Accepted preprint of the edited article published as
Rohlfing, Ingo (2020): The Choice between Crisp and Fuzzy Sets in Qualitative Comparative Analysis and the Ambiguous Consequences for Finding Consistent Set Relations. *Field Methods* 32 (1): 75-88.
https://doi.org/10.1177/1525822x19896258

The author information on the title page below might be outdated.
Check for current information on Ingo Rohfing:

ORCID: 0000-0001-8715-4771

The choice between crisp and fuzzy sets in Qualitative Comparative Analysis and the ambiguous consequences for finding consistent set relations

Ingo Rohlfing[1]

3/6/2021

Abstract

Empirical researchers using Qualitative Comparative Analysis can work with crisp, multi-value and fuzzy sets. The relative advantages of crisp and multi-value sets have been discussed in the QCA literature. There has been little reflection on the more frequent decision between crisp and fuzzy sets for which there often is no theoretical guidance. A review shows that researchers often prefer fuzzy over crisp sets, sometimes because they contain more information. This meets with the argument that fuzzy sets produce more conservative consistency measures and constitute tougher tests. In my paper, I demonstrate analytically and with data from published QCA studies that the relationship between crisp sets, fuzzy sets and the consistency score is ambiguous. It depends on the distribution of cases whether the consistency value is more or less conservative for fuzzy sets than for crisp sets. I outline the implications of the ambiguous relationship for empirical research.

Keywords: consistency, necessity, QCA, set relation, sufficiency

[1] Cologne Center for Comparative Politics, Universität zu Köln, `i.rohlfing@uni-koeln.de`. For comments and support, I am grateful to four anonymous reviewers and the editor, Dennis Bereslavskiy, Nancy Deyo, Kim Sass Mikkelsen, Cyrill Otteni, Ayjeren Rozyjumayeva and Jan Schwalbach. Special thanks go to the anonymous participant at a QCA panel at the conference of the German International Relations Section of the DVPW. The participant was asking the panelists about the choice between crisp and fuzzy sets and made me started thinking about this issue. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement nr. 638425).

Empirical researchers using Qualitative Comparative Analysis can work with crisp, multi-value and fuzzy sets.[2] Empirical researchers must work with crisp sets if the concept of interest is generically binary. If the concept is multinomial, one can choose between a multivalue set or calibrate the variable into multiple crisp sets. The advantages of both forms of calibration have been discussed in the methods literature (Vink and van Vliet 2009; Schneider and Wagemann 2012:section 10.2; Thiem 2013; Vink and van Vliet 2013).

There has been little reflection on the much more frequently encountered choice as to the use of a crisp or fuzzy set, given that it is possible to assign cases to at least three ordinal ranks. In a review of 26 empirical QCA studies published in 2016 and referenced in the Social Sciences Citation Index, I show that 17 articles do not discuss the choice of the set type (see section 4). This is a problem of itself because this design decision should be justified. Among the nine articles that do discuss the choice, five justify the utilization of fuzzy sets with the goal of retaining more information than one could with crisp sets. Of 24 studies that could have used fuzzy sets, 18 do so, with only six opting for crisp sets.

The goal of retaining more information dovetails with an argument in favor of fuzzy sets. It is argued that fuzzy sets are superior because they capture more granular information than crisp sets. As a consequence of this, the consistency value is claimed to be more conservative for fuzzy sets than crisp sets (Schneider and Wagemann 2007:225-228; 2012:68-69). In empirical research, the parameter of consistency plays a key role in inferring whether the data is consistent with the claim that a necessary or sufficient relation is in place (Ragin 2006). If that argument about fuzzy sets was correct, QCA researchers should calibrate variables into fuzzy sets because more rigorous tests are preferable to easier ones (King, Keohane and Verba 1994:chapter 1).

In this paper, I demonstrate that the choice between crisp and fuzzy sets has *ambiguous* consequences because fuzzy-set consistency scores are not necessarily more conservative. All four scenarios are possible in a comparison of crisp and fuzzy sets: first; the set relation is consistent, regardless of whether one uses crisp or fuzzy sets; second, the relation is not consistent for any set type; third, we observe a set relation using crisp sets, but not fuzzy sets; fourth, the fuzzy-set relation is consistent while the crisp-set relation is not. Based on a brief introduction to the parameter of consistency (section 2), I demonstrate and explain in section 3 that the ambiguous consequences of the set choice are rooted in the construction of the consistency formula (Ragin 2006). The consistency measure assigns different weights to consistent and inconsistent cases in a crisp-set and a fuzzy-set analysis. The unequal weighting is intentional and seen as an advantage of the consistency score over previous measures (Ragin 2006:295). The weights are limited to 0 and 1 in a crisp-set study. In fuzzy-set QCA (fsQCA), the weights can vary between 0 and 1, reflecting partial set membership values. The discrepancy between the weights in crisp-set QCA (csQCA) and fsQCA accounts for the indeterminate relationship between the set type and consistency score.

