Exploring associations with mathematics dispositions: a novel perspective combining measurement and set-theoretic analytical approaches

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

In this paper we approach modelling mathematics dispositions from a different methodological perspective in order to shed more light into the complex interplay between teaching practices and students' learning outcomes. We draw on survey data from around 5000 students from Year 7 to 11 (age 11 to 16) from 40 Secondary schools in England. Our methodological approach includes Rasch modelling to produce measures of attitudinal outcomes as well as students' perceptions of pedagogic practices. We then employ fuzzy-set Qualitative Comparative Analysis (fs/QCA) to explore the relationships between students' characteristics and the perceived type of teaching they receive in mathematics. We use two measures of 'transmissionist teaching' which aim to quantify the degree to which teaching practices are perceived as 'teacher-centred'. One measure gives the students' perceptions and the other gives the teacher's perspective. We find that different configurations of student and teacher perceptions of transmissionist teaching are associated with high and low mathematical dispositions for different year groups and for boys and girls. We discuss the methodological merits of this approach along with the substantive educational implications of these findings.

1

1. Introduction: Overview of Literature and Context of this study

Mathematics as a topic in the school curriculum is considered important, if not vital, for students' potential access to further study in Science, Technology, Engineering and Mathematics (STEM) subjects, and also many fields in the Social Sciences (e.g. Business, Economics, etc) in Higher education, and consequently to their educational and socioeconomic life opportunities. The need to promote a mathematically engaged society is also well documented in literature and recent policy documents in the UK (OFSTED, 2011; Smith, 2017) and globally (see Mevarech and Kramarski, 2014). The Advisory Committee on Mathematics Education (ACME) (ACME, 2012, 2009) recognised this important issue and advocated "tackling the perceptions of mathematics" as a particularly important issue placing emphasis on the importance of mathematics as a "powerful analytical tool", with inherent "pervasiveness" (