Making degreeness count in QCA: Problems with fuzzy sets and a linguistic solution

International Journal of Social Research Methodology

ISSN: 1364-5579 (Print) 1464-5300 (Online) Journal homepage: www.tandfonline.com/journals/tsrm20

Making degreeness count in QCA: problems with fuzzy sets and a linguistic solution

Roel Rutten

To cite this article: Roel Rutten (31 Aug 2025): Making degreeness count in QCA: problems with fuzzy sets and a linguistic solution, International Journal of Social Research Methodology, DOI: 10.1080/13645579.2025.2554825

To link to this article: https://doi.org/10.1080/13645579.2025.2554825

Published online: 31 Aug 2025.

INTERNATIONAL JOURNAL OF SOCIAL RESEARCH METHODOLOGY
https://doi.org/10.1080/13645579.2025.2554825

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Making degreeness count in QCA: problems with fuzzy sets and a linguistic solution

Roel Rutten [a,b]

[a]Tilburg School of Social and Behavioural Sciences, Tilburg University, Tilburg, The Netherlands; [b]ERAC (European Regional Affairs Consultants), Eindhoven, The Netherlands

ABSTRACT

QCA prefers fuzzy sets because they capture all empirical variation. However, fuzziness plays no role in minimizing the truth table, which follows threshold (crisp) logic also in fsQCA. Fuzzy set values only matter for calculating set-relationships, where they present a problem. Fuzzy set-relationships are calculated over all cases, whereas crisp set-relationships only consider cases that provide corroborating or contradicting within-case (i.e. crisp) causal evidence. When developing causal explanations, fsQCA thus dialogues incompatible within-case and cross-case causal evidence. In response, this paper suggests capturing degreeness with linguistic hedges – e.g. somewhat, moderately, considerably and very – but calculating crisp set-relationships. This avoids incompatible within-case and cross-case causal evidence. The crossover point sits between intensifying and diluting hedges, but researchers can make degreeness count by setting lower or higher crossover points to investigate (i) necessity in degree, (ii) explanatory power of individual conditions and (iii) differences between solutions for higher and lower crossover points.

ARTICLE HISTORY
Received 9 December 2024
Accepted 22 August 2025

KEYWORDS
Calibration; fuzzy sets; linguistic hedges; QCA; variation

Most researchers, inside and outside QCA (Qualitative Comparative Analysis), agree that fuzzy sets more accurately capture the degreeness of social reality than do crisp sets (Ragin, 2008, p. 141). On this notion, the overwhelming majority of QCA studies nowadays use fuzzy sets (Rohlfing, 2020). However, QCA remains a threshold method also in its fuzzy guise (Goertz, 2020, p. 182). The degreeness that fuzzy sets capture does not play a role in the truth table analysis. Fuzzy set QCA (fsQCA) does not actually deliver on its promise of making degreeness count. Also, fsQCA follows a threshold logic for which all set-membership values >0.5 are considered 'true' and all those <0.5 'false' (Ragin, 2008, p. 138). Moreover, fuzzy sets in QCA are not fuzzy in the sense that Zadeh (1965) intended them. Fuzzy sets (of the Zadeh-kind) model vagueness, a non-probabilistic kind of uncertainty (Zadeh, 1972). However, calibration (to precise) eliminates fuzziness (vagueness). A calibrated fuzzy set is, in fact, an

CONTACT Roel Rutten r.p.j.h.rutten@tilburguniversity.edu Tilburg School of Social and Behavioural Sciences, ERAC (european Regional Affairs Consultants), PO Box 90153, Tilburg 5000 LE, The Netherlands

oxymoron (Smithson, 2016); however, it is what makes fuzzy sets compatible with QCA's threshold logic (Ragin, 2000, p. 170) and allows cross-case comparison. Nor does fuzzy logic play any role in QCA. QCA's solution makes propositional if, then-statements (Ragin, 2008, p. 153). The only way in which the degreeness that fuzzy sets capture plays a role in QCA is in the calculation of set-relationships. Contrary to crisp set-relationships, fuzzy set-relationships are calculated over all cases in the population and not just those in the 'relevant cells' (e.g. the (1,1) and the (1,0) cells for sufficiency) (Ragin, 2008, pp. 39–41). However, this conflating of logics leads to a number of problems that have hitherto gone (largely) unnoticed in QCA. Nor is the marriage between fuzzy sets and QCA's threshold logic a very fortunate one. This paper outlines the problems of fuzzy sets in QCA and suggests an alternative to actually make degreeness count in the form of linguistic hedges.

Making a successful case for problems with fuzzy sets in QCA depends on taking an ontological perspective. Fuzzy sets in QCA are valued for their capturing all empirical variation, which is an epistemological argument (Pula, 2021; Rohlfing, 2020). At the same time, calibration in QCA distinguishes between relevant and irrelevant variation, which is an ontological argument (Ragin, 2000, p. 169; Rutten, 2023). It asks: What does it mean to be a case of X and then identifies an empirical value that corresponds to that meaning. QCA's fuzzy sets define just three thresholds (the threshold of full non-membership, the crossover point and the threshold of full membership) and assume all variation between them to be equally relevant (Ragin, 2000, pp. 161–162). However, the degreeness of social reality implies that cases may also be qualitatively different from one another between these three thresholds. Ontologically, more semantically relevant thresholds may be defined (Goertz, 2020, pp. 179–183; Lakoff, 2014, p. 11), and not all empirical variation between full non-membership and full membership may be analytically relevant (Lakoff, 2014, p. 13). That is, fuzzy sets bring some of the very empiricist, variable-based thinking into QCA that the comparative method opposes (Finn, 2022; Porpora, 2018; Pula, 2021; Ragin, 1987/2014, pp. 165–166). This paper suggests linguistic hedges as a superior alternative to fuzzy sets in QCA. Linguistic hedges are qualitative thresholds that form an ordinal scale, as opposed to fuzzy sets behaving as ratio scales. For computational purposes, linguistic hedges are crispified to conform to QCA's threshold logic, while fsQCA problematically conflates fuzzy (degreeness) and crisp (threshold) logic.

