Fuzzy-set Qualitative Comparative Analysis Summary[1]

Arthur P. Tomasino
Bentley University

Fuzzy-set Qualitative Comparative Analysis (fsQCA) is a relatively new but accepted method for understanding case-based research[2]. The technique utilizes a holistic perspective exploring the similarities and differences across cases. Theory and evidence are linked by the presence and absence of outcomes in configurations of causal conditions allowing for the assessment of the multiple conjunctions of causes of an outcome (Ragin, 2014).

FsQCA is based on set theory, focusing on the degree to which any one variable is a subset or superset of another. Sets are conventionally thought of as dichotomous (or "Crisp") and cases under study are either "in" or "out" of the set. For example, the set of employed workers would conventionally be represented by a binary variable with two values, 1 ("in", i.e. employed) and 0 ("out", i.e., not employed[3]). In contrast fuzzy-sets allow for membership in the set between 0 and 1 (with 0 representing full non-membership and 1 representing full membership). Using the previous example, a "part-time" employee, would have a membership in the fuzzy-set "employed" somewhere between 0 and 1 (as defined by the researcher) as the worker is neither fully employed (1) or fully not employed (0).

Set membership	Membership code
fully in	**1**
more in than out	Between 1 and 0.5
maximally ambiguous	.5
more out than in	Between 0.5 and 0
fully out	0

Table 1 - Set Membership Coding

Although it is tempting to view membership as a continuous variable, fuzzy-set membership is more than a continuous variable but is rather a *calibration* of a set of variables related to the degree of membership within a category. Through calibration a single variable or multiple variables are assigned to fuzzy-sets representing both a qualitative and quantitative assessment of case characteristics ((Ragin, 2006), p8). It requires strong conceptual and theoretical guidance to assign membership to each set forming a set-theoretic relationship between the fuzzy-set and theory. Using substantial theoretic knowledge relevant to set membership the resulting fuzzy-set is then a fine-grained measure of carefully calibrated case variables ((Ragin, 2000), p7).

The basic concepts of fsQCA (Table 2) differ from standard statistical analysis techniques relying on correlations to determine causality and significance tests to access generalize-ability. FsQCA focuses on the analysis of *necessary and sufficient* conditions and a *set-theoretic perspective* to determine causality of outcomes. Complex causality or multiple conditions leading to an outcome (equifinality) inherently exist in the analysis. If the causal conditions represent system descriptors (variable or factors) then the different combinations of descriptors leading to the outcome are the possible states in which the system can exist.

[1] The description of fsQCA is based on (Kent, 2008; Ragin, 2006, 2007, 2009; Ragin & Giesel, 2008; C. Schneider & Wagemann, 2010; C. Q. Schneider & Grofman, 2006)
[2] See (C, 2005; Marx, Rihoux, & Ragin, 2014; Ragin, 2006, 2008; Ragin & Rihoux, 2009)
[3] This is a subtle point in set theory. In the example "0" does not mean unemployed but rather completely out of the set of employed workers. A case could be "0" in "employed" and still be a part-time employee.

FsQCA is best applied when the researcher assumes (or believes) that complex causality is present and the population of cases is too low for statistical techniques. Typically, the researcher has previously gained knowledge of the population and uses this knowledge to define, specify and measure the key concepts applicable to the system under study.

fsQCA Concept	Description
Complex Causality	FsQCA focuses on complex causality or multiple interacting conditions that create system outcomes.
Equifinality	Different conditions can lead to the same outcome (equifinality).
Qualitative data	Analyzed data are qualitative in nature. Data expresses membership of cases in sets.
Interpretation	The interpretation of results in terms of necessary and sufficient conditions.
Set-theoretic Relationships	Conceptualization of relations between conditions and outcomes as a set relation (not a covariation).
Iterative Analysis	Iterative FsQCA analysis requires redefinition of conditions and potentially adding or deleting cases to uncover sets of causal conditions

Table 2 - The six basic concepts of importance in fsQCA (C. Q. Schneider & Grofman, 2006)

