Political Methodology

Committee on Concepts and Methods

Working Paper Series

34

June 2012

Unifying Configurational Comparative Methodology

Generalized-Set Qualitative Comparative Analysis

Alrik Thiem

Swiss Federal Institute of Technology Zurich

(thiem@sipo.gess.ethz.ch)

C&M
The Committee on Concepts and Methods
www.concepts-methods.org

IPSA
International Political Science Association
www.ipsa.org

CIDE
Research and Teaching in the Social Sciences
www.cide.edu

The C&M working paper series are published by the Committee on Concepts and Methods (C&M), the Research Committee No. 1 of the International Political Science Association (IPSA), hosted at CIDE in Mexico City. C&M working papers are meant to share work in progress in a timely way before formal publication. Authors bear full responsibility for the content of their contributions.

The Committee on Concepts and Methods (C&M) promotes conceptual and methodological discussion in political science. It provides a forum of debate between methodological schools who otherwise tend to conduct their deliberations at separate tables. It publishes two series of working papers: "Political Concepts" and "Political Methodology."

Political Concepts contains work of excellence on political concepts and political language. It seeks to include innovative contributions to concept analysis, language usage, concept operationalization, and measurement.

Political Methodology contains work of excellence on methods and methodology in the study of politics. It invites innovative work on fundamental questions of research design, the construction and evaluation of empirical evidence, theory building and theory testing. The series welcomes, and hopes to foster, contributions that cut across conventional methodological divides, as between quantitative and qualitative methods, or between interpretative and observational approaches.

Submissions. All papers are subject to review by either a member of the Editorial Board or an external reviewer. Only English-language papers can be admitted. Authors interested in including their work in the C&M Series send their paper to wp[at]concepts-methods.org. They may also seek initial endorsement by one editorial board member.

The C&M webpage offers full access to past working papers.

www.concepts-methods.org

Introduction

Qualitative Comparative Analysis (QCA) is becoming increasingly popular with social scientists across disparate areas of research. Its suitability for analyzing set-theoretic instead of correlational hypotheses has been one of the main reasons for its proliferation. In particular, applications from sociology (e.g., Amenta and Halfmann, 2000; Brown and Boswell, 1995; Dixon, Roscigno and Hodson, 2004; Hodson, Roscigno and Lopez, 2006) and political science (e.g., Avdagic, 2010; Berg-Schlosser, 2008; Lilliefeldt, 2012; Thiem, 2011) continue to top the list of publication figures (Thiem and Duşa, 2012*b*). For about ten years now, however, QCA has also diversified into a set of distinct variants. Textbooks and applied research now differentiate between crisp-set QCA (csQCA) (Rihoux and De Meur, 2009), multi-value QCA (mvQCA) (Cronqvist and Berg-Schlosser, 2009) and fuzzy-set QCA (fsQCA) (Ragin, 2009). Each of these variants is associated with a specific set-data type - crisp sets for csQCA, multi-value sets for mvQCA and fuzzy sets for fsQCA - whose analysis has required different software with tailored routines. Until recently, the fs/QCA software (Ragin and Davey, 2009) has been the sole option for fsQCA and mvQCA could only be handled by Tosmana (Cronqvist, 2011), but either software has been capable of processing csQCA from the beginning. More recently, the fuzzy package for STATA (Longest and Vaisey, 2008) has been introduced as an alternative to fs/QCA and the two R packages QCA (Duşa and Thiem, 2012) and QCA3 (Huang, 2012) now also offer extensive functionality for all three QCA variants. Although software functionality has differed considerably, end-users have almost always been able to carry out all the steps which their analyses have necessitated.

Despite many methodological developments over the last decade, however, one considerable problem has hitherto been left unaddressed. If the data did not fit the set types associated with any of its three variants, QCA could either not be employed at all or researchers were forced to accept a loss of information by recalibrating their sets into a processable format. Most significantly, while data types for mvQCA and csQCA could be reconciled in Tosmana and those for fsQCA and csQCA in fs/QCA, multi-value and fuzzy sets have been incongruous. The theoretical dissolution of this incompatibility is the secondary objective of this paper.

The primary objective pushes the argument further. It is to be demonstrated that all three QCA variants can in fact be treated as special cases of a more general variant, which shall be referred to as generalized-set QCA (gsQCA). At the core of gsQCA is the multi-level fuzzy set, a set-data type which not only combines the three existing ones under a single framework, but also allows to retain the established truth table construction and Boolean minimization procedures. In this connection, it is also to be illustrated that neither does mvQCA offer a middle way between csQCA and fsQCA (Herrmann and Cronqvist, 2009), nor is the most serious point of criticism raised against it by Schneider and Wagemann (2012) as well as Vink and van Vliet (2009) methodologically justified (Thiem, forthcoming). mvQCA fits squarely into the set-theoretic approach of its two sister variants.

The argument is structured around three main sections. Firstly, the debate on the commonalities and differences between csQCA, mvQCA and fsQCA is reviewed. The second section then introduces an approach to combining the existing crisp-set, multi-value and fuzzy-set data types in a single analysis. Finally, the third section generalizes the method illustrated in the preceding section by introducing gsQCA. This consecutive approach is taken for a simple reason. The combination of all existing set-data types in a single analysis is novel, but still adheres to a set of common rules within which current set-theoretic methodology operates, such as the single-membership principle.[1] This will not hold true any more in gsQCA. The conclusions recapitulate the argument.

The State of the Debate

Crisp-set QCA (csQCA), introduced to a wider audience of social scientists by Ragin (1987), is the first of the three variants to have appeared in the toolbox of comparative social science methodology.[2] Until the publication of Ragin (2000), csQCA was still referred to as QCA because back then no ambiguities as to what the acronym denoted had existed. With twenty applications in published journal articles in 2011, it remains the most-often applied variant (Thiem and Duşa, 2012*a*).

The roots of csQCA lie in electrical engineering, where algorithms have been developed for reducing switching circuits to their most economical form. Analogously, Ragin's intention was to simplify complex social-scientific data in a holistic manner using the switching algebra of electrical engineering. Switching algebra, which is just one branch of Boolean algebra, is fundamentally different in some important respects from the linear algebra that underlies the vast majority of statistical models which are usually applied in social science research. Many special cases of a Boolean algebra exist, such as propositional logic, switching algebra and set algebra, but any set $\mathcal{S}$ of elements $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \ldots\}$ and two binary operations $+$ (Boolean sum) and $\cdot$ (Boolean product) form a Boolean algebra if, and only if, the operations are commutative:

$$\mathbf{S}_1 + \mathbf{S}_2 = \mathbf{S}_2 + \mathbf{S}_1 \text{ and } \mathbf{S}_1 \cdot \mathbf{S}_2 = \mathbf{S}_2 \cdot \mathbf{S}_1; \tag{P1}$$

each of the operations distributes over the other:

$$\mathbf{S}_1 \cdot (\mathbf{S}_2 + \mathbf{S}_3) = \mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 \text{ and } \mathbf{S}_1 + \mathbf{S}_2 \cdot \mathbf{S}_3 = (\mathbf{S}_1 + \mathbf{S}_2) \cdot (\mathbf{S}_1 + \mathbf{S}_3); \tag{P2}$$

identity elements 0 and 1 exist for + and · :

$$0 + \mathbf{S}_1 = \mathbf{S}_1 \text{ and } 1 \cdot \mathbf{S}_1 = \mathbf{S}_1; \tag{P3}$$

and for each $\mathbf{S}_1 \in \mathcal{S}$ there exists an $\mathbf{s}_1 \in \mathcal{S}$ (complement) such that:

$$\mathbf{S}_1 + \mathbf{s}_1 = 1 \text{ and } \mathbf{S}_1 \cdot \mathbf{s}_1 = 0. \tag{P4}$$