In section 4, I compare the crisp-set and fuzzy-set consistency values of 267 truth table rows collected from ten published QCA articles. The results demonstrate that the ambiguous relationship between set types and consistency scores is present in empirical data. I outline the implications of this finding in the concluding section. First, fuzzy sets contain more information than crisp sets, but this does not necessarily mean that fsQCA is more rigorous than csQCA. Second, if one opts for one type of set and fails to find a consistent set relation, nothing speaks against running a second analysis with the other type of set *as long as* one is transparent about it and reports the results of both studies. Empirical researchers should only conclude that no set relation exists for the relationship of interest if no consistent relation is found with either set type.

[2] A fourth type is the generalized set (Thiem 2014), which I omit from the following discussion. To my knowledge, it has never been used in empirical work.

2 Consistency in Qualitative Comparative Analysis

The consistency score, sometimes referred to as an inclusion score, plays a central role in empirical QCA work. Researchers use it as one, if not exclusive, criterion for designating an observed set relation as consistent with the claim that it represents a relation of sufficiency. Other criteria, which have been more recently developed, are that there should be at least one typical case for the set relation of interest and that set membership values should not be skewed (Schneider and Wagemann 2012:chapter 9). I focus on the consistency parameter in the following because a sufficiently high consistency score is required for the calculation of coverage as a measure for the empirical importance or relevance of a set relation (Ragin 2006). One should only ask about the a set relation's relevance after having inferred that a set relation is given in the first place.

In the following, I focus on sufficient relations because this is the primary focus of most QCA articles and the articles that I reviewed. A complete discussion and analysis of necessary relationships can be found in the online supplement. [3]. For a set relation of sufficiency, the formula is:

$$\frac{\sum_{i=1}^{N} min(x_i, y_i)}{\sum_{i=1}^{N} x_i} \qquad (1)$$

Following the definition of sufficiency that the term is a subset of the outcome, formula 1 measures the share of the cases' total membership in the condition that is implied by the cases' total membership in the outcome. The convention is that the value should be 0.75 or more to designate the observed relationship as being consistent with a claim of sufficiency (Schneider and Wagemann 2012:127-129).

The idea that cases should have unequal weight in calculating consistency is integral to the consistency formula (Ragin 2006; Schneider and Wagemann 2012:sections 5.2 to 5.4). The larger the difference between its membership in the term and the outcome, the more weight an inconsistent case receives. For consistent cases, higher membership in the term also means that their weight increases. The higher the membership in the term, the less likely it is that the case has a membership in Y that is at least as large. By the same token, a case that has low membership in the term is likely to display a set membership in Y that qualifies it as a consistent case. In an analysis of sufficiency, if a case has a membership of 0.1 in the term, any membership value of 0.1 or more in the outcome makes it a consistent case. In contrast, a case with a membership of 0.9 in the term is only consistent when it has an outcome membership of at least 0.9. For this reason, a case with high membership in the term should count more in calculating consistency than a low-membership case.

3 An analytical comparison of crisp-set and fuzzy-set consistency for sufficiency

In empirical and methodological work on QCA, there has been little consideration of the choice between crisp and fuzzy sets given that such a choice is possible. I reviewed 26 QCA articles published in 2016 to collect information on the use of set types and the choice between crisp and fuzzy sets (see the online supplement for the references). The articles were selected based on a topic search in the Web of Science database: "QCA" OR "Qualitative Comparative Analysis" OR

[3] The reproduction material is accessible here: `http://doi.org/10.17605/OSF.IO/FYQX4`

