Peters G and Fontaine G (eds) *Handbook of Research Methods and Applications in Comparative Policy Analysis*. Cheltenham: Edward Elgar, 254–76.

Thomann E and Ege J (2020) Qualitative comparative analysis (QCA) in public administration. In: *Oxford Research Encyclopedia of Politics*. Available at: https://oxfordre.com/politics/display/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1444

Thomann E and Maggetti M (2017) Designing research with qualitative comparative analysis (QCA): Approaches, challenges, and tools. *Sociological Methods & Research* 49(2): 356–86.

Thomann E, Ege J and Paustyan E (2022) Approaches to qualitative comparative analysis and good practices: A systematic review. *Swiss Political Science Review* 28: 557–80.

Turnbull B (2002) Program theory building: A strategy for deriving cumulative evaluation knowledge. *American Journal of Evaluation* 23(3): 275–90.

Turner K, Jepson R, MacDonald B, et al. (2018) Developing and refining a programme theory for understanding how twenty mile per hour speed limits impact health. *Journal of Transport & Health* 10: 92–110.

Vellema S, Ton G, De Roo N, et al. (2013) Value chains, partnerships and development: Using case studies to refine programme theories. *Evaluation* 19(3): 304–20.

Verweij S and Gerrits L (2013) Understanding and researching complexity with qualitative comparative analysis: Evaluating transportation infrastructure projects. *Evaluation* 19(1): 40–55.

Verweij S and Vis B (2021) Three strategies to track configurations over time with qualitative comparative analysis. *European Political Science Review* 13(1): 95–111.

Wagemann C and Schneider CQ (2015) Transparency standards in qualitative comparative analysis. *Qualitative & Multi-Method Research* 13(1): 38–42.

Seweryn Krupnik holds the position of an Assistant Professor at the Jagiellonian University (JU), Poland. He is also a Research Fellow at the Center for Evaluation and Analysis of Public Policies of the JU.

Anna Szczucka is a Research Fellow at the Center for Evaluation and Analysis of Public Policies of the Jagiellonian University and a Researcher at University of Science and Technology in Krakow, Poland.

Monika Woźniak is a Ph.D. candidate at the Jagiellonian University in Kraków, Poland. She specialises in transformative innovation policies and evaluation of public programmes.

Valérie Pattyn is an Associate Professor at the Institute of Public Administration of Leiden University (The Netherlands) and is partially affiliated with KU Leuven Public Governance Institute (Belgium).