Most notably, Boolean-algebraic distributivity is false in linear algebra. Postulate (P4) relates to the defining characteristic of csQCA in that cases can only have either full membership in a set or none at all. Expressed in set-algebraic notation, the union $\mathbf{S}_1 \cup \mathbf{s}_1$ of literals $\mathbf{S}_1$ and $\mathbf{s}_1$ of some set $\mathbf{S}_1$ form the universal set $\mathbf{U}$, whereas their intersection $\mathbf{S}_1 \cap \mathbf{s}_1$ the empty set $\varnothing$. From these four postulates, a number of theorems can be derived, the two most important of which in the context of the set-algebraic logic of QCA lead to the elimination of as many sets as possible from Boolean product terms and the elimination of as many terms as possible from a Boolean sum of product terms. More precisely, for any sets $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$ that form part of a Boolean algebra, $\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{s}_1 \cdot \mathbf{S}_2 = \mathbf{S}_2$ and $\mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{s}_1 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_3 = \mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{s}_1 \cdot \mathbf{S}_3$ (cf. McCluskey, 1965: 114ff.).

If $\mathbf{U}$ represents the universal set with any subset denoted by $\mathbf{S}$, $\mathbf{S}$ can itself be represented by its characteristic function $\mu_{\mathbf{S}}$ as given in Equation (5). This function describes some general property shared by all subsets of $\mathbf{U}$ (Klir, St. Clair and Yuan, 1997: 63).

$$\mu_{\mathbf{S}}(u) = \begin{cases} 1 & \text{if } u \in \mathbf{S} \\ 0 & \text{if } u \notin \mathbf{S} \end{cases} \tag{5}$$

If the Boolean-algebraic system is restricted to the two values 0 and 1, these numbers may not only represent symbolic indicators of membership and non-membership in $\mathbf{S}$, but also numerical indicators of the degree to which some case $u_i \in \mathbf{U}$ also belongs to $\mathbf{S}$. Mapping cases into the binary set $\mathbb{B} = \{0, 1\}$ by means of characteristic functions therefore allows the representation of superset and subset relations as functional inequalities of the form $\mu_{\mathbf{S}_1}(u) \leq \mu_{\mathbf{S}_2}(u)$.

Operations on sets can also be written as operations on their characteristic function.

For example, the characteristic function of the union of sets $\mathbf{S}_1$ and $\mathbf{S}_2$ can be expressed by means of the parallel maximum of the individual characteristic functions as given in Equation (6).

$$\mu_{\mathbf{S}_1 \cup \mathbf{S}_2}(u) = \max\left[\mu_{\mathbf{S}_1}(u), \mu_{\mathbf{S}_2}(u)\right] \tag{6}$$

Analogously, the characteristic function of the intersection of $\mathbf{S}_1$ and $\mathbf{S}_2$ can be constructed using the parallel minimum of the individual characteristic functions as given in Equation (7).

$$\mu_{\mathbf{S}_1 \cap \mathbf{S}_2}(u) = \min\left[\mu_{\mathbf{S}_1}(u), \mu_{\mathbf{S}_2}(u)\right] \tag{7}$$

The list of characteristic functions involving the conjunction of all literals of the relevant subsets $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \ldots\} \in \mathbf{U}$ and a summary statement about the outcome value of each element in this list is called a truth table (cf. Ragin, 1987: 87). There exist $d = \prod_{j=1}^{k} p_j$ characteristic functions of p-valued k input sets. These functions are more commonly referred to in QCA as configurations. In the social sciences, configurations represent exhaustive combinations of properties characterizing units of analysis, such as people, organizations or states. For illustration, a hypothetical truth table with three input sets $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$, also called conditions, and their corresponding outcome value OUT (also called truth value) is presented in Table 1. Three binary-value conditions yield $2^3 = 8$ configurations, each of which is indexed under $\mathcal{C}_a$. The minimum number of cases n that is usually required for the respective outcome value is given in addition, but it is no essential part of the truth table.

Table 1: Hypothetical Truth Table

$\mathcal{C}_a$	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_3$	OUT	n
1	1	1	1	1	≥ 1
2	1	1	0	1	≥ 1
3	1	0	1	1	≥ 1
4	1	0	0	1	≥ 1
5	0	1	1	0	≥ 1
6	0	1	0	C	≥ 2
7	0	0	1	?	0
8	0	0	0	?	0

A condition is the analogue of an independent variable, but the outcome value is not analogous to the dependent variable. The dependent variable is captured by the outcome set, which does not show up in the truth table. Instead, the outcome value represents a truth value indicating whether the evidence is consistent with a hypothesis about the existence of a subset relation between each configuration and the outcome set or not. Configurations 1-4 are true and support this hypothesis (OUT = 1), configuration 5 is false and does not (OUT = 0). Mixed evidence exists for configuration 6 (OUT = C). If at least two cases conform to one configuration, but the evidence neither sufficiently supports nor confutes the hypothesis, contradictions arise. No empirical evidence at all exists for configurations 7 and 8 (OUT = ?). If a particular configuration contains no or too few cases, it is called a logical remainder.

The canonical (Boolean) sum f_1 resulting from the truth table presented in Table 1 is given by Equation (8). It consists of four fundamental (Boolean) products, each of which corresponds to a true configuration from Table 1.

$$f = \overbrace{\mathbf{C}_1 \cdot \mathbf{C}_2 \cdot \mathbf{C}_3}^{\mathcal{C}_1} + \overbrace{\mathbf{C}_1 \cdot \mathbf{C}_2 \cdot \mathbf{c}_3}^{\mathcal{C}_2} + \overbrace{\mathbf{C}_1 \cdot \mathbf{c}_2 \cdot \mathbf{C}_3}^{\mathcal{C}_3} + \overbrace{\mathbf{C}_1 \cdot \mathbf{c}_2 \cdot \mathbf{c}_3}^{\mathcal{C}_4} \tag{8}$$

If two such products differ on only one condition, then this condition is redundant and can be eliminated so that a simpler term results. Equation (8) can be reduced in two passes as shown in Figure 1. In the first pass, the four fundamental products with three literals can be reduced to four simpler terms with just two literals. In the second pass, these four terms can then be reduced at once to a single term with just one literal. No further reduction is possible. Condition $\mathbf{C}_1$ is the only factor which is essential in explaining the outcome ($\mathbf{C}_1 \rightarrow \mathbf{O}$).

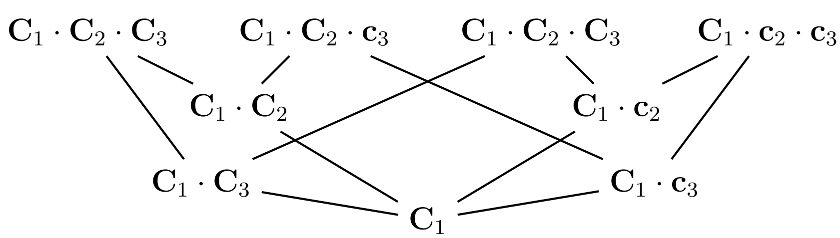

Figure 1: Boolean minimization of Equation (8).