Creating set-theoretic relationships, calibrating fuzzy-sets, and performing fsQCA analysis employs a four step process. First, measures for the elements of our analysis are created as combinations quantitative and qualitative data. For example, measures may consist of a combination of survey questions and respondent demographics. A careful calibration of the measures follows, defining set-theoretic memberships. The third step uses fsQCA to create configurations of measures representing causes for outcomes of interest. The fourth step uses fuzzy-set consistency and coverage to validate (or invalidate) these configurations. In the fifth step, the resulting configurations are interpreted. In the following sections summarize each of these steps. We presetn both QCA and fsQCA for clarity of explanation.

Measures and Membership

Membership is a measure of the relation of a condition to a set. For example, using the previous employed example, a researcher may define the set “employed” to include the U.S. Bureau of Labor Statistics defining “employed” as working between 34 and 40 hours per week[4]. “Not employed” is a much more complicated concept based on profession, salaries, and the State in which work is performed, but based on statistics the researcher might conclude working under 10 hours is “not employed” and calibrate the remaining cases using an S-curve as shown in Figure 1. Theoretical concepts, at the discretion of the researcher, would be the basis for calibrations.

[4] http://www.bls.gov/news.release/empsit.t08.htm

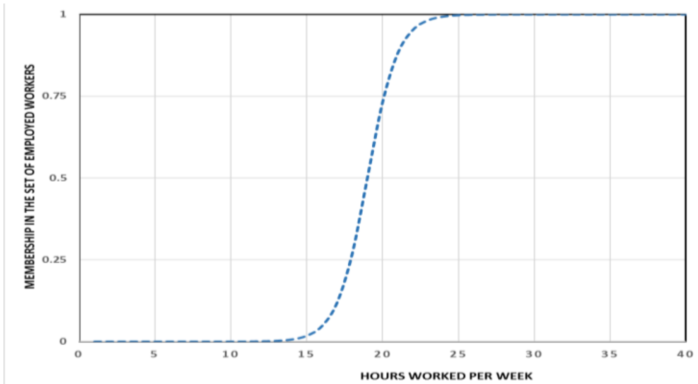

Figure 1 - Membership Calibration of the Set "Employed"

Numerous techniques have emerged for calibrating membership and many are specific to the research and cases, as summarized in Table 3.

Technique	Description
Likert Scales	*Likert scales* that access conditions such as "strongly agree" to "strongly disagree" simply translate to membership scores. A 5-point Likert scale would translate to membership values of (0, .25, .5, .75, 1).
Direct Membership	*Direct membership* uses the full membership, full non-membership and crossover to anchor calibration. Intermediate membership can then be calculated using exponents and probabilities to create a smooth S-shaped calibration curve (Ragin, 2007).
Indirect Membership	*Indirect membership* relies on the researcher's knowledge and grouping of cases according to their membership in a target set. Many times this type of calibration is open and revised as a study proceeds. Typically, theory based, calibration is at a minimum documented in detail giving meaning to the calibration and further fsQCA analysis.
Counting	*Counting* derives membership using the proportion of positive answers to total answers to derive membership. For example, the proportion of number of attributes that are associated with a set and exhibited by a case to the total number of attributes defining the set defines membership

Table 3 - Calibration Techniques for Set Membership

Of critical importance, independent of calibration technique(s) used, is the detailed documentation of the techniques so resulting causal conditions from an fsQCA analysis can be explained and evaluated (C. Schneider & Wagemann, 2010).

Configurations and Outcomes

Truth tables represent all the logically possible combinations of the conditions, or *configurations* resulting in an *outcome*. It is essentially the representation of the empirical data from a study in tabular form. For example a system with outcome Z and causal conditions A, B, and C, might have observed data as shown in the table below.