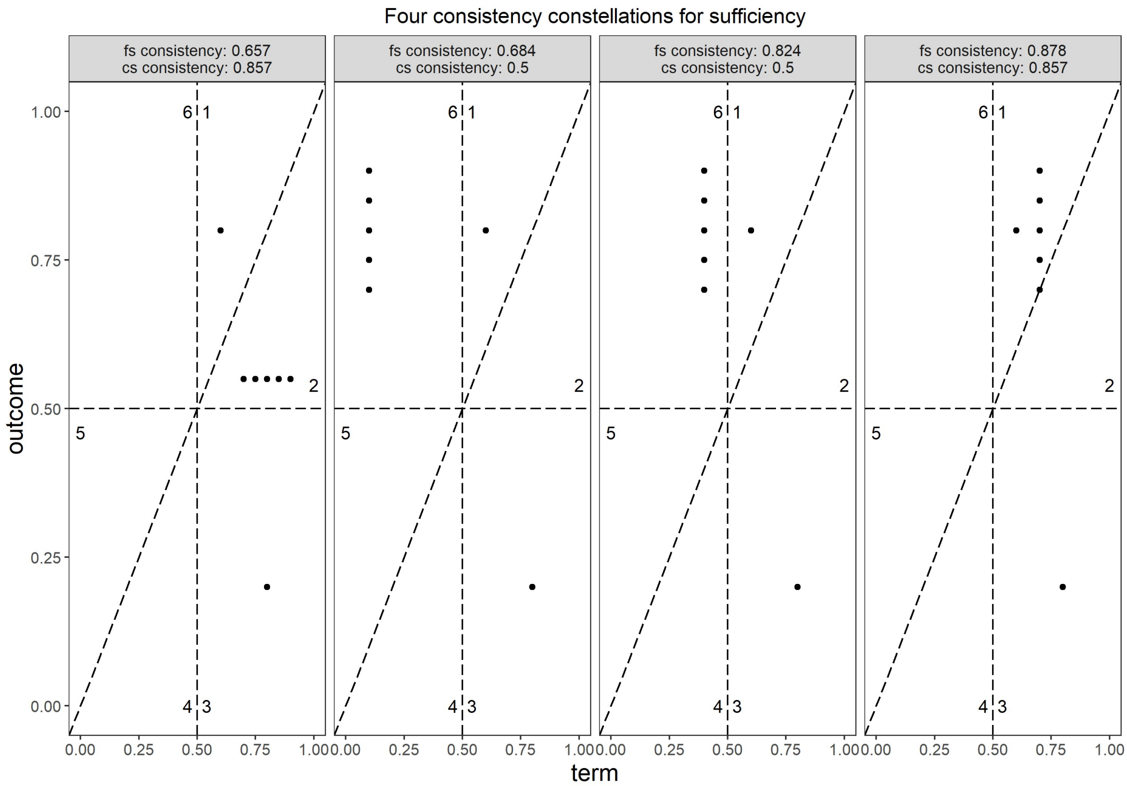
"fsQCA" OR "fuzzy set QCA" OR "fuzzy-set QCA". I limited the search to the fields of International Relations, Political Science, Public Administration and Sociology. The 26 articles are all empirical articles from 2016 that belong to these fields. The type of study was coded based on the author's own description of the study as a crisp-set or fuzzy-set analysis. The codes are based on my reading of the analysis and are exploratorily derived. For crisp-set studies, I looked at the outcome concept of a study and decided whether it would have been possible to capture the concept with a fuzzy set. Based on my reading, the choice of the set type for the outcome is decisive as to whether researchers call a study "crisp" or "fuzzy" and for how the other sets are conceptualized. In the appendix, I explain in more detail what fuzzy set could have been used instead of a crisp set.

Table 1: Review of 26 articles published in 2016

article	type of study	reason for set	fuzzy set possible
Andrews et al.	fuzzy	partial membership	-
Berlin	fuzzy	none	-
Beyens et al.	crisp	sample size	no
Borgna	fuzzy	partial membership	-
Cassani et al.	fuzzy	partial membership	-
Cristofoli et al.	fuzzy	partial membership	-
Dixon et al.	crisp	none	yes
Elliott et al.	fuzzy	none	-
Fagerholm	crisp	theory	yes
Grant et al.	fuzzy	none	-
Hanley et al.	fuzzy	none	-
Hinterleitner et al.	fuzzy	none	-
Jano	fuzzy	none	-
Keudel-Kaiser	crisp	none	no
Kim et al.	fuzzy	none	-
Kolkmann et al.	crisp	none	yes
Kort et al.	fuzzy	partial membership	-
Laux	fuzzy	none	-
Li	crisp	calibration challenges	yes
Lucidarme et al.	crisp	none	yes
Ortiz et al.	crisp	transparency	yes
Pullum	fuzzy	none	-
Schwarz	fuzzy	none	-
Stevens	fuzzy	none	-
Wang	fuzzy	none	-
Zimmermann	fuzzy	none	-