While the logic of classical crisp sets forms the basis of many mathematical concepts, a considerable number of social-scientific concepts do not neatly fit into the binary logic of a Boolean algebra. In order to address this shortcoming, Zadeh (1965) proposed the concept of the fuzzy set, which is at the core of the fsQCA variant. The most important feature of fuzzy set theory is the invalidity of the fourth Boolean-algebraic postulate (P4). In contrast to membership in a crisp set, membership in a fuzzy set is not limited to the binary set $\mathbb{B} = \{0, 1\}$ but extends to the unit interval $\mathbb{U} = [0, 1]$. With fuzzy sets, it becomes possible for a case to have non-zero membership in **S** and its complement **s** at the same time. In this sense, a crisp set represents a special case of a fuzzy set (Clark et al., 2008: 57).

The fsQCA analogue of the characteristic function in csQCA is the membership function. As the only restriction on membership functions is a unit interval co-domain, they can take an infinite number of forms. For example, the membership function suggested by Ragin (2008) and used in the fs/QCA software (Ragin and Davey, 2009) is given in Equation (9), where τ_e denotes the threshold for full exclusion from **S**, τ_c the crossover threshold at the point of maximally ambiguous set membership in **S** and τ_i the threshold for full inclusion in **S**.

$$\mu_{\mathbf{S}}(u) = \begin{cases} 1/\left(1 + \mathrm{e}^{-\left[(u-\tau_c)\left(\frac{-\log(19)}{\tau_e-\tau_c}\right)\right]}\right) & \text{if } u < \tau_c, \\ 1/\left(1 + \mathrm{e}^{-\left[(u-\tau_c)\left(\frac{\log(19)}{\tau_i-\tau_c}\right)\right]}\right) & \text{if } u \geq \tau_c \end{cases} \tag{9}$$

All variants of QCA require the construction of truth tables, but with fuzzy-set data, truth tables cannot be constructed directly because the possible number of configurations would be infinite. The truth table algorithm introduced in Ragin (2008) thus considers truth table configurations as the outer corners of a vector space. Although each case can have membership in one, several or all dimensions of this space, it can only be a strong instance, that is, have membership above 0.5, of one configuration. As usually more cases than only strong instances have some membership in a configuration, its outcome value depends on the partial configuration membership of all cases.

About four years after fsQCA had been introduced, Cronqvist (2004) presented the Tosmana software for processing multi-value sets. The first substantive application of mvQCA has come from Balthasar (2006). Since then, however, only six further articles have been published. This relatively low number may be explained by the fact that, so far, mvQCA has been viewed with a considerable degree of suspicion in the methodological literature (cf. Schneider and Wagemann, 2012; Vink and van Vliet, 2009). Why this suspicion is unjustified is explained in more detail elsewhere (Thiem, forthcoming), but mvQCA is a generalization of csQCA, albeit on a dimension that is different from the one which generalizes csQCA to fsQCA. How these dimensions relate to each other nonetheless will be fully elaborated in the remainder of this article.

After the introduction of fsQCA (Ragin, 2000) and mvQCA (Cronqvist and Berg-Schlosser, 2009; Cronqvist, 2011), the classical approach to configurational comparative thinking popularized by Ragin (1987) has come to be referred to as csQCA. Together, these three variants now make up the entire set of QCA techniques that have been used in more than 230 substantive applications published in academic journal articles since 1984 (Thiem and Duşa, 2012*b*: 3). The next section will demonstrate that 1) csQCA, mvQCA and fsQCA are closer related to each other than is usually acknowledged in the literature and 2) that mvQCA fits squarely into the framework of set-theoretic methodology.

Combining csQCA, mvQCA and fsQCA

This section demonstrates how csQCA, mvQCA and fsQCA can be combined in a single analysis. In order to reconcile the different notational systems and terminologies that have so far existed, three definitions are given first. The set data structure that has traditionally been referred to as a *crisp set* becomes a *binary-level crisp set*, a *multi-value set* a *multi-level crisp set*, and a *fuzzy set* a *binary-level fuzzy set*. Note that the prefix *binary-level* does not refer to the number of distinct states of set membership but the number of distinct states of extreme set membership. Even though membership in a fuzzy set is graded, it only has two extreme states for the concept that is represented by this set.

Now that the three basic set structures have been defined, a common notational system is required. The most serious point of criticism raised by Vink and van Vliet (2009), and echoed by others (Schneider and Wagemann, 2012), has been that mvQCA is difficult to understand in set-theoretic terms. Such difficulties do not arise from mvQCA as a method, but they are a consequence of the different systems of notation currently used in csQCA and fsQCA on the one hand, and mvQCA on the other. Traditionally, *membership notation* has been used for csQCA and fsQCA, but *category notation* for mvQCA. Parallels have often been drawn between the numbers used in these two systems, although they are not at all comparable, thereby leading to much confusion on the side of applied QCA users.

Membership and category-notation can be unified by putting them on a common denominator. The resulting system based on such a common denominator will be referred to as *level-score notation* in gsQCA. Level-score notation borrows from category notation in that curly brackets indicate distinct qualitative values. In addition, however, level-score notation appends to each level the set membership score of the case in the distinct level. When expressed in general terms, the set membership score s_i of a case i in category $\{v_l\}$ of set $\mathbf{S}_j$ is given by some membership function $\mu_{\mathbf{S}_j}(x_i\,\{v_l\})$, where $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, k$ and $l = 1, 2, \ldots, p$. The level $\{v_l\}$ can be designated by any symbol as long as it is a unique identifier within j. The expression $\{v_l\}\, s_i$ is then called a level-score term, consisting of a distinct category identifier and a membership score.

Table 2 presents ten hypothetical cases $i = 1, 2, \ldots, 10$, three conditions $\mathbf{C}_j$ with $j = 1, 2, 3$ and an outcome set $\mathbf{O}$. Condition $\mathbf{C}_1$ is a binary-level crisp set with two levels and $\mathbf{C}_2$ a multi-level crisp set with three levels. Their membership functions $\mu_{\mathbf{C}_1}(x_i\{0,1\})$ and $\mu_{\mathbf{C}_2}(x_i\{\alpha,\beta,\gamma\})$ have mapped the base variable values x_i of cases $i = 1, 2, \ldots, 10$ into the binary set $\mathbb{B} = \{0, 1\}$. Condition $\mathbf{C}_3$ as well as the outcome set $\mathbf{O}$, in contrast, are proper binary-level fuzzy sets. Their respective membership functions $\mu_{\mathbf{C}_3}(x_i\{0,1\})$ and $\mu_{\mathbf{O}}(x_i\{0,1\})$ have mapped the base variable values into the unit interval $\mathbb{U} = [0, 1]$. While membership functions play an important role in the set calibration process, their exact specification is irrelevant for the argument.

For example, the set membership score of $i = 1$ on condition set $\mathbf{C}_1$, level $v_2 = 1$, is given by $\mu_{\mathbf{C}_1}(x_1\{v_2\}) = \mathbf{C}_1\{v_2\}c_1 = \mathbf{C}_1\{1\}0$. In contrast, its set membership score on the same condition but level $v_1 = 0$, is given by $\mu_{\mathbf{C}_1}(x_1\{v_1\}) = \mathbf{C}_1\{v_1\}c_1 = \mathbf{C}_1\{0\}1$. The advantage of binary-level crisp set data is that membership scores for the other level need not be explicitly provided because with only two levels, the set membership score in the other level is always identified by Boolean negation. More precisely, for any $l = z$ of binary-level crisp set $\mathbf{S}$, $\mathbf{S}\{v_z\}s_i = 1 - \mathbf{S}\{v_{l \neq z}\}s_i$. In the case of $\mathbf{C}_1$, level $v_1 = 0$ thus carries the meaning of applying the logical NOT operator on $\mathbf{C}_1\{1\}$.[3]

With regards to the binary-level fuzzy set $\mathbf{C}_3$, the set membership score for $i = 1$, level $v_2 = 1$ is given by $\mu_{\mathbf{C}_3}(x_1\{v_2\}) = \mathbf{C}_3\{v_2\}c_1 = \mathbf{C}_3\{1\}0.4$, whereas its set membership score for level $v_1 = 0$ is given by $\mu_{\mathbf{C}_3}(x_1\{v_1\}) = \mathbf{C}_3\{v_1\}c_1 = \mathbf{C}_3\{0\}0.6$. As in the case of binary-level crisp sets, membership scores for the other level of a binary-level fuzzy set need not be explicitly provided in Table 2 because two levels allow full membership score identification by means of Boolean negation.