QCA Truth Table					fsQCA Truth Table			
A	**B**	**C**	**Z**		**A**	**B**	**C**	**Z**
0	0	0	**0**		.2	1	.1	**.4**
0	0	1	**1**		1	1	1	**.7**
0	1	0	**1**		.8	.5	1	**.1**
0	1	1	**0**		.5	.7	0	**1**
1	0	0	**1**		.5	.5	.5	**0**
1	0	1	**1**		1	.6	.2	**.1**
1	1	0	**0**		0	1	0	**1**

1	1	1	**0**		.1	1	1	**.5**

Table 4 - Example Truth Tables for QCA and fsQCA

QCA analyzes crisp (only "0" or "1") membership and fsQCA continuous (or "fuzzy") membership. Since QCA uses crisp data, its truth table is usually constructed for every possible combination of inputs A, B, and C. If a combination is not empirically observed it can be deleted from the table or designated as a "don't care" (either 0 or 1, denoted by "X") based on the knowledge of the researcher. Truth tables for fsQCA only contain the observed empirical data, as there are infinite possible combinations of inputs.

A solution formula is a way of expressing the results (the configuration of conditions exhibiting the outcome) of QCA or fsQCA analysis. Letters (or strings) linked by Boolean operators represent outcomes and their causally relevant conditions. Formulas use Boolean, rather than arithmetic, operators. The three basic Boolean operators are logical OR (+), logical AND (*), and logical NOT (~). Each operator is defined the same for QCA and fsQCA as in Table 5.

Logical Operator	Symbol	Description	Equation	Use
NOT	~	Negation of the original value	~X = 1-X	QCA & fsQCA
AND	*	Set intersection – calculated as the minimum value of two (or more) sets	X * Y = min(X,Y)	QCA & fsQCA
OR	+	Set union – calculated as the maximum of two (or more) sets	X + Y = max(X,Y)	QCA & fsQCA
CONCENTRATION	Conc()	Expands values to "very" by squaring the original value. For example, a person with member ship of .8 in "tall" converts to a membership of .64 in "very tall".	Conc(X) = X^2	fsQCA only
DILATION	Dil()	Transforms values to "more or less" in a set. For example, a person with a .36 membership in "rich" dilates to a membership of .6 in "more or less rich".	Dil(X) = $X^{1/2}$	fsQCA only

Table 5 - Boolean Operators for QCA and fsQCA

Combining variables using operators represents sets of the causal conditions relating to the output results. From Table 4, using the operators in Table 5, the solution formula for Z (in QCA) is as follows.

Y ← ~A*~B*C + ~A*B*~C + A*~B*~C + A*~B*C

FsQCA produces similar solution formulas. The sign ← indicates a logical relationship.

Necessary and Sufficient Conditions

When analyzing complex causality, fsQCA determines the necessary and sufficient causal conditions to produce an outcome. For necessary conditions the relevant causal condition are present in all instances of an outcome. In contrast, sufficient conditions represent causal complexity because they exist only in combinations with other conditions. Stated in tandem, necessity and sufficiency provide a complete understanding of causality, as summarized in Table 6.

Conditions	Description
Necessary and Sufficient	A condition is *necessary and sufficient* if it is the only condition producing an outcome. Similarly if it is absent then the outcome is also absent.
Necessary but not Sufficient	A condition is *necessary but not sufficient* if it is contained in all combinations producing the outcome. Again, if it is absent the condition is also absent. In QCA, a necessary condition occurs when the outcome is present (1) and the condition, or input, is also present (1). In fsQCA necessity is indicated when the membership of the input (x_i) is greater than the membership of the outcome (y_i). The set of

	cases containing the input condition subsumes the output set.
Sufficient	A condition is *sufficient but not necessary* if it is capable of producing the outcome by itself, but at the same time, other combinations of conditions can also produce the outcome. In QCA whenever a sufficient condition exists (1) the outcome is also present (1). In fsQCA sufficiency is indicated when the membership of the input (x_i) is less than or equal to the membership of the outcome (z_i). The set of cases containing the input condition is a subset of the output set.
Neither Necessary nor Sufficient	A condition is *neither necessary nor sufficient* if it produces an output only if combined with other conditions.