To make a stronger case for problems with fuzzy sets, the paper starts by reiterating QCA's logic of causal inference and how it investigates qualitative differences (thresholds) rather than gradual ones (degreeness). The paper then identifies four problems of fuzzy sets in QCA that all hinge around fuzzy sets behaving as ratio scales. The next section, on measurement and calibration, explains how calibration is always about setting thresholds and that calibrating linguistic hedges is no different. However, minding the problems with fuzzy sets in QCA, linguistic hedges are not assigned set-membership values. The following section gives guidance on how to actually calibrate linguistic hedges. The subsequent section explains how linguistic hedges allow QCA studies to account for degreeness in causal explanation. The conclusion draws the argument of the paper together. But however one thinks about thresholds and degreesness, the fact remains that fuzzy set values do not play a role in minimizing the truth table in fsQCA. In response, this paper develops an alternative in the form of linguistic hedges to make degreeness count in QCA.

Causal explanation in QCA

QCA is a comparative case study method. It looks for invariances between cases with the same outcome and applies difference-making logic to minimize explanatory conditions (Rubinson et al., 2024). For example, suppose a case has characteristics A, B and C and also outcome Y. So the logical proposition: if [A•B•C] then Y, is true. (Where • means logical AND.) Suppose another case has conditions A and B but not C and also outcome Y. Now the logical proposition: if [A•B•~C], then Y is also true. Applying propositional logic, both statements can be simplified to: if [A•B] then Y. When conditions [A•B] are present, having (C) or not-having (~C) condition C does not matter for having outcome Y. The invariance across cases is between [A•B] and Y, while [C OR ~C] is logically redundant. How do we know that A, B, C and Y are present or absent? Because the degree to which cases have the characteristics of A, B, C and Y is above or below a threshold set by the researcher (Rubinson et al., 2024). Put differently, a case is a case (an instance) of Y if it has enough of the characteristics of Y (epistemology) to qualify as an instance (a case) of Y (ontology) (Rutten, 2024, p. 122). In other words, calibration makes QCA a threshold method. Calibration decides whether a case 'sits' above or below the threshold (the crossover point). A QCA study investigates whether being above or below the threshold for X makes a difference for being above or below the threshold for Y (Ragin, 1987/2014, p. 86). Also in fsQCA, the crossover point (threshold) is what counts analytically. Having a degree of membership above the crossover point (>0.5) establishes a case as an instance of Y (true), while a degree of membership <0.5 establishes it as not-an instance of Y (false) (Ragin, 2008, p. 138). The crossover point in fsQCA distinguishes between relevant and irrelevant variation in the same way as 1s and 0s do in crisp set QCA (csQCA).

Setting a crossover point does not deny that social reality is experienced in degrees. Being not a case of Y does not mean that the case has none of the characteristics of Y. Just not enough to qualify as an instance of Y. Social reality is not dichotomous but it is full of relevant thresholds (Vis, 2012). Many concepts are at once dimensional and categorical (Smithson, 2016). All households have a degree of income but either they make ends meet at the end of the month or they do not. All people have a degree of narcissism, but only beyond a threshold do we recognize someone as a narcissist. Being above or below a threshold makes a substantial difference for how cases (households, individuals) fare. It is therefore analytically relevant to ask the question: Does being above the threshold for X make a difference for being above the threshold for Y? Put differently, QCA investigates qualitative differences, not gradual differences as do statistical methods, and this is true for both csQCA and fsQCA.

Another way of saying this is that singular (within-case) causal evidence in QCA is crisp. Cases corroborate or contradict the statement: if X, then Y based on their 'sitting' above or below the thresholds for X and Y (Beach & Pedersen, 2018; Goertz, 2017, pp. 49–52). On this (crisp) basis, cases are assigned to truth table rows. Membership in degree allows cases to sit anywhere in the property space (Ragin, 2008, p. 128) but they are assigned to truth table rows (the end points of the property space) based on their sitting above or below the crossover point (Schneider & Wagemann, 2012, p. 98). That is, cases populating a truth table row provide singular causal evidence corroborating or contradicting the logical proposition that the row makes based on their 'sitting' above or

below the crossover point (threshold) (Ragin, 1987/2014, pp. 87–89). The subset consistency of the row produces general (cross-case) causal evidence for the row (the statement it makes). Declaring a row 'true' (1) or 'false' (0) follows from dialoguing singular (cases sitting above or below the threshold for Y) and general (subset consistency) causal evidence (Ragin, 2008, p. 130). After the truth table minimization, researchers 'go back' to the cases to connect the configurations in the solution (i.e. general causal evidence in the form of cross-case patterns) to singular causal evidence (provided by the cases that a configuration covers) (Beach & Rohlfing, 2018). The truth table minimization abstracts from singular causal evidence and merely describes cross-case patterns. Going back to the cases, researchers must make it plausible that singular causal evidence from the cases corroborates the general causal claim implied by 'their' configuration. But researchers go back to cases that provide corroborating or contradicting (crisp) singular causal evidence (Ragin, 2008, p. 121). This dialogue of singular and general causal evidence is critical to QCA's approach to causal explanation (Finn, 2022). It is what makes QCA a case-based method and what distinguishes it from quantitative (variable-based) methods that heavily prioritize general causal evidence (Rutten, 2022, p. 1216). But here is also where csQCA and fsQCA deviate, because crisp and fuzzy sets produce different kinds of general causal evidence.

Problem 1: conflating crisp and fuzzy logic

Because they are calculated over different cases, crisp and fuzzy subsets are different things. A fuzzy subset-relationship captures the distribution of cases to the left and right of the X = Y-diagonal of an XY-plot. Instead, a crisp subset-relationships expresses the proportion of corroborating and contradicting cases in the (1,1) and (1,0) cells of a 2×2 table (Ragin, 2008, pp. 39–41). A case contributes to fuzzy subset consistency when $X \leq Y$ and this produces three important discrepancies with crisp subset consistency. (i) Cases inconsistent in degree ($X > Y$ but X and $Y > 0.5$) reduce fuzzy subset consistency even though their singular causal evidence corroborates the logical proposition: if X, then Y – because both X and Y 'sit' above the threshold. (ii) Equifinal cases ($X \leq Y$ but $X < 0.5$) contribute to fuzzy subset consistency even though their singular causal evidence contradicts the logical proposition: if X, then Y – because X 'sits' below the threshold. Ontologically, X is not present. (iii) Irrelevant cases (both X and $Y < 0.5$) contribute to fuzzy subset consistency when $X \leq Y$. However, these cases neither corroborate nor contradict the logical proposition: if X, then Y – because, ontologically, X is not present and Y has not occurred (Rutten, 2023).