Table 2: Data Table of Set-Level and Configuration Membership Scores

	Set-Level Data				Configuration $\mathcal{C}_{a\,=\,\dots}$ from Table 3											
i	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_3$	$\mathbf{O}$	1	2	3	4	5	6	7	8	9	10	11	12
1	$\{1\}0$	$\{\alpha\}1$	$\{1\}0.4$	$\{1\}0.7$	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.4	0.0	0.0
2	$\{1\}0$	$\{\gamma\}1$	$\{1\}0.2$	$\{1\}0.1$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.2
3	$\{1\}1$	$\{\beta\}1$	$\{1\}0.8$	$\{1\}0.5$	0.0	0.8	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	$\{1\}1$	$\{\beta\}1$	$\{1\}1.0$	$\{1\}0.2$	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	$\{1\}1$	$\{\gamma\}1$	$\{1\}0.6$	$\{1\}1.0$	0.0	0.0	0.4	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0
6	$\{1\}0$	$\{\gamma\}1$	$\{1\}0.1$	$\{1\}0.9$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.1
7	$\{1\}1$	$\{\alpha\}1$	$\{1\}0.7$	$\{1\}0.3$	0.3	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	$\{1\}1$	$\{\alpha\}1$	$\{1\}0.9$	$\{1\}0.4$	0.1	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	$\{1\}0$	$\{\beta\}1$	$\{1\}1.0$	$\{1\}0.0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
10	$\{1\}1$	$\{\gamma\}1$	$\{1\}0.0$	$\{1\}1.0$	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
				$\lvert\mathcal{C}_a\rvert$	0.4	1.8	1.4	1.6	0.2	0.6	0.6	1.0	1.7	0.4	0.0	0.3
				$\lvert\mathcal{C}_a \cap \mathbf{O}\{1\}\rvert$	0.4	0.7	1.4	0.7	0.2	0.6	0.6	0.0	1.0	0.4	0.0	0.2
				$\lvert\mathcal{C}_a \cap \mathbf{O}\{0\}\rvert$	0.4	1.3	0.4	0.4	0.2	0.4	0.3	1.0	0.9	0.3	0.0	0.3

Finally, the set membership score for $i = 1$ on condition $\mathbf{C}_2$, level $v_1 = \alpha$, is given by $\mu_{\mathbf{C}_2}(x_1\{v_1\}) = \mathbf{C}_2\{v_1\}c_1 = \{\alpha\}1$. In contrast, its set membership score in levels $v_2 = \beta$ and $v_3 = \gamma$ is given by $\mu_{\mathbf{C}_2}(x_1\{\beta, \gamma\}) = \mathbf{C}_2\langle\{\beta\}c_1\{\gamma\}c_1\rangle = \mathbf{C}_2\langle\{\beta\}0\{\gamma\}0\rangle = \mathbf{C}_2\{\beta, \gamma\}0$.[4] No membership scores have to be provided for levels v_2 and v_3 because the multi-level set is not only crisp but its levels are also mutually exclusive. Unlike for binary-level sets, however, there is no other way of representation for $\mathbf{C}_2$ that is as efficient as that used in Table 2. The reason is simply that if it is known of which category a case is a member, it is also known that it cannot be a member of another category. More precisely, for any $l = z$, $\{v_z\}c_i = 1 - (\{v_1\}c_i + \{v_2\}c_i + \ldots + \{v_{l \neq z}\}c_i + \ldots + \{v_k\}c_i)$. The opposite logic does not hold because even if the category of which a case is not a member was identified, the information would be insufficient to determine the category of which it was a member. As a result, more information is needed and the representation of this information can never be more efficient than when the member category is used because exactly $p - 1$ level-score terms would be required.

In summary, if crisp or fuzzy sets are also binary-level, or if multi-level sets are also crisp, mutually exclusive and represented by membership-identifying level-score terms, all level-specific set membership scores are fully determined. With the common notational system of level-scores, a hierarchical representation of subset relations between these three different types of sets becomes possible. It is shown in Figure 2.

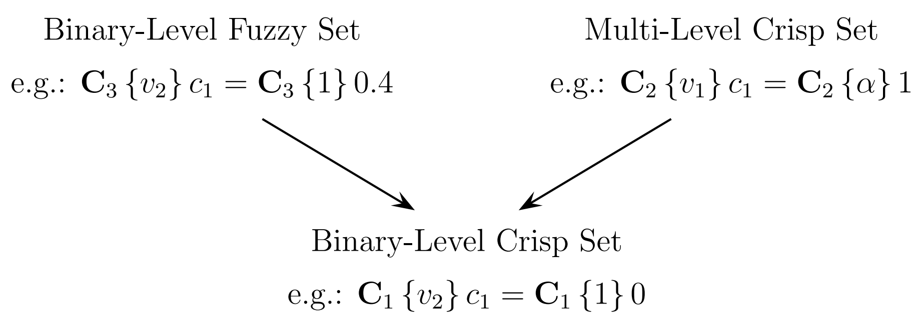

Figure 2: Hierarchy of Set-Data Types

Binary-level crisp sets are special cases of both binary-level fuzzy sets and multi-level crisp sets on different dimensions. Binary-level crisp sets share with multi-level crisp sets the characteristic of dichotomous membership scores - either 0 or 1, and they share with binary-level fuzzy sets the characteristic of having only two categories. Binary-level fuzzy sets and multi-level crisp sets, however, share no commonalities because each has a dimension that the other is not able to incorporate. With regards to the point of criticism raised by Vink and van Vliet (2009) that the status of mvQCA as a distinct set-theoretic method stands in doubt, level-score notation thus illustrates more clearly that mvQCA is, in fact, a generalization of csQCA and therefore a full-fledged member of the QCA family of comparative configurational techniques.

With complete information on conditions $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ and the outcome set $\mathbf{O}$, the truth table can be derived. It consists of $d = \prod_{j=1}^{k} p_j$ configurations, where p_j is again the total number of levels in set j. The complete truth table of the data presented in Table 2 is shown in Table 3. It lists all $d = 12$ configurations and the associated outcome value (OUT) for each level of $\mathbf{O}$. True configurations are coded as OUT = 1, false configurations as OUT = 0 and logical remainders as OUT = ?. A configuration's outcome value is determined by the number of cases with a set membership score of more than 0.5 in it (n) as well as its sufficiency inclusion score (Incl_S).[5] The membership scores of each case in each of the 12 configurations $\mathcal{C}_a$ with $a = 1, 2, \ldots, 12$ are shown in the right part of Table 2. Cells with a gray background indicate the highest membership of a case in the associated configuration that exceeds a value of 0.5. The scalar cardinality $|\mathcal{C}_a|$ of the configuration represented by $\mathcal{C}_a$ is calculated as given in Equation (10) and presented in Table 2.

$$|\mathcal{C}_a| = \sum_{i=1}^{n} \min[\mathbf{C}_1\{v_l\}c_i, \mathbf{C}_2\{v_l\}c_i, \mathbf{C}_3\{v_l\}c_i] \tag{10}$$