Table 6 - Necessary and Sufficient Conditions

Necessary conditions are always present when the outcome is present. A condition is sufficient if when present it produces the output but the output may be present when the sufficient condition is not present. Since sufficiency relates to the combinations of conditions that create an output the main analysis from QCA and fsQCA is a determination of the list of **sufficient conditions** for a specific outcome[5].

Consistency and Coverage

Sufficient conditions are determined from consistency and coverage measures in fsQCA. Consistency and coverage are measures of the fit of possible sufficient conditions to explain an outcome. Whenever a sufficient condition is present, the outcome is also present. The condition may also hold for the majority of cases, but not all.

For QCA *consistency* is simply the proportion of cases in which the condition produces the outcome to the number of cases with the condition. Consistency measures the subset of cases with the condition and outcome to all cases with the condition. *Coverage* is the proportion of cases that contain the condition to the total number of cases in which the outcome is present. Therefore, coverage assesses the degree to which conditions “account for” the outcome. When many configurations of conditions exist to the outcome, coverage will be very small for a particular configuration and its importance or relevance will also be small ((Ragin, 2008), p. 44).

Similarly, for fsQCA consistency is the proportion of cases with condition membership less than or equal to the output membership to the total number of cases with the outcome membership greater than zero. *High consistency values indicate the condition is sufficient for the output.* Likewise, coverage is the proportion of cases with condition membership less that than or equal to the output membership to the total number of cases where the membership of the output is greater than zero.

It is important to note that when all the membership scores for the condition are less than the membership for the outcome the consistency is unity (1) and the condition is *completely sufficient* and the cases with the outcome subsume the cases exhibiting the condition. If only a few cases have the condition membership greater than the outcome membership, the consistency is close to unity. *Therefore when utilizing fsQCA, analysis conditions considered potentially sufficient for an outcome typically have consistencies greater than .8* (Ragin, 2009). Table 7 summarizes the sufficiency, consistency, and coverage for QCA and fsQCA analysis.

---

[5] Necessity can also be determined but is less used since configurations with very low membership in the outcome can be necessary, by definition, but have little meaning towards causality, representing an outlier or error.

QCA	fsQCA
Sufficiency	
X ⊆ Y *(X is a subset of Y)*	Xi ≤ Yi *(Xi, Yi is membership)*
Consistency	
∑(Xi → Yi) / ∑(Xi)	∑(Xi ≤ Yi) / ∑(Xi)
Coverage	
∑(Xi → Yi) / ∑(Yi)	∑(Xi ≤ Yi) / ∑(Yi)

Table 5 - Sufficiency, Consistency, and Coverage for QCA and fsQCA

Therefore, fsQCA (or QCA) produce configurations of conditions, represented as solutions formulas, related to an outcome of interest. The solutions formulas are then assessed as to their casual strength to the outcome by determining consistency and coverage. Although consistency and coverage may suggest causality, the determination of causality is still predicated on an interpretation of the configurations and fsQCA results.

Interpretation of QCA and fsQCA results

Configurations, solutions formulas, consistency and coverage all constitute the set relations important to complex causality just as significance and strength are important in correlational analysis. Consistency, like significance, can support or disprove a hypothesis. For example, a hypothesized configuration with low consistency has a weak subset relationship and the hypothesis is not supported. Coverage, as in correlational strength, indicates the importance of a set-theoretic relationship. Similar to correlational analysis, where it is possible to have significant but weak correlation, in set-theoretic analysis it is possible to have highly consistent configurations with low coverage. Therefore it is important for researchers, when using set-theoretic analysis to confirm and support their results with strong theoretical foundation and substantive knowledge (Ragin, 2008).

Interpretation of results can be enhanced by creating tables of the solutions formulas with their resulting consistencies and coverages, along with the numbers of cases in each configuration. This can provide both a view and comparison of each causal path to the outcome. Additionally, the overlap of cases across configurations can also provide insight into the uniqueness of each path. Sawyer, Fedorowicz, and Tomasino (2015) provide a good example of such analysis and interpretation (shown in Figure 2). Similarly, the references included at the end of this paper provide other examples and additional information.