The review shows that 17 articles do not discuss the choice of the set type. This is a problem of itself because the choice of a set type is a research design decision that should be justified. The reasons for the choice of the set type vary among the nine articles that do discuss them. The most common argument for fuzzy sets is that they capture partial set membership values and capture more information than crisp sets (five articles).[4] In the methods-oriented literature, the choice between crisp and fuzzy sets is rarely discussed. It has been argued that fuzzy sets contain more information and therefore produce more conservative consistency values, that is, they make it harder to achieve sufficiently high consistency values (Schneider and Wagemann 2007:225-228; 2012:68-69). The reasoning is based on a plot similar to figures 1 and 2 with six zones numbered clockwise from 1 to 6. The numbering and case labels I use in the following have been introduced by Schneider and Rohlfing (2013). Figure 1 demonstrates with four hypothetical, stylized distributions of cases that the relationship between the set type and the consistency value is ambiguous.

Figure 1: Possible constellations between crisp-set and fuzzy-set consistency for sufficiency



The left panel represents the argument that fuzzy sets produce more conservative consistency measures because only one typical case out of six is typical in csQCA and fsQCA. In the second panel, neither crisp sets nor fuzzy sets achieve a consistency value of more than 0.75. The third panel captures a constellation where the crisp set score is more conservative than the fuzzy-set consistency value. For reasons of completeness, the panel on the right shows that it is also possible to derive crisp-set and fuzzy-set consistency values that are above 0.75.

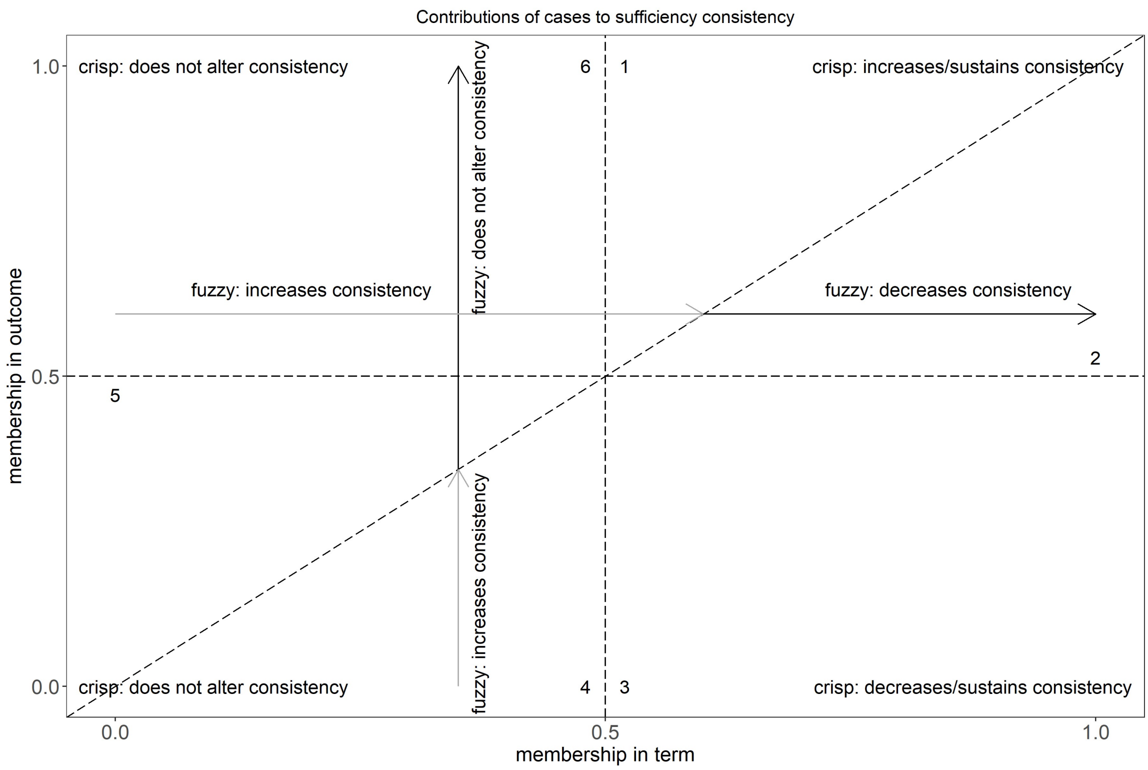
The four panels indicate that the fuzzy-set consistency value depends on the exact location of cases in a so called enhanced XY plot. This sensitivity derives from the weight that the consistency formula assigns to cases when working with fuzzy sets. In figure 2, I compare the weights of cases in csQCA and fsQCA in the order in which the six cells are numbered. I exclude the fuzzy-set scenario in which cases are full members or non-members of the term and the outcome because