The scalar cardinality $|\mathcal{C}_a \cap \mathbf{O}\{v_l\}|$ of each configuration and the respective level of the outcome set $\mathbf{O}\{v_l\}$ is given in Equation (11) and also presented in Table 2.

$$|\mathcal{C}_a \cap \mathbf{O}\{v_l\}| = \sum_{i=1}^{n} \min[\min[\mathbf{C}_1\{v_l\}c_i, \mathbf{C}_2\{v_l\}c_i, \mathbf{C}_3\{v_l\}c_i], \mathbf{O}\{v_l\}o_i] \tag{11}$$

The sufficiency inclusion score of each configuration $\text{Incl}_S(\mathcal{C}_a)$ is the ratio between the two as specified in Equation (12) and presented under column Incl_S in Table 3 for each level of $\mathbf{O}$.

$$\text{Incl}_S(\mathcal{C}_a) = \frac{|\mathcal{C}_a \cap \mathbf{O}\{v_l\}|}{|\mathcal{C}_a|} \tag{12}$$

All outcome values given in Table 3 should be uncontroversial. No score apart from those indicating perfect inclusion could be considered as indicating quasi-sufficiency (e.g. 0.85).[6] The number of cases required in each configuration to not be coded as a logical remainder is a single case because with only ten cases the total number is rather small. Configuration $\mathcal{C}_{11}$ shows no inclusion score because no case has a positive membership in it. Expressed differently, $|\mathcal{C}_{11}| = 0$. The canonical sum of fundamental

Table 3: Truth Table

					Incl_S		OUT [a]		
$\mathcal{C}_a$	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_3$	n	$\mathbf{O}\{1\}$	$\mathbf{O}\{0\}$	$\mathbf{O}\{1\}$	$\mathbf{O}\{0\}$	i
1	1	α	0	0	1.00	1.00	?	?	-
2	1	β	1	2	0.39	0.72	0	0	3, 4
3	1	γ	0	1	1.00	0.29	1	0	10
4	1	α	1	2	0.44	0.25	0	0	7, 8
5	1	β	0	0	1.00	1.00	?	?	-
6	1	γ	1	1	1.00	0.67	1	0	5
7	0	α	0	1	1.00	0.50	1	0	1
8	0	β	1	1	0.00	1.00	0	1	9
9	0	γ	0	2	0.59	0.53	0	0	2, 6
10	0	α	1	0	1.00	0.75	?	?	-
11	0	β	0	0	-	-	?	?	-
12	0	γ	1	0	0.67	1.00	?	?	-

[a] number of cases cut-off: 1; inclusion cut-off: 0.9

product terms which is to be explained for all true configurations with respect to $\mathbf{O}\{1\}$ is given by Equation (13).

$$f_{\{1\}} = \overbrace{\mathbf{C}_1\{1\}\mathbf{C}_2\{\gamma\}\mathbf{C}_3\{0\}}^{\mathcal{C}_3} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}_2\{\gamma\}\mathbf{C}_3\{1\}}^{\mathcal{C}_6} + \overbrace{\mathbf{C}_1\{0\}\mathbf{C}_2\{\alpha\}\mathbf{C}_3\{0\}}^{\mathcal{C}_7} \tag{13}$$

Using the theorems of Boolean algebra, canonical sum (13) can be reduced to minimal sum (14).

$$\mathbf{C}_1\{1\}\mathbf{C}_2\{\gamma\} + \mathbf{C}_1\{0\}\mathbf{C}_2\{\alpha\}\mathbf{C}_3\{0\} \rightarrow \mathbf{O}\{1\} \tag{14}$$

In usual QCA parlance, this means that the second level of $\mathbf{C}_1$ (the presence of $\mathbf{C}_1$) in conjunction with the third level of $\mathbf{C}_2$ (the presence of $\mathbf{C}_2\{\gamma\}$), or the first level of $\mathbf{C}_1$ (the absence of $\mathbf{C}_1$) in conjunction with the first level of $\mathbf{C}_2$ (the presence of $\mathbf{C}_2\{\alpha\}$) and the first level of $\mathbf{C}_3$ (the absence of $\mathbf{C}_3$) are sufficient for the second level of $\mathbf{O}$ (the presence of $\mathbf{O}$).

In summary, mvQCA can be easily combined with csQCA and fsQCA if the set-data type that is associated with each variant is brought into the standardized system of level-score notation. Parameters of fit remain as valid as truth table construction and minimization procedures. The next section will now show how to extend this system further to include multi-level fuzzy sets, which reconcile the two dimensions that are covered separately by binary-level fuzzy sets and multi-level crisp sets. In order to distinguish them from the existing variants, analyses with multi-level fuzzy sets will be referred to as generalized-set Qualitative Comparative Analysis (gsQCA).

Generalized-Set QCA

The approach taken by gsQCA builds on *multi-level fuzzy sets*, which represent the generalization of the three other set-data types introduced in the preceding section. A multi-level fuzzy set is more general because it combines the category dimension of multi-level crisp sets with the graded membership dimension of binary-level fuzzy sets. Figure 3 illustrates how some multi-level fuzzy set $\mathbf{S}_j$ should be conceived of geometrically. Only the first three levels of $\mathbf{S}_j$ can be visualized, but theoretically the number of levels is not limited.

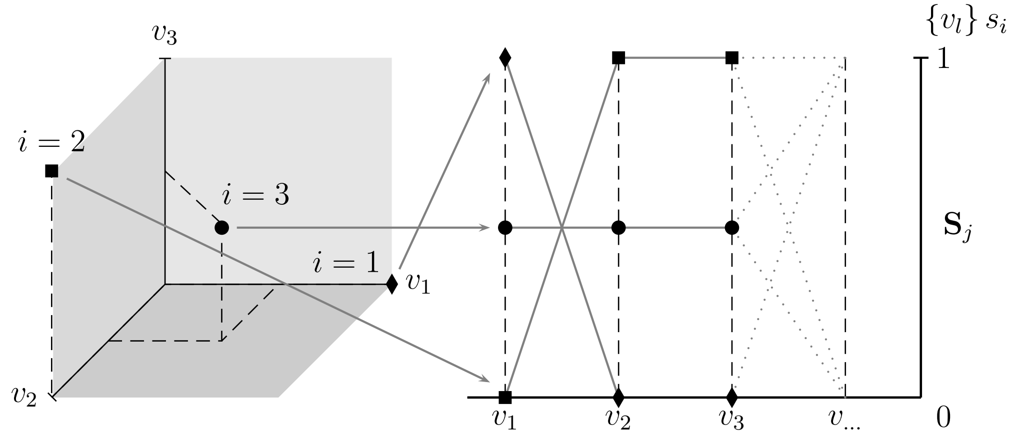

Figure 3: Geometric Representation of A Multi-Level Fuzzy Set

In Figure 3, three cases are shown, each represented by a different plot marker. A diamond designates case $i = 1$, a square case $i = 2$ and a circle $i = 3$. The three levels of $\mathbf{S}_j$, $\{v_l\}$ with $l = 1, 2, 3$ form a three-dimensional vector space in the unit interval. In contrast to multi-level crisp sets, which have mutually-exclusive levels and are restricted to membership or non-membership, a case in a multi-level fuzzy set can possess membership in any number of levels, and to different degrees, at that. For example, case $i = 1$ can only be distinguished by its notation from a case in a multi-level crisp set because it has full membership in the first level, but no membership in the second and third level of $\mathbf{S}_j$. In contrast, case $i = 2$ has no membership in the first level, but it is a full member of the second and third level. Case $i = 3$ most clearly reflects the characteristics of a multi-level fuzzy set, with a partial membership of 0.5 in each of $\mathbf{S}_j$'s three levels. Equation (15) summarizes these relations in level-score notation.[7]

$$\begin{aligned} &1\colon \mathbf{S}_j\langle\{v_1\}1.0\{v_2\}0.0\{v_3\}0.0\rangle \\ &2\colon \mathbf{S}_j\langle\{v_1\}0.0\{v_2\}1.0\{v_3\}1.0\rangle \\ &3\colon \mathbf{S}_j\langle\{v_1\}0.5\{v_2\}0.5\{v_3\}0.5\rangle \end{aligned} \tag{15}$$