Performance (Perf_Gov4) as f(PSN_type, Orgl, Tech, Fin, Stake, Elect_ag, Mandate)

Config #	PSN Type	Governance Factors				Legal Auth		Raw Coverage	Unique Coverage	Consistency	#
		Orgl	Tech	Stake	Fin	Elect_ag	mandate				
1	X	X	1	1	X	0	1	0.449565	0.215362	0.915044	17
2	Court	1	1	1	X	1	0	0.145217	0.067826	0.936449	7
3	Police	0	1	1	1	1	0	0.067246	0.028985	1.00000	1
4	X	1	1	1	0	1	0	0.162319	0.046667	0.980736	4
5	X	0	X	1	0	0	1	0.256232	0.022029	0.975717	9
6	Court	1	0	1	0	0	0	0.015652	0.015652	1.00000	1
"X"means the element does not matter to the configuration								Solution coverage: 0.746377			
"0" means the element is NOT included in the configuration								Solution consistency: 0.935343			
"1" means the element IS included in the configuation											

Configuration Results Table

	1	2	3	4	5	6
1	17				6	
2		7		3		
3			1			
4				4		
5					9	
6						1

Case distribution by Configuration

Figure 2 - Example of Solution Formula Tables for fsQCA Results Interpretation

In essence, fsQCA provides a framework for comparing cases and configurations an interpreting their contribution to causality of an outcome. Patterns of consistency, inconsistency and coverage gauge the strength of the set-relationship, which through theory and knowledge can be used to assess complex causality. In contrast to correlational analysis techniques fsQCA, through analysis of necessary and sufficient conditions can be used, in concert with substantive theoretical foundations and knowledge, can be used to determine causality, in particular when equifinality is present.

References

C, R. (2005). From Fuzzy Sets to Crips Truth Tables.

Kent, R. (2008). Using fsQCA: A Brief Guide and Workshop for Fuzzy-Set Qualitative Comparative Analysis. *Department of Marketing, University of Sterling*, 2008-2010.

Marx, A., Rihoux, B., & Ragin, C. (2014). The origins, development, and application of Qualitative Comparative Analysis: the first 25 years. *European Political Science Review : EPSR, 6*(1), 115-142. doi: http://dx.doi.org/10.1017/S1755773912000318

Ragin, C. (2000). *Fuzzy-set social science*: University of Chicago Press.

Ragin, C. (2006). Set relations in social research: Evaluating their consistency and coverage. *Political Analysis, 14*(3), 291-310.

Ragin, C. (2007). Fuzzy sets: calibration versus measurement. In J. M. Box-Steffensmeier, H. E. Brady & D. Collier (Eds.), *The Oxford Handbook of Political Methodology* (pp. 174-198). Oxford: The Oxford University Press.

Ragin, C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*: Wiley Online Library.

Ragin, C. (2009). Qualitative comparative analysis using fuzzy sets (fsQCA). In B. Rihoux & C. C. Ragin (Eds.), *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques* (pp. 87-121). Thousand Oaks, CA.: Sage Publications.

Ragin, C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*: Univ of California Press.

Ragin, C., & Giesel, H. (2008). User's guide to fuzzy-set/qualitative comparative analysis. *University of Arizona*, 87.

Ragin, C., & Rihoux, B. (2009). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*: Sage.

Sawyer, S., Fedorowicz, J., & Tomasino, A. (2015). *Patterns of Governance among Inter-organizational Coordination Hubs*. Paper presented at the International Conferance of Information Systems, Fort Worth, Texas.

Schneider, C., & Wagemann, C. (2010). Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comparative Sociology, 9*(3), 397-418.

Schneider, C. Q., & Grofman, B. (2006). *It might look like a regression equation… but it's not! An intuitive approach to the presentation of QCA and fs/QCA results'.* Paper presented at the Conference on "Comparative Politics: Empirical Applications of Methodological Innovations", Sophia University, Tokyo (Japan).