Consequently, fsQCA conflates crisp and fuzzy logic. Cases are assigned to truth table rows on the basis of their crisp singular causal evidence (i.e. to which corner of the property space they 'belong'). However, the subset consistency of truth table rows in fsQCA (producing general causal evidence) is calculated using fuzzy-set logic. The same goes for the solution (Ragin, 2008, p. 140). However, researchers 'go back' to the cases to look for corroborating singular causal evidence on the basis of crisp (threshold) logic (Schneider & Wagemann, 2012, p. 307). Recognizing the problem, fsQCA distinguishes between cases 'inconsistent in degree' and cases that are logically contradictory (Schneider & Wagemann, 2012, p. 127). However, this distinction only underlines that fsQCA collapses two different logics. From a crisp (threshold) logic perspective, it makes

no sense to infer corroborating general causal evidence from equifinal and irrelevant cases. Such cases have nothing to say about the logical proposition: if X, then Y (Schneider & Wagemann's, 2012, pp. 191–193) follow an entirely fuzzy argument, suggesting that cases inconsistent in degree *artificially* increase consistency if interpreted as supporting an if, then-statement (emphasis added).) To the extent that causal inference in QCA is about dialoguing singular and general causal evidence (Ragin, 2008, p. 121), fsQCA is dialoguing two different logics. Its singular and general causal evidence are not compatible (Beach & Pedersen, 2018; Rutten, 2024, p. 139). The observation that the fuzzy set of X ⊆ the fuzzy set of Y does not corroborate the proposition that: if X, then Y.

Problem 2: assumed linearity of fuzzy sets

All fuzzy set scales behave as ratio scales (Ragin, 2000, p. 155; Smithson & Verkuilen, 2006, p. 44). Therefore, equidistance between cases in set-membership values assumes semantical equidistance (Rutten, 2024, p. 133; Smithson, 2016, p. 179). The difference between fuzzy set-membership values 0.1 and 0.2 is assumed to have the same semantical meaning as the difference between 0.7 and 0.8. Fuzzy set-relationships are calculated on this basis. However, this assumption is often highly questionable – even with ratio scale source variables. Say a monthly income of £1,000 equals 0.1 membership in the fuzzy set of rich people, £2,000 equals a fuzzy set-membership of 0.2, £7,000 equals 0.7 and £8,000 equals 0.8. The numerical distance between 0.1 and 0.2 versus 0.7 and 0.8 is the same. However, semantically, the difference between £1,000 and £2,000 is of a qualitatively different nature than the difference between £7,000 and £8,000. Going from £1,000 to £2,000 is the difference between having and not having to rely on food banks. Going from £7,000 to £8,000 is the difference between buying a BMW 4-series or a 5-series. To the extent that equating numerical and semantical equidistance is flawed, calculating fuzzy set-relationship is an exercise in spurious specificity.

Another way of putting this is that many QCA researchers seem to think of fuzzy sets as calibrated variables (Lakoff, 2014; Pappas & Woodside, 2021; Smithson & Verkuilen, 2006, pp. 15–16). They approach fuzzy sets epistemologically as if their aim is to capture all empirical variation between their 'floor' and 'ceiling' (Ragin, 2008, p. 33). However, semantically, fuzzy sets in QCA are only concerned with three thresholds: the threshold of full non-membership (floor), the crossover point and the threshold of full non-membership (ceiling) (Schneider & Wagemann, 2012, p. 28). Yet, the degreeness of social reality suggests there may be (many) more relevant thresholds (Lakoff, 1987, pp. 68–71). Cases just above the crossover point are qualitatively different from cases just below the threshold of full membership. Rich people just above the crossover point can afford regular leisure and luxury but still rely on income from their jobs. Instead, rich people just below the threshold of full membership are financially independent. These differences are better captured as a number of thresholds than as variation on a ratio scale (Mohan, 2019, p. 140). As the above example suggests, numerical equidistance may not suggest semantical equidistance. Calibration attaches empirical values to meaning (Ragin, 2000, p. 9). However, capturing all empirical variation on a ratio scale may not be meaningful (Ragin, 2000, p. 157; Smithson & Verkuilen, 2006, p. 44). There may well be more meaningful thresholds than the ratio scale of a fuzzy set suggests (Lakoff, 2014).

If degreeness in social reality comes in the form of ontologically defined thresholds, the assumption of fuzzy sets that empirical variation can be expressed on a ratio scale is flawed (Smithson & Verkuilen, 2006, pp. 87–88).

Problem 3: QCA's fuzzy sets are not fuzzy

Fuzzy sets of the Zadeh-kind model vagueness. They model a non-probabilistic kind of uncertainty – which may plausibly be expressed on a ratio scale. The statement that today is a warm day or that the meal was good are fuzzy statements because warm does not specify a temperature and rating the quality of a meal is even more imprecise. However, there is nothing fuzzy (in this sense) about fuzzy sets in QCA because they are precisely calibrated (Lakoff, 2014; Smithson, 2016). Ragin, (2000, p. 7), effectively argues this point. To calibrate (to precise) eliminates fuzziness (vagueness), making a calibrated fuzzy set an oxymoron. In fuzzy sets of the Zadeh-kind, 25°C (77 F) may be what most people (in the UK) consider a warm day and, hence, days of 25°C have full (1) membership in the fuzzy set of warm days. But there is still an element of uncertainty; we do not know if warm equals 25°C. However, in fsQCA, the threshold of full membership in the fuzzy set of warm days is precisely calibrated at 25°C. There is no uncertainty about what does or does not count as a warm day (Ragin, 2000, p. 170). In fsQCA, 24°C is calibrated as having 0.9 membership in the fuzzy set of warm days not because of any uncertainty about what counts as a warm day but because a fully warm day has been precisely defined at 25°C. This is the difference between Zadeh's (1972) fuzzy numbers (capturing non-probabilistic uncertainty) and fsQCA's calibrated variables (capturing the degree to which cases have a precisely defined characteristic) (Mohan, 2019, p. 125; Smithson & Verkuilen, 2006, pp. 15–16).