As a specific example of a multi-level fuzzy set, imagine the set "employment status" $\mathbf{ES}$ with the three distinct levels "unemployed" $\mathbf{ES}\{u\}$, "employed" $\mathbf{ES}\{e\}$ and "self-employed" $\mathbf{ES}\{s\}$. Now imagine two people, p_1 and p_2, the former of whom only works two days a week for a company, but is otherwise at home for childcare, whereas the latter works four days a week for a company, but has an independent business on Fridays. As a matter of fact, it is impossible to describe either employment status accurately by using a binary-level fuzzy set that can only represent one category. Even less suitable is a binary-level crisp set because none of the two cases fully matches any of the categories. Also a multi-level crisp set is inappropriate because each case's employment status cannot be captured by one category only. A multi-level fuzzy set solves this problem. Using level-score notation, p_1 is most accurately described by p_1: $\mathbf{ES}\langle\{u\}0.6\{e\}0.4\{s\}0\rangle$, p_2 by p_2: $\mathbf{ES}\langle\{u\}0\{e\}0.8\{s\}0.2\rangle$. Note that level membership scores in multi-level fuzzy sets need not sum up to unity. If the standard working week consists of five days over which membership scores yield unity - a realistic assumption - but p_2 is self-employed on Fridays as well as Saturdays, the most accurate description of this case is p_2: $\mathbf{ES}\langle\{u\}0\{e\}0.8\{s\}0.4\rangle$.[8]

The concept of the multi-level fuzzy set has several consequential implications, from the perspective of set-theoretic principles in general, and for QCA in particular. Boolean negation does not apply to multi-level fuzzy sets because a case's membership in any category of such a set is unrelated to its membership in all other categories. How much a person is unemployed is independent of the degree to which she is employed if there is a distinction between being employed and being self-employed. As a result, the notational representation of a multi-level fuzzy set has to include all possible level-score terms, not just any single term as is the case for binary-level crisp and fuzzy sets, or the term that represents the category in which a case has full membership as at least needed for multi-level crisp sets. Figure 4 illustrates the extended hierarchy of set-data types now in place under gsQCA. The addition of multi-level fuzzy sets includes as direct special cases both binary-level fuzzy sets and multi-level crisp sets,

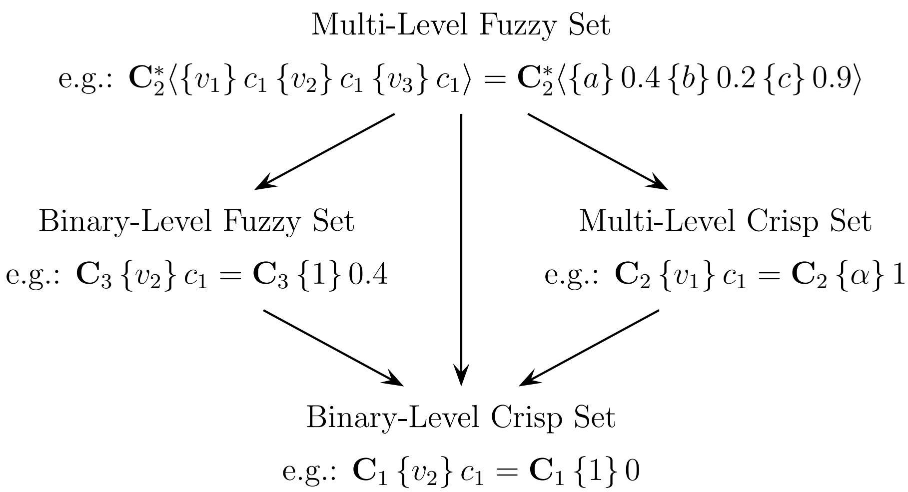

Figure 4: Extended Hierarchy of Set-Data Types

and indirectly also binary-level crisp sets. In consequence, multi-level fuzzy sets break down the dimensional incompatibility between multiple levels and graded membership.

A new data table incorporating the multi-level fuzzy set $\mathbf{C}_2^*$ is presented in the left part of Table 4. This set incorporates the three levels $\mathbf{C}_2^*\{a\}$, $\mathbf{C}_2^*\{b\}$ and $\mathbf{C}_2^*\{c\}$. Condition $\mathbf{C}_1$ remains a binary-level crisp set and $\mathbf{C}_3$ as well as the outcome set $\mathbf{O}$ binary-level fuzzy sets.

Table 4: Data Table of Set-Level and Configuration Membership Scores, with Multi-Level Fuzzy Set

	Set-Level Data				Configuration $\mathcal{C}_a = \ldots$ from Table 5											
i	$\mathbf{C}_1$	$\langle \quad \mathbf{C}_2^* \quad \rangle$	$\mathbf{C}_3$	$\mathbf{O}$	1	2	3	4	5	6	7	8	9	10	11	12
1	$\{1\}0$	$\{a\}\,0.1\,\{b\}\,0.2\,\{c\}\,0.7$	$\{1\}0.4$	$\{1\}0.7$	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.6	0.1	0.2	0.4
2	$\{1\}0$	$\{a\}\,0.7\,\{b\}\,0.1\,\{c\}\,0.4$	$\{1\}0.2$	$\{1\}0.1$	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.1	0.4	0.2	0.1	0.2
3	$\{1\}1$	$\{a\}\,0.5\,\{b\}\,0.7\,\{c\}\,0.1$	$\{1\}0.8$	$\{1\}0.5$	0.2	0.7	0.1	0.5	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0
4	$\{1\}1$	$\{a\}\,0.1\,\{b\}\,0.4\,\{c\}\,0.8$	$\{1\}1.0$	$\{1\}0.2$	0.0	0.4	0.0	0.1	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0
5	$\{1\}1$	$\{a\}\,0.9\,\{b\}\,1.0\,\{c\}\,0.8$	$\{1\}0.6$	$\{1\}1.0$	0.4	0.6	0.4	0.6	0.4	0.6	0.0	0.0	0.0	0.0	0.0	0.0
6	$\{1\}0$	$\{a\}\,0.7\,\{b\}\,0.3\,\{c\}\,0.3$	$\{1\}0.1$	$\{1\}0.9$	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.1	0.3	0.1	0.3	0.1
7	$\{1\}1$	$\{a\}\,0.6\,\{b\}\,0.9\,\{c\}\,0.8$	$\{1\}0.7$	$\{1\}0.3$	0.3	0.7	0.3	0.6	0.3	0.7	0.0	0.0	0.0	0.0	0.0	0.0
8	$\{1\}1$	$\{a\}\,0.3\,\{b\}\,0.1\,\{c\}\,0.4$	$\{1\}0.9$	$\{1\}0.4$	0.1	0.1	0.1	0.3	0.1	0.4	0.0	0.0	0.0	0.0	0.0	0.0
9	$\{1\}0$	$\{a\}\,0.8\,\{b\}\,0.2\,\{c\}\,0.8$	$\{1\}1.0$	$\{1\}0.0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.8	0.0	0.8
10	$\{1\}1$	$\{a\}\,0.7\,\{b\}\,0.9\,\{c\}\,0.6$	$\{1\}0.0$	$\{1\}1.0$	0.7	0.0	0.6	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
				$\lvert\mathcal{C}_a\rvert$	1.7	2.5	1.5	2.1	1.9	2.6	1.5	0.6	1.3	1.2	0.6	1.5
				$\lvert\mathcal{C}_a \cap \mathbf{O}\{1\}\rvert$	1.7	1.7	1.5	1.8	1.9	1.6	0.9	0.4	1.0	0.3	0.6	0.6
				$\lvert\mathcal{C}_a \cap \mathbf{O}\{0\}\rvert$	0.6	1.7	0.5	1.5	0.6	2.0	0.9	0.6	0.8	1.2	0.4	1.4