4 A review of calibration strategies in QCA research by de Block and Vis (2018) can be read in similar ways.

there is no difference to a crisp-set study. I recommend taking a look at the online supplement containing animated plots that illustrate the following arguments.

In csQCA, there are only four possible constellations. Consistent or typical cases are located in the upper-right corner of figure 2). In a fuzzy-set perspective, the upper-right cell is cut into zones 1 and 2. Zone 1 above the diagonal contains typical cases consistent with a claim of sufficiency. In comparison with csQCA, these typical cases contribute *less* to the consistency formula because membership in the term and the outcome is below 1. Cases below the diagonal in zone 2 qualify as inconsistent or deviant for consistency in degree in fsQCA. These cases also make a fuzzy-set analysis more conservative because they reduce consistency in fsQCA and contribute positively to consistency in csQCA (see section 2). Among all cases in cell 2, the biggest discrepancy between csQCA and fsQCA occurs for cases that are full members of the term and just so members of the outcome.

Figure 2: Comparison of consistency for crisp and fuzzy sets for sufficiency



In csQCA, deviant cases for consistency in kind in cell 3 reduce consistency because they contribute 0 to the numerator and 1 to the denominator of the consistency formula. In fsQCA, cases in cell 3 are also deviant in kind, but contribute more than 0 to the numerator and less than 1 to the denominator. A case in cell 3 therefore reduces the fuzzy-set consistency score less than it does in a crisp-set analysis, making the fuzzy-set consistency score less conservative. The vertical arrow shows that the degree to which fuzzy-set consistency is reduced by deviant cases for consistency in zone 3 (and 2) depends on the case's membership in the outcome. Keeping membership in the term and the membership values of all other cases fixed, the consistency value increases, the higher the case's membership in the outcome up to the point where membership in X and Y becomes equal and the case becomes a typical case. The consistency score does not change regardless of how much larger the membership of a typical case in Y is compared to its membership in the term.

Cases in the left half of the plot are formally consistent with a claim of sufficiency in csQCA. However, they do not affect the consistency score because their contribution to the numerator and denominator of the formula is 0. A more nuanced perspective is needed for the discussion of fsQCA. Cases in cell 4 decrease fuzzy-set consistency because they add less to the numerator than the denominator. For this reason, cases in cell 4 make the fuzzy-set value more conservative than the crisp-set value. Cases in zones 5 and 6 are consistent cases in fsQCA. They increase the fuzzy-set consistency score and make it less conservative in comparison with csQCA.

The horizontal arrow shows the way in which the consistency score depends on the case's membership in the term, keeping its membership in the outcome and all other cases' membership values fixed. The larger the membership of a consistent case in the term (cells 1, 5 and 6), the higher its contribution to the numerator and denominator and the higher the fuzzy-set consistency value. In total, the comparison of the crisp-set and fuzzy-set perspective shows that there is no straightforward relationship between the type of set and the consistency value. Whether the crisp-set or fuzzy-set consistency score is higher depends on the exact location of the cases in the XY plot.

4 A comparison of crisp and fuzzy sets in empirical studies

Based on the analysis of Web of Science, I choose ten empirical QCA studies published in 2016 that use fuzzy sets and for which the relevant part of the analysis could be reproduced (see the appendix). Crisp-set studies have to be set aside because one can transform fuzzy sets into crisp sets by using the cross-over point of the fuzzy sets for crisp-set calibration, but not the other way round. Ten QCA articles might seem like a small number, but the unit of analysis is the consistency value for individual truth table rows as non-minimized sufficient or non-sufficient terms of the outcome. For all ten articles, I am able to extract consistency values for 267 truth table rows for the positive and negated outcome. The fuzzy-set truth tables are derived by calibrating fuzzy sets using the calibration information presented in the articles. The crisp-set truth tables are based on sets that are calibrated by only using the cross-over point of the fuzzy-set calibration. Since all cases are members of the same row in a crisp-set and fuzzy-set truth table, it is possible to compare the consistency values for the same rows and determine whether the crisp-set or fuzzy-set value is higher (if any).