Consequently, there is no point of maximum ambiguity in fuzzy sets of the QCA-kind. In fuzzy logic, there is a temperature where a day could qualify just as well as a warm day as a mild day. This is the point of maximum ambiguity in fuzzy sets of the Zadeh-kind. Such days cannot conclusively be defined as more a member of the set of warm days or more a member of the set of mild days (which makes 0.5 a meaningful fuzzy number). It is the point where days crossover from being more a member of the set of mild days to being more a member of the set of warm days. In fsQCA's calibrated variables, there is no such point (Rutten, 2024, p. 118). Given that, semantically speaking, the crossover point is precisely defined, a day cannot be too mild to qualify as a mild day and yet not warm enough to qualify as a warm day. Consequently, researchers must decide whether, e.g. 20°C counts as more of a -warm day or more of a mild day – whether 20°C is above or below 0.5. This difference between fuzzy sets of the Zadeh-kind and the QCA-kind is illustrated by the fact that, nowadays, QCA researchers are encouraged to avoid calibrating 0.5-cases (Rubinson et al., 2024).

In sum, thresholds in QCA are precisely defined. They are calibrated, which eliminates fuzziness. What QCA calls fuzzy sets are in fact better thought of as crisp sets wherein cases can have precisely defined degrees of membership (Goertz, 2020, pp. 147–149; Ragin, 2000, pp. 154–157; Smithson & Verkuilen, 2006, p. 7) – and this is what enables cross-case comparison. The unambiguous connection between

empirical values and semantical meaning (i.e. calibration) defeats the notion of fuzziness (Lakoff, 2014; Mohan, 2019, p. 48 and, interestingly; Ragin, 2000, p. 7) to instead define multiple thresholds between full non-membership and full membership (Mohan, 2019, p. 126).

Problem 4: QCA's outcomes are threshold, not fuzzy-logic statements

The outcome of an fsQCA study (the solution) still is a number of propositional (crisp) statements. Fuzzy logic is very different. Consider the following examples of fuzzy logic statements:

(i) I am fairly hungry and it is almost dinner time, so I will not have an extra meal now.
(ii) I am very hungry, it is approaching dinner time, so I will have a tiny extra meal now.
(iii) I am considerably hungry and it is nowhere near dinner time, so I will have a good extra meal now.

Even with the fuzzy sets 'hungry', 'dinner time' and 'meal', fsQCA cannot produce statements like these. Because of the truth table minimization following propositional logic, in fsQCA the above statements would become:

(i) not-hungry AND dinner time → not-extra meal,
(ii) hungry AND dinner time → not-extra meal,
(iii) hungry AND not-dinner time → extra meal,

based on the cases sitting above (>0.5) (intensifying hedges) or below (<0.5) (diluting hedges) the threshold. It loses all the nuance of the initial fuzzy logic statements because degrees of membership (captured by linguistic hedges (Zadeh, 1972)) play no role in developing the fsQCA solution (Mohan, 2019, p. 188). They only do so in calculating set-relationships, where they are of dubious merit (see above). Of course, fuzzy sets are epistemologically richer than crisp sets (Ragin, 2008, p. 141). However, this richness is analytically irrelevant in the truth table minimization and may be misleading when developing general causal evidence. Not only can it develop general causal evidence that is incompatible with singular causal evidence from the same cases. It can also produce logical contradictions such as simultaneous subset relations ($X \subseteq Y$ and $\sim X \subseteq Y$). For this, fsQCA has developed the PRI (proportional reduction in inconsistency) (Schneider & Wagemann, 2012, p. 242). However, from a threshold perspective, PRI is merely an epistemological patch for an ontological problem.

In sum, the overarching problem is the assigning of set-membership values to degrees of membership in fsQCA. Calibration helps to know whether a country is rich or poor rather than just richer than another country (Ragin, 1987/2014, xxiv– xxv). However, when it comes to differences in degree between cases, fuzzy sets behaving as ratio scales present a problem. The connection between concepts and measurements, on the one hand, and 'things' (e.g. cases) in the world, on the other hand, is always fraught with difficulties. Concepts and their measurements are human categories that do not neatly

overlap with 'things' in the world (Porpora, 2018; Pula, 2021; Rutten, 2023). Nor are most concepts themselves linear (Lakoff, 1987, pp. 141–148). Something that the QCA literature acknowledges where it suggests defining concepts (sets) from both their positive (X) and negative (~X) poles. For example, poor households (X) are defined in terms of their being unable to make ends meet, while not-poor households (~X) are defined in terms of their enjoying leisure and luxury. Particularly below the crossover point, cases will have characteristics of both X and ~X (Rubinson & Rutten, 2023, p. 342). To express (measurements of) differences between cases ('things' in the world) as fuzzy set-membership values imposes a linearity that does not match with the nature of these differences. Neither conceptually (concepts are not linear) nor empirically (degrees of membership form an ordinal scale). That is, fuzzy set-membership values compound the problem of connecting meaning (ontology) and measurement (epistemology) (Lakoff, 2014). Calibrating linguistic hedges recognizes that cases may differ from one another in kind (also between the fuzzy set-membership values 0 and 1) but rejects fsQCA's added assumption that these differences are meaningfully expressed on a ratio scale.

Measurement and calibration

The fact that most QCA researchers value fuzzy sets more than crisp sets evidences the strength of 'epistemological instincts' among social scientists (Porpora, 2018). The idea that all empirical variation is relevant and must be captured. However, QCA is a threshold method (Goertz, 2020, p. 182), and the function of thresholds is to distinguish between relevant and irrelevant variation – on the basis of the definition of the concept (set) (Ragin, 2008, p. 78; Rubinson et al., 2024). For example, calibrating the set of poor households, the first question to ask is: What does it mean to be a poor household? The answer to the question is a definition (meaning), not an empirical value (measurement). For instance, a poor household is one that cannot make ends meet. The next question is: How do we know a poor household when we see one? This is an empirical question that requires connecting an empirical value to the definition. Put differently, calibration connects measurement to meaning in a going back and forth between measurements and definition (Ragin, 2008, p. 78). The result is a truth value. Applied to a particular household, the statement: This is a poor household, may be true or false (crisp set) or partially true (fuzzy set). The problem, as identified above, is that numerical truth values, i.e. fuzzy set-membership values, do not match with QCA's threshold logic. However, eliminating fuzzy set values does not alter the process of calibration; it remains a dialogue between meaning and measurement (Goertz & Mahoney, 2012; Greckhamer et al., 2018; Ragin, 2000, pp. 160–167).