The complete truth table that results from applying the truth table algorithm used in the preceding section to the set-level data in the left part of Table 4 is shown in Table 5. Notice that, on the one hand, some cases have membership above 0.5 in more than one configuration as indicated by more than one gray cell in each row of Table 4 and the multiple appearance of the same case identifier in the case column i of Table 5. For example, case $i = 5$ has membership in configurations $\mathcal{C}_2$, $\mathcal{C}_4$ and $\mathcal{C}_6$. On the other hand, however, it can also be observed that case $i = 8$ has no membership of above 0.5 in any configuration. Generally, if a case's membership score does not exceed 0.5 in at least one category of a multi-level fuzzy set, it will not have strong membership in any truth table configuration. Thus, with multi-level fuzzy sets, not only does Boolean negation not help in identifying all level membership scores, but also can cases have membership above 0.5 in more than one configuration or even none at all. This may seem counter-intuitive at first to most researchers familiar with QCA. However, the example of people's employment status above has demonstrated that some cases can only be suitably represented by assigning partial membership to more than a single category.

Table 5: Truth Table

					Incl_S		OUT [a]		
$\mathcal{C}_a$	$\mathbf{C}_1$	$\mathbf{C}_2^*$	$\mathbf{C}_3$	n	$\mathbf{O}\{1\}$	$\mathbf{O}\{0\}$	$\mathbf{O}\{1\}$	$\mathbf{O}\{0\}$	i
1	1	a	0	1	1.00	0.35	1	0	10
2	1	b	1	3	0.68	0.68	0	0	3, 5, 7
3	1	c	0	1	1.00	0.33	1	0	10
4	1	a	1	1	0.86	0.71	1/0	0	5, 7
5	1	b	0	1	1.00	0.32	1	0	10
6	1	c	1	3	0.62	0.77	0	0	4, 5, 7
7	0	a	0	2	0.60	0.60	0	0	2, 6
8	0	b	1	0	0.67	1.00	?	?	-
9	0	c	0	1	0.77	0.62	0	0	1
10	0	a	1	1	0.25	1.00	0	1	9
11	0	b	0	0	1.00	0.67	?	?	-
12	0	c	1	1	0.40	0.93	0	1	9

[a] number of cases cut-off: 1; inclusion cut-off: 0.85/0.90

In order to further examine the consequences of these features of multi-level fuzzy sets, consider what happens if the cut-off for the sufficiency inclusion score on whose basis the outcome value for each configuration is established is set to 0.85 in Table 5. This value is certainly not below commonly accepted levels. The canonical sum for all true configurations with respect to $\mathbf{O}\{1\}$ is then given by Equation (16).

$$f_{\{1\}} = \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{a\}\mathbf{C}_3\{0\}}^{\mathcal{C}_1} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{b\}\mathbf{C}_3\{0\}}^{\mathcal{C}_3} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{a\}\mathbf{C}_3\{1\}}^{\mathcal{C}_4} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{c\}\mathbf{C}_3\{0\}}^{\mathcal{C}_5} \tag{16}$$

Minimization of these four fundamental products yields the minimal sum with two prime implicants given in Equation (17). Condition $\mathbf{C}^*_2$ is redundant in the first, second and fourth, and condition $\mathbf{C}_3$ is redundant in the first and third fundamental products.

$$\mathbf{C}_1\{1\}\mathbf{C}^*_2\{a\} + \mathbf{C}_1\{1\}\mathbf{C}_3\{0\} \rightarrow \mathbf{O}\{1\} \tag{17}$$

If, however, the cut-off for the sufficiency inclusion score is raised to 0.9, $\mathcal{C}_4$ drops out and the reduced canonical sum for the remaining configurations is now given by Equation (18).

$$f_{\{1\}} = \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{a\}\mathbf{C}_3\{0\}}^{\mathcal{C}_1} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{b\}\mathbf{C}_3\{0\}}^{\mathcal{C}_3} + \overbrace{\mathbf{C}_1\{1\}\mathbf{C}^*_2\{c\}\mathbf{C}_3\{0\}}^{\mathcal{C}_5} \tag{18}$$

Condition $\mathbf{C}^*_2$ is again redundant across these three fundamental products. The resulting minimal sum is given by Equation (19).

$$\mathbf{C}_1\{1\}\mathbf{C}_3\{0\} \rightarrow \mathbf{O}\{1\} \tag{19}$$

The minimal sum consists of a single prime implicant that covers three configurations: $\mathcal{C}_1$, $\mathcal{C}_3$ and $\mathcal{C}_5$. However, these configurations only contain $i = 10$ as their single strong case. It is the only case that has membership above 0.5 in all configurations. With multi-level fuzzy sets, complex solutions therefore need not cover at least as many empirically strong cases as there have been fundamental products in the canonical sum. It is well possible that a single case conforms to more than one configuration.

While the theoretical implications of multi-level fuzzy sets are not to be underestimated, the practical consequences for applied research are to be seen elsewhere. For one, a relatively small number of social-scientific concepts lend themselves to being represented as multi-level fuzzy sets. In the vast majority of cases, multi-level crisp sets will suffice. The main contribution of this article to applied research should thus be seen in the introduction of a new notational system which allows the joint processing of either binary-level set structure and multi-level crisp sets.

Conclusions

This article set out to demonstrate that the distinction between csQCA, mvQCA and fsQCA which has existed so far in the literature on comparative configurational methodology can be broken down by reconciling all variants with each other under a single framework. The approach whereby this is made possible has been referred to as generalized-set QCA (gsQCA). The fundamental concept of gsQCA is the multi-level fuzzy set, a set-data type which incorporates the multi-dimensionality of multi-level crisp sets and the graded-membership principle of binary-level fuzzy sets, so that all three existing QCA variants become special cases of gsQCA. If all conditions have two levels, one for the presence and one for the absence of the attribute represented by the set, for each of which the membership function maps cases into the binary set $\mathbb{B} = \{0, 1\}$, gsQCA collapses to csQCA. If at least one condition has more than two mutually-exclusive levels for each of which the membership function maps cases into the binary set $\mathbb{B} = \{0, 1\}$, but in only one of which cases can have membership, and all remaining conditions have two levels, one for the presence and one for the absence of the attribute represented by the set, for each of which the membership function maps cases into the binary set $\mathbb{B} = \{0, 1\}$, gsQCA becomes mvQCA. If at least one condition in the data has two levels, one for the presence and one for the absence of the attribute represented by the set, for each of which the membership function maps cases into the unit interval $\mathbb{U} = [0, 1]$ according to the principle of Boolean negation and all remaining conditions have two such levels for each of which the membership function maps cases into the binary set $\mathbb{B} = \{0, 1\}$, gsQCA turns into fsQCA. With gsQCA, researchers are therefore not forced any longer to accept a loss of information by having to turn continuous base variables into binary-level crisp sets when their core conditions are multi-level crisp sets. The replacement of membership and category-notation by gsQCA's level-score notation allows the reconciliation of these two set-data types.