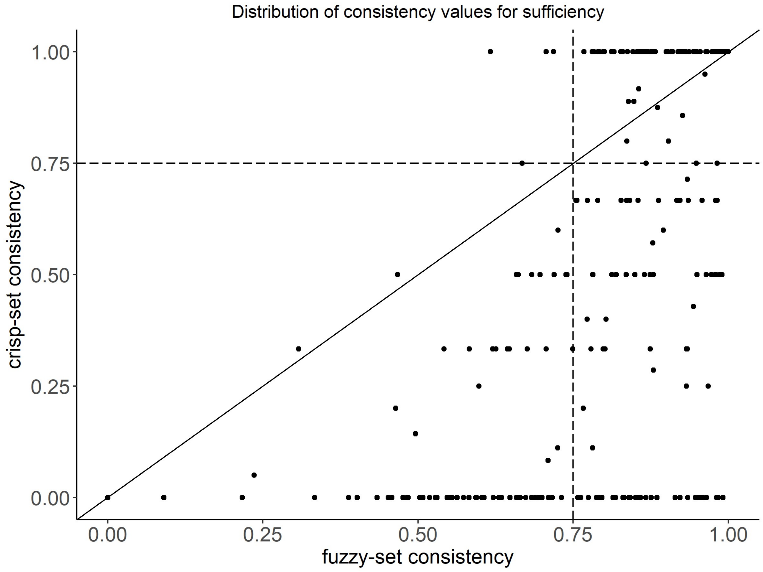
I do not derive a solution from a crisp-set or fuzzy-set truth table for two reasons. First, the main exercise of this section is to compare the consistency values for crisp sets and fuzzy sets, which does not require producing the results of a full-fledged truth table analysis. Second, if the consistency score of the same row differs, it is likely that the findings of an empirical study do differ because different input is likely to produce a different solution.[5] This is particularly likely to be the case if the consistency score for one set type is below the conventional minimum and above it for the other set type.

Figure 3 presents the consistency values of fuzzy-set and crisp-set truth tables. The diagonal conveys whether the crisp-set or the fuzzy-set consistency score is higher. The distribution supports the argument that the relationship between the set type and consistency score is ambiguous because cases are located above the line - crisp-set consistency is higher - and below the line - fuzzy-set consistency is higher. Of all 267 truth table rows, 82 have a larger crisp-set value and about twice as many, 167, have a larger consistency score in the fuzzy-set analysis. These numbers are also based on rows that have a value of less than 0.75 for both set types and would be set aside as not consistent in either analysis. When I limit the comparison to rows that are above 0.75 for both set types (the upper-right quadrant in figure 3), 76 rows have a crisp-set consistency value that is larger than the fuzzy-set score (rows above the diagonal in the quadrant). Depending on where an empirical researcher would fix the consistency threshold for assigning outcome values to truth table rows,

[5] It is likely that the solutions differ, but form set relations with each other (Schneider and Wagemann 2012:section 11.2.3).

these are rows one could assign as consistent in a crisp-set perspective and not consistent in a fuzzy-set perspective. This would occur if the threshold would be set at a consistency value that lies between the crisp-set and fuzzy-set consistency value of a truth table row. In contrast, only eight rows have a fuzzy-set consistency value that exceeds the crisp-set score.

Figure 3: Comparison of sufficiency consistency for truth table rows (n=267)



In a third step, it is valuable to look at the cases that are in the lower-right and upper-left quadrants of figure 3. The upper-left quadrant includes four rows that have a consistency value of at least 0.75 for crisp sets and less than 0.75 for fuzzy sets. These rows fall below the conventional minimum of 0.75 and should always be designated as not consistent for the outcome. In a crisp-set analysis, in contrast, these four rows could be taken as consistent conditional on the chosen consistency threshold. This is certainly the case for three rows because they have a consistency score of 1. In the lower-right quadrant, we find 86 rows with a fuzzy-set consistency of more than 0.75 and less than 0.75 for crisp sets. As holds true for crisp sets, it is very likely that empirical researchers would designate some of these rows consistent and include them in fsQCA and exclude them as inconsistent in csQCA. The analysis of empirical data from published QCA articles demonstrates that there is no clear-cut relationship between the set type and the consistency values one derives for crisp-set and fuzzy-set relations. Figure 3 shows that this is likely to produce divergent results and inferences because crisp-set and fuzzy-set consistency values can fall on different sides of the conventional thresholds.