Calibration may pertain to cases and source variables. Calibrating cases is always 'manual'. In this example, researchers investigate for every household in their case population whether it can make ends meet. Whether it can afford the basic costs of living as the researcher defined them (e.g. food, housing, fuels, clothing). Or the degree to which the household can afford these costs. For example, a fully poor household cannot afford any of the above costs, while a moderately poor household can only afford those basic costs but nothing else. For manual calibration of fuzzy sets, researchers decide the number of degrees of membership (e.g. three, four or six), define them, connect them to empirical values, and assign them a set-membership value. Calibrating source variables,

researchers truncate the variation in the source variable by setting a 'floor' and a 'ceiling' as well as a crossover point. Variation between the floor (the threshold of full non-membership) and the ceiling (the threshold of full membership) is relevant, and variation beyond these thresholds is not (Ragin, 2008, p. 84). Source variables can be calibrated manually; however, the direct and indirect methods of calibration are most commonly used for calibrating source variables (Ragin, 2008, p. 85). The direct method of calibration defines three thresholds, the thresholds of full (non-)membership and the crossover point, and connects each threshold to an empirical value. These values are then entered into a logarithmic function that converts each empirical value into a unique set-membership value. The indirect method defines six thresholds (degrees of membership), i.e. 0, 0.2, 0.4, 0.6, 0.8 and 1, and connects each threshold to an empirical value. It then converts all empirical values between two thresholds into unique set-membership values in a proportionate way.

All three calibration methods connect thresholds (meaning) to empirical values (measurements). However, they have important shortcomings. They assume that differences between thresholds can be meaningfully expressed on a ratio scale (Smithson & Verkuilen, 2006, p. 29). However, numerical equidistance of fuzzy set-membership values does not imply semantical equidistance between various degrees of membership (see above). The problem with the direct and indirect methods of calibration is that source variables may be ratio scaled but degrees of membership are not (Smithson, 1987, p. 27). Manual calibration connects thresholds to degrees of membership in a mostly arbitrary way. Whether a degree of membership, e.g. very poor household, receives set-membership value 0.8, 0.9 or even 0.8695 does not substantially matter for computing fuzzy set-relationships. There is no direct connection between set-membership values and the meaning of the degree of membership; however, fuzzy set-membership values actually accomplish this (Lakoff, 2014; Mohan, 2019; Smithson, 2016). That is, fuzzy sets invite researchers to collapse meaning into measurement and think in terms of empirical variation. Moreover, there may be multiple relevant thresholds between 0 and 1 to indicate qualitatively different states (Finn, 2022; Pula, 2021; Smithson, 1987, p. 80). (See the above poor household example.)

Likert scales are also frequently used as a basis for calibration because they already have meaning (Ragin, 2000, p. 169). In fact, as Likert scales are ordinal, they define qualitative states, i.e. thresholds. However, it is bad practice to mechanically assign set-membership values to the five points of a Likert scale – as in, e.g. 1 = 0, 2 = 0.33, 3 = 0.5, 4 = 0,67 and 5 = 1. First, 0.5-cases are ignored in the truth table minimization, so calibrating them is the QCA equivalent of throwing away data (Rutten, 2024, p. 119). Instead, researchers must decide whether the neutral Likert value is above or below the crossover point (more in than out, or more out than in the set) (Rubinson et al., 2024). For example, a household that neither agrees nor disagrees (neutral value) to the question: This household can make ends meet, is, in fact, a poor household because it cannot make ends meet at least some of the time. Consequently, a five-point Likert scale is calibrated into four degrees of set-membership as, e.g. not-poor household (1), moderately poor household (2 and 3), considerably poor household (4) and fully poor household (5).

That is, calibration is always about setting thresholds by dialoguing meaning and measurement. In that sense, calibrating linguistic hedges is no different from calibrating fuzzy set-membership values. However, when calibrating linguistic hedges, researchers:

- Explicitly acknowledge that qualitative differences (thresholds) between cases form an ordinal scale rather than interval or ratio scale.
- Think carefully about the number of thresholds (linguistic hedges) the definition of their concept (set) plausibly allows. Which, again, is no different from calibrating fuzzy set-membership values (Ragin, 2000, p. 167); however, it challenges the direct connection between empirical values and set-membership values that fuzzy sets suggest (Lakoff, 2014; Smithson & Verkuilen, 2006, p. 44) and their assumption that all variation between 0 and 1 is equally relevant (Lakoff, 2014; Smithson, 1987, p. 27).
- Do not assign set-membership values to linguistic hedges. Instead, for computing set-relationships and to align with QCA's threshold logic, researchers crispify linguistic hedges. Normally, the crossover point sits between diluting and intensifying hedges (Zadeh, 1972). So {not, marginally, somewhat and moderately} become 0 (false) and {considerably, mostly, very and fully} become 1 (true). However, researchers may raise or lower the crossover point to benefit from the nuance of their linguistic hedges (see below).

Calibrating using linguistic hedges

The idea of fuzzifying sets with linguistic hedges was first developed by Zadeh himself (Zadeh, 1972) and elaborated by Lakoff (1973). Also, Ragin (2000, pp. 176–178) cursorily discussed linguistic hedges in connection to calibration but the idea never caught on in QCA. As always, calibration is a dialogue between concepts (meaning) and cases (measurement). A dialogue one can start at either end. Calibrating linguistic hedges, researchers first decide on the number of hedges – which is a function of what is semantically meaningful, not of the fine-grainedness of the source data (Ragin, 2000, p. 167). It is also a compromise between two considerations. First, comparability between cases (concepts, sets). Even though some concepts are more fine-grained than others, for comparison, they all need to be described by the same linguistic hedges (Rutten, 2022; Smithson & Verkuilen, 2006, p. 27). Also, when using sets with different degrees of membership in the same study (see below). However, the number of linguistic hedges that one can meaningfully define is limited (Lakoff, 2014). How many qualitatively different states can cases have and still be members of the same overarching set? Second, the limitations of the (English) language. There are only so many hedges available to form an ordinal scale where people still intuitively understand the correct order of the hedges. This puts the maximum number of linguistic hedges at eight, which is unproblematic given that many manually calibrated fuzzy sets use fewer than eight degrees of membership (Ragin, 2008, p. 31) and that continuous fuzzy sets are, in fact, a form of spurious specificity (see above).