The introduction of gsQCA opens up a new perspective on the QCA family of configurational comparative methods. Most significantly, the three existing variants should not be seen any longer as standing in competition, but as being special cases of a general QCA approach for the analysis of specific combinations of set-data types. No substantial disadvantages result from this generalization. Under gsQCA, all parameters of fit (mainly inclusion and coverage), truth table construction and Boolean minimization procedures, as well as the derivation of complex, intermediate and parsimonious solutions remain valid. If there is a cost to be found, it is the more elaborate system of level-score notation which is required under gsQCA. In consideration of the added benefit the method brings, however, this should be but a symbolic price to pay.

Notes

[1]According to the single-membership principle, each case can have a membership score of at most 0.5 in more than one truth table configuration, but it can only have membership above 0.5 in a single configuration.

[2]To the knowledge of the author, the first application of csQCA has been presented by Ragin, Mayer and Drass (1984).

[3]The logical NOT is usually indicated in csQCA and fsQCA either by lower-case letters, in contrast to upper-case letters, by a tilde sign preceding the set name or a prime. These indicators do not apply in level-score notation.

[4]Greek letters have been used for $\mathbf{C}_2$ in order to reiterate that the identifier of a level need not be a numeric value. Angled brackets are used when more than one level-score term is provided for a case within an expression.

[5]*Inclusion* corresponds to what Ragin (2006) refers to as *consistency*.

[6]The outcome value coding of the truth table is not based on probabilistic criteria.

[7]These expressions may be collapsed to 1: $\mathbf{S}_j\langle\{v_1\}1.0\{v_2, v_3\}0.0\rangle$, 2: $\mathbf{S}_j\langle\{v_1\}0.0\{v_2, v_3\}1.0\rangle$ and 3: $\mathbf{S}_j\{v_1, v_2, v_3\}0.5$.

[8]As another example, imagine **COL** to represent the RGB color space with the three additive primary colors red $\mathbf{COL}\{r\}$, green $\mathbf{COL}\{g\}$ and blue $\mathbf{COL}\{b\}$ as its three levels. RGB color spaces closely resemble the human visual system in the perception of colors. Where does "purple" as a case fit in? The membership score of purple is exactly 0.5 in both $\mathbf{COL}\{r\}$ and $\mathbf{COL}\{b\}$, but it is 0 in $\mathbf{COL}\{g\}$. Using level-score notation, purple is thus identified by purple: $\mathbf{COL}\langle\{r\}0.5\{g\}0\{b\}0.5\rangle$.

References

Amenta, Edwin and Drew Halfmann. 2000. "Wage Wars: Institutional Politics, WPA Wages, and the Struggle for US Social Policy." *American Sociological Review* 65(4):506–528.

Avdagic, Sabina. 2010. "When Are Concerted Reforms Feasible? Explaining the Emergence of Social Pacts in Western Europe." *Comparative Political Studies* 43(5):628–657.

Balthasar, Andreas. 2006. "The Effects of Institutional Design on the Utilization of Evaluation." *Evaluation* 12(3):353–371.

Berg-Schlosser, Dirk. 2008. "Determinants of Democratic Successes and Failures in Africa." *European Journal of Political Research* 47(3):269–306.

Brown, Cliff and Terry Boswell. 1995. "Strikebreaking or Solidarity in the Great Steel Strike of 1919: A Split Labor Market, Game-Theoretic, and QCA Analysis." *American Journal of Sociology* 100(6):1479–1519.

Clark, Terry D., Jennifer M. Larson, John N. Mordeson, Joshua D. Potter and Mark J. Wierman. 2008.

Cronqvist, Lasse. 2004. *Presentation of TOSMANA. Adding Multi-Value Variables and Visual Aids to QCA.* COMPASSS: WP 2004-20.
URL: *http://www.compasss.org/files/WPfiles/Cronqvist2004.pdf*

Cronqvist, Lasse. 2011. *Tosmana - Tool for Small-N Analysis, Version 1.3.2.* Trier: University of Trier.

Cronqvist, Lasse and Dirk Berg-Schlosser. 2009. Multi-Value QCA (mvQCA). In *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, ed. Benoît Rihoux and Charles C. Ragin. London: Sage Publications pp. 69–86.

Dixon, Marc, Vincent J. Roscigno and Randy Hodson. 2004. "Unions, Solidarity, and Striking." *Social Forces* 83(1):3–33.

Duşa, Adrian and Alrik Thiem. 2012. *QCA: Qualitative Comparative Analysis.* R package version 1.0-4.

Herrmann, Andrea Monika and Lasse Cronqvist. 2009. "When Dichotomisation becomes a Problem for the Analysis of Middle-Sized Datasets." *International Journal of Social Research Methodology* 12(1):33–50.

Hodson, Randy, Vincent J. Roscigno and Steven H. Lopez. 2006. "Chaos and the Abuse of Power." *Work and Occupations* 33(4):382–416.

Huang, Ronggui. 2012. *QCA3: Yet Another Package for Qualitative Comparative Analysis.* R package version 0.0-5.

Klir, George J., Ute H. St. Clair and Bo Yuan. 1997. *Fuzzy Set Theory: Foundations and Applications.* Upper Saddle River, NJ: Prentice Hall.

Lilliefeldt, Emelie. 2012. "Party and Gender in Western Europe Revisited: A Fuzzy-Set Qualitative Comparative Analysis of Gender-Balanced Parliamentary Parties." *Party Politics* 18(2):193–214.

Longest, Kyle C. and Stephen Vaisey. 2008. "fuzzy: A program for Performing Qualitative Comparative Analyses (QCA) in Stata." *Stata Journal* 8(1):79–104.

McCluskey, Edward J. 1965. *Introduction to the Theory of Switching Circuits.* Princeton: Princeton University Press.

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

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

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

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, Charles C. 2009. Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA). In *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, ed. Benoît Rihoux and Charles C. Ragin. London: Sage Publications pp. 87–121.

Ragin, Charles C. and Sean Davey. 2009. *Fuzzy-Set/Qualitative Comparative Analysis fs/QCA 2.5*. Tucson: Department of Sociology, University of Arizona.

Ragin, Charles C., Susan E. Mayer and Kriss A. Drass. 1984. "Assessing Discrimination: A Boolean Approach." *American Sociological Review* 49(2):221–234.

Rihoux, Benoît and Gisèle De Meur. 2009. Crisp-Set Qualitative Comparative Analysis (csQCA). In *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*, ed. Charles C. Ragin and Benoît Rihoux. London: SAGE pp. 33–68.

Schneider, Carsten Q. and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis (QCA)*. Cambridge: Cambridge University Press.

Thiem, Alrik. 2011. "Conditions of Intergovernmental Armaments Cooperation in Western Europe, 1996-2006." *European Political Science Review* 3(1):1–33.

Thiem, Alrik. forthcoming. "Clearly Crisp, and Not Fuzzy: A Comment on Vink and van Vliet (2009)." *Field Methods* xx(x):xx–xx.

Thiem, Alrik and Adrian Duşa. 2012*a*. *QCA: A Package for Qualitative Comparative Analysis*. Swiss Federal Institute of Technology Zurich: Unpublished manuscript.

Thiem, Alrik and Adrian Duşa. 2012*b*. *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer.

Vink, Maarten P. and Olaf van Vliet. 2009. "Not Quite Crisp, Not Yet Fuzzy? Assessing the Potentials and Pitfalls of Multi-value QCA." *Field Methods* 21(3):265–289.

Zadeh, Lofti A. 1965. "Fuzzy Sets." *Information and Control* 8(3):338–353.