5 Conclusion

The ambiguous relationship between crisp sets and fuzzy sets and their consistency scores has two implications for QCA research. First, no set type produces more conservative consistency values *per se*. Empirical researchers who decide for fuzzy sets should not believe that the higher degree of information that fuzzy sets capture constitute more rigorous tests. This is a difference to quantitative research for which the dichotomization of continuous variables is discouraged (MacCullum, Zhang, Preacher and Rucker 2002). Second, my review of QCA studies and the review by de Block and Vis (2018) indicate that empirical researchers do not usually have theoretical reasons to prefer one set type over another; that decision is justified on different ground, if it is justified at all (see table 1).

This a legitimate strategy as long as researchers understand that the *failure* to find a consistent set relation with one type of set allows for the possibility of achieving consistency using the other type of set. If a fuzzy-set analysis does not produce a solution because all truth table rows are inconsistent, the proper conclusion is that there is no *fuzzy*-set relation because there could still be a consistent crisp-set relation (and vice versa). Inferring that there is no set relation present could mean committing a false-negative conclusion because one might find evidence for a set relation when working with crisp sets. The assumption that there is no set relation *per se* is only correct if a crisp-set *and* the fuzzy-set analysis do not produce a solution.[6]

For empirical QCA research, the implication is that one should be open to the possibility of making different findings with different set types. In a *first* step, a researcher should make an informed decision between crisp and fuzzy sets (and, possibly, multivalue sets). This decision could be based on the nature of the concepts (inherently binary vs gradual), the availability of sufficient data for the calibration of fuzzy sets and the research interest in set relations between differences in kind (crisp) as opposed to differences in degree (fuzzy). If a consistent set relation is found with the preferred set type, there is no need to also work with the second set type and the empirical analysis can be terminated. The generated result stands for itself; it can neither be invalidated when one cannot derive a solution using the other set type, nor does it become stronger if one is able to derive a solution for both set types. If, however, no consistent set relation is observed with the initially chosen set type, one can run the same analysis again with the other set type in a *second* step. The second step automatically turns the study into an *exploratory* analysis because change of the set type and continued analysis of the data has its cause in the first-stage finding of a result that is too inconsistent (a "null finding") and the goal of getting a consistent result. It is *indispensable* that an empirical researcher makes the two-step procedure transparent because the reader needs to know how much and in what way the data was probed before one was able to derive a consistent set relation from the data (Nosek and Lakens 2014). This two-step procedure ensures that one avoids false-negative conclusions about the presence of a set relation when the relation is in place for one type of set but not the other, and one happened to start the analysis with the type for which no solution can be derived.

References

de Block, Debora and Barbara Vis. 2018. "Addressing the Challenges Related to Transforming Qualitative into Quantitative Data in Qualitative Comparative Analysis". *Journal of Mixed Methods Research* advance access.

King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Nosek, Brian A. and Daniël Lakens. 2014. "Registered Reports". *Social Psychology* 45 (3): 137-141.

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

Schneider, Carsten Q. and Ingo Rohlfing. 2013. "Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research". *Sociological Methods & Research* 42 (4):559-97.

Schneider, Carsten Q. and Claudius Wagemann. 2007. *Qualitative Comparative Analysis (QCA) und Fuzzy Sets. Ein Lehrbuch für Anwender und alle, die es werden wollen*. Barbara Budrich.

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

[6] Depending on the distribution of the cases, it might also be that only a specific mix of crisp and fuzzy sets might yield a solution.

Thiem, Alrik. 2013. "Clearly Crisp, and Not Fuzzy: A Reassessment of the (Putative) Pitfalls of Multi-Value QCA". *Field Methods* 25 (2):197-207.
Thiem, Alrik. 2014. "Unifying Configurational Comparative Methods: Generalized-Set Qualitative Comparative Analysis". *Sociological Methods & Research* 43 (2):313-37.
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):265-89.
Vink, Maarten P. and Olaf van Vliet. 2013. "Potentials and Pitfalls of Multi-Value QCA: Response to Thiem". *Field Methods* 25 (2):208-13.