The eight linguistic hedges describe how many of the characteristics of the concept (the set) a case has. A case's degree of membership in a set – identified by a linguistic hedge – captures the degree to which a case has the relevant characteristics. As argued, the crossover point sits between diluting and intensifying hedges ($\leq$ moderately = 0; $\geq$ considerably = 1). However, researchers can raise or lower the crossover point to define the crisp set of at least mostly-X ($\leq$ considerably = 0; $\geq$ mostly = 1) or the crisp set of at least moderately-X ($\leq$ somewhat = 0; $\geq$

moderately = 1). When using fewer than eight hedges, it is important to preserve the meaning of the hedges across different sets. Using three hedges, researchers calibrate, e.g. fully a case of X, at least considerably a case of X and at most moderately a case of X. The prefixes 'at least' and 'at most' enable comparing sets with fewer hedges to those with more hedges. See Table 1 for an example. Note that empirical values in the example are not distributed proportionally to emphasize that not all empirical variation between fully a case of X and not-a case of X is equally relevant. Depending on the nature of the concept, researchers can calibrate fewer (at least two) or more (at most eight) linguistic hedges. (Table 1 shows just a few possible examples of calibrating linguistic hedges, and they do not have to be symmetrical around the crossover point.)

All thresholds, of course, have a degree of ambiguity, and linguistic hedges are no exception. Carefully defining each hedge solves ontological ambiguity, but it may still leave multiple plausible empirical values to connect to the definition (epistemological ambiguity). Therefore, researchers may want to perform robustness and sensitivity tests for their hedges. At least for those hedges that are used as crossover points in the analysis to crispify sets.

In sum, calibrating linguistic hedges follows the same logic as calibrating fuzzy (and crisp) sets (Ragin, 2008, p. 80, see; Grillitsch et al., 2023, pp. 1458–1459 for an example). Except that linguistic hedges are not assigned set-membership values. On the one hand, linguistic hedges capture the degreeness of social reality, on the other hand, they conform to QCA's threshold logic by crispifying the hedges at a crossover point that the researcher chooses. This, in fact, opens new research avenues in QCA.

Investigating degreeness with linguistic hedges

Fuzzy set values do not play a role in minimizing the truth table because QCA relies on propositional (crisp) logic. However, linguistic hedges offer a way to make the degreeness of social reality count – and calculate crisp set-relationships to avoid conflating crisp and fuzzy logic. As argued, the crossover point 'sits' between intensifying and diluting linguistic hedges. All intensifying hedges become 1s and all diluting hedges become 0s (see above). We now have a crisp set but can still use the linguistic hedges when going back to the cases. We have all that, analytically, the truth table minimization requires (i.e. crisp sets) without assuming that social reality is dichotomous. However, as argued above, we can also raise or lower the crossover point. Lowering or raising this threshold offers three possibilities to make the degreeness of social reality count in QCA (Figure 1).

First, because of QCA's threshold logic, conditions are either necessary or not-necessary for the outcome. However, often times it may be necessary to just have the condition in degree for cases to achieve the outcome. For example, being a conservative is not necessary for voting for a far-right party. However, people without any conservative views will not vote far-right so perhaps a degree of conservatism is necessary (but not sufficient) for people to vote far-right. A QCA researcher may lower the bar on 'being conservative' and set the crossover point at, e.g. moderately conservative to inspect whether having this condition is necessary for voting far-right. Of course, lowering the threshold for conservatism makes the set of conservative people larger and, therefore, by definition a more consistent superset of the set of far-right voters. Researchers must have

Table 1. Linguistic hedges, meaning and measurements.

				examples of other possible calibrations		
linguistic hedges		definition	empirical values	two hedges	four hedges	five hedges
intensifying hedges	fully poor household	must choose between food and clothing	≤ £900	at least considerably	at least very	fully
	very poor household	cannot afford rent after food and clothing	£1,000–£1,100			at least mostly
	mostly poor household	can only afford food, clothing and rent	£1,100–£1,400		at least considerably	
	considerably poor household	can only afford basic cost of living in most of the months	£1,400–£1,500			considerably
diluting hedges	moderately poor household	can afford cost of living every month with no savings	£1,500–£1,600	at most moderately	at most moderately	at most moderately
	somewhat poor household	can afford cost of living with little savings	£1,600–£1,700			
	marginally poor household	can afford occasional leisure and luxury	£1,700–£2,000		at most marginally	
	not poor household	can afford regular leisure and luxury	≥ £ 2,000			not

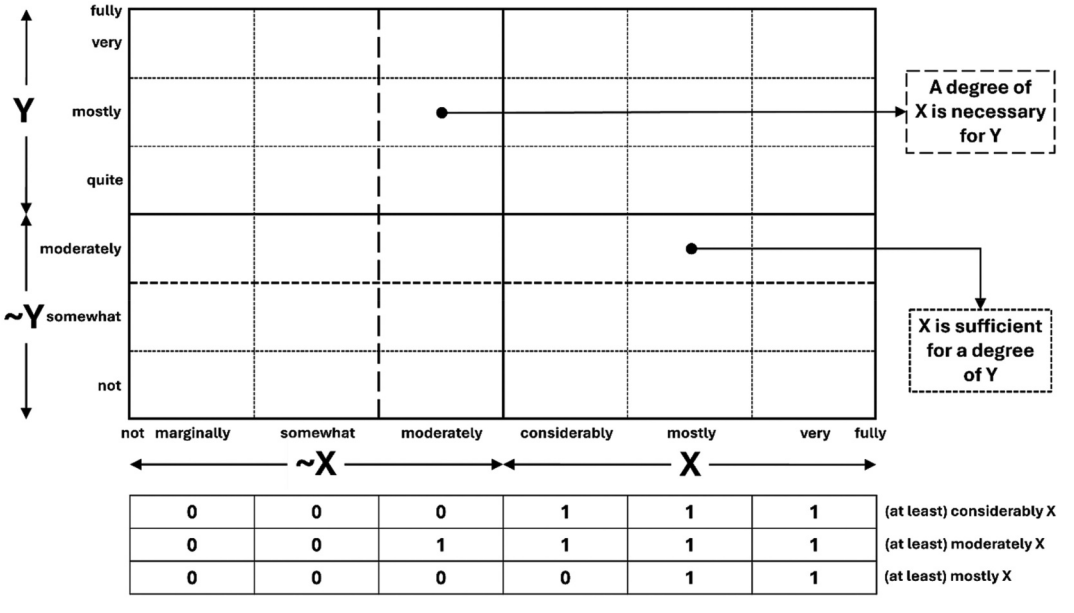

Figure 1. Making degreeness count.

Table 2. Solutions for degrees of the outcome.

study	moderately-Y	considerably-Y	mostly-Y	interpretation
(i)	A + B + C	A•B + A•C	A•B•C	upgradeable path
(ii)	A•B + A•C	A•B•C	—	dead-end path
(iii)	A	B•C	D•E	unique paths

good substantive arguments for lowering the threshold. At the same time, it would be very meaningful to find that, e.g. being (at least) moderately conservative is necessary for voting (at least) considerably far right. Lowering the threshold to investigate necessity in degree mimics what NCA (Necessary Condition Analysis) does (Vis & Dul, 2018) but, unlike NCA, stays within QCA's set-analytical logic.

Second, single conditions are almost never sufficient for the outcome. However, by lowering the bar on Y, researchers may investigate the explanatory power of each condition. If (at least) considerably-X is sufficient for (at least) moderately-Y (lowering the threshold for Y a little), X may be argued to have strong explanatory power. Just having X is almost sufficient for having Y. However, if X is only sufficient for marginally-Y, its explanatory power is weak. Again, thus enlarging the set of Y makes X a more consistent subset by default, so researchers must have good substantive reasons for doing so. At the same time, it is meaningful to know how close to the target any one condition gets by itself.

Third, researchers can do separate truth table analyses for two or three degrees of the outcome: e.g. moderately-Y, considerably-Y and mostly-Y. These are three qualitatively different states of the outcome that may have different explanations. Suppose three QCA studies would find the following solutions for three degrees of Y (see, Table 2).

In this example, the paths in Study (i) can be upgraded to achieve higher degrees of the outcome. Instead, the path in Study (ii) is a dead-end path that does not allow cases to achieve the highest degree of the outcome. Cases on Path (ii) will have to do something different (path breaking) to progress from considerably-Y to mostly-Y, whereas cases on Path (i) can focus on doing the same thing better (augmenting 'their' path). This is relevant for, for example, policy decisions. Perhaps, a human-capital development path may support regional development all the way to regions becoming mostly developed, whereas a path that focuses on building infrastructure may only get regions as far as becoming considerably developed. Example (iii) shows unique paths for each degree of Y, which suggests that achieving a higher degree of the outcome requires specific characteristics (e.g. skills) that do not 'serve' lower degrees of the same outcome. Again, lowering and raising the bar for Y will (by default) produce solutions of the kind suggested in Table 2. Researchers need good reasons to raise or lower the bar. However, findings like those in Table 2 (and if they are interpretable against theoretical and substantive knowledge) acknowledge the degreeness of social reality more so than does current practice in fsQCA. Of course, one can lower and raise the crossover point in fsQCA without first defining linguistic hedges. But one would have to justify the alternate settings of the crossover point and so effectively define thresholds (linguistic hedges).

One may object that changing the crossover point is often suggested as a robustness test (Rutten, 2022). However, calibrating thresholds is something different. Linguistic hedges calibrate qualitatively different kinds of cases. The empirical values connected to them are far apart, and QCA is not expected to be

robust for major changes to the crossover point (Skaaning, 2011). Calibrating different cases one expects to find different solutions. Instead, robustness tests assess the solution's sensitivity to small changes to the crossover point (Greckhamer et al., 2018, p. 490). Small changes reflect plausible alternate empirical values for the same semantical meaning. Small changes to the crossover point still calibrate qualitatively the same kinds of cases and solutions should be robust to them (Rutten, 2022). Instead, investigating linguistic hedges as suggested above is part of QCA's dialogue between ideas and evidence. It helps interpret what configurations mean and develops richer and more nuanced causal explanations (Cooper & Glaeser, 2016, p. 456).

Finally, notice that considerably-Y refers to cases that are at least considerably-Y and moderately-Y refers to cases that are at least moderately-Y. The above does not suggest performing different truth table analyses on cases that are only considerably-Y versus cases that are only moderately-Y. That is possible, of course, but may dramatically reduce the case population. Crisp sets do not deny that there is truth in degree behind the crisp statements 'true' (1) and 'false' (0), only that this variation is analytically irrelevant for the research question at hand. The benefit of linguistic hedges is that they capture this degreeness and that researchers can easily make it count in their investigations. Just as it may be analytically relevant to investigate whether X makes a difference for Y, so it may be relevant to investigate, e.g. whether moderately-X makes a difference for considerably-Y. Moreover, linguistic hedges work (in this sense) no different than fuzzy sets. When in fsQCA, one finds that X is sufficient for Y, researchers interpret this (in terms of case-level causal claims) as being at least 0.5-X is sufficient for being at least 0.5-Y (Ragin, 2008, pp. 131–133).

Conclusion

This paper suggested linguistic hedges as a superior alternative to fuzzy sets in QCA. Conceptualizing the degreeness of social reality not as fsQCA's calibrated variables but as linguistic hedges opens up new avenues for set-analytical investigation. Fuzzy sets in QCA are, in fact, not fuzzy (in the Zadeh-sense) because they are calibrated. Nor do fuzzy set-membership values play any role in QCA's principal analytical tool, i.e. the truth table minimization. Nothing of analytical value is lost by abandoning fuzzy sets in favour of linguistic hedges (Lakoff, 2014). In fact, there may well be more semantically relevant thresholds than just the three defined in QCA's fuzzy sets (floor, crossover point, ceiling). What is gained with linguistic hedges is congruence between singular and general causal evidence and a consistent adherence to QCA's threshold logic throughout a study. Moreover, per the above examples, linguistic hedges allow investigating the degreeness of social reality in a way that is beyond current QCA practice, crisp or fuzzy. Whether one accepts this paper's criticism of fuzzy sets and acknowledges linguistic hedges as an alternative may largely depend on whether one prioritizes ontology or epistemology (Porpora, 2018). Researchers who primarily think epistemology will value how QCA's fuzzy sets capture all empirical variation (e.g. Schneider and Wageman 2012: 193). Those who primarily think ontology will appreciate how linguistic hedges are semantically meaningful thresholds that distinguish between relevant and irrelevant variation also between fsQCA's floor,

crossover point and ceiling. How linguistic hedges capture the degreeness of social reality by defining multiple kinds of qualitatively different cases rather than assuming that, between 0 and 1, differences can be meaningfully expressed on a ratio scale. But however one thinks about ontology and epistemology, QCA is a threshold method, fuzzy sets are calibrated (i.e. not fuzzy in the Zadeh-sense) and the degreeness they capture is not put to use in the truth table minimization. Lowering and raising the crossover point with linguistic hedges acknowledges social reality's degreeness more profoundly than does fsQCA practice and suggests a way forward for all.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Roel Rutten is an associate professor in the department of Organization Studies of Tilburg University and a senior research consultant at European Regional Affairs Consultants. He is a leading QCA expert and has published a handbook and several papers about this method. He has applied QCA to study the geography of innovation and uses the method for policy evaluations.

ORCID

Roel Rutten http://orcid.org/0000-0002-5933-2996

References

Beach, D., & Pedersen, R. (2018). Selecting appropriate cases when tracing causal mechanisms. *Sociological Methods & Research*, *47*(4), 837–871. https://doi.org/10.1177/0049124115622510

Beach, D., & Rohlfing, I. (2018). Integrating cross-case analysis and process tracing in set-theoretic research: Strategies and parameters of debate. *Sociological Methods & Research*, *47*(1), 3–36. https://doi.org/10.1177/0049124115613780

Schneider, C., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge University Press.

Cooper, B., & Glaeser, J. (2016). Exploring robustness of set theoretical findings: An illustration from the sociology of education. *International Journal of Social Research Methodology*, *19*(4), 445–459. https://doi.org/10.1080/13645579.2015.1033799

Finn, V. (2022). A qualitative assessment of QCA: Method stretching in large-N studies and temporality. *Quality and Quantity*, *56*(5), 3815–3830. https://doi.org/10.1007/s11135-021-01278-5

Goertz, G. (2017). *Multimethod research, causal mechanisms, and case studies: An integrated approach*. Princeton University Press.

Goertz, G. (2020). *Social science concepts and measurements*. Princeton University Press.

Goertz, G., & Mahoney, J. (2012). Concepts and measurement: Ontology and epistemology. *Social Science Information*, *51*(2), 205–216. https://doi.org/10.1177/0539018412437108

Greckhamer, T., Furnari, S., Fiss, P., & Aguilera, R. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization*, *16)4*(4), 482–495. https://doi.org/10.1177/1476127018786487

Grillitsch, M., Sotarauta, M., Fitjar, B., Haus-Reve, R., Kolehmainen, S., Kurikka, J., Lundquist, H., Martynovich, K., Monteilhet, M., Nielsen, S., Nilsson, H., Rekers, M., Sponanen, J. Stihl, L., & Sopanen, S. (2023). Agency and economic change in regions: Identifying routes to new path

development using qualitative comparative analysis. *Regional Studies*, *57*(8), 1453–1468. https://doi.org/10.1080/00343404.2022.2053095
Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, *2*(4), 458–508. https://doi.org/10.1007/BF00262952
Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. University of Chicago Press.
Lakoff, G. (2014). Set-theory and fuzzy sets: Their relationship to natural language - An interview with George Lakoff. *Qualitative and Multi-Method Research*, *12*(1), 9–14.
Mohan, C. (2019). *An introduction to fuzzy set theory and fuzzy logic*. MV Learning.
Pappas, I., & Woodside, A. (2021). Fuzzy-set Qualitative Comparative Analysis (fsQCA): Guidelines for research practice in information systems. *International Journal of Information Management*, *58*(102310), 1–23. https://doi.org/10.1016/j.ijinfomgt.2021.102310
Porpora, D. (2018). Critical realism as a relational sociology. In F. Dépelteau (Ed.), *The palgrave handbook of relational sociology* (pp. 413–429). Palgrave Macmillan.
Pula, B. (2021). The logico-formalist turn in comparative and case study methods: A critical realist critique. *International Journal of Social Research Methodology*, *24*(6), 739–751. https://doi.org/10.1080/13645579.2020.1823063
Ragin, C. (2000). *Fuzzy-set social science*. The University of Chicago Press.
Ragin, C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago press.
Ragin, C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*. University of California Press. (Original work published 1987)
Rohlfing, I. (2020). The choice between crips 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
Rubinson, C., Gerrits, L., Rutten, R., & Greckhamer, T. (2024). *Avoiding common errors in QCA: A short guide for new practitioners* (pp. 1–6). http://Compasss.org/wp-content/uploads/2019/07/Common_Errors_in_QCA.pdf
Rubinson, C., & Rutten, R. (2023). Concepts before numbers: Teaching effective calibration practices for qualitative comparative analysis. In A. Ruth, A. Wutich, & R. Bernard (Eds.), *The handbook of teaching qualitative and mixed research methods: A step-by-step guide for instructors* (pp. 340–344). Routledge.
Rutten, R. (2022). Applying and assessing large-N QCA: Causality and robustness from a critical realist perspective. *Sociological Methods & Research*, *51*(3), 1211–1243. https://doi.org/10.1177/0049124120914955
Rutten, R. (2023). Uncertainty, possibility and causal power in QCA. *Sociological Methods & Research*, *52*(4), 1707–1736. https://doi.org/10.1177/00491241211031268
Rutten, R. (2024). *Qualitative comparative analysis: Learning from cases*. Edward Elgar.
Skaaning, S. (2011). Assessing the robustness of crisp and fuzzy-set QCA results. *Sociological Methods & Research*, *40*(2), 391–408. https://doi.org/10.1177/0049124111404818
Smithson, M. (1987). *Fuzzy set analysis for behavioral and social sciences*. Springer.
Smithson, M. (2016). Fuzzy sets and fuzzy logic in the human sciences. In C. Kahraman, U. Kaymak, & A. Yazici (Eds.), *Fuzzy logic in its 50th* (pp. 175–186). Springer.
Smithson, M., & Verkuilen, J. (2006). *Fuzzy set theory applications in the social sciences*. Sage.
Vis, B. (2012). The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociological Methods & Research*, *41*(1), 168–198. https://doi.org/10.1177/0049124112442142
Vis, B., & Dul, J. (2018). Analyzing relationships of necessity not just in kind but also in degree: Complementing fsQCA with NCA. *Sociological Methods & Research*, *47*(4), 872–899. https://doi.org/10.1177/0049124115626179
Zadeh, L. (1965). Fuzzy sets. *Information & Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X
Zadeh, L. (1972). A fuzzy-set theoretic interpretation of linguistic hedges. *Journal of Cybernetics*, *2*(3), 4–34. https://doi.org/10.1080/01969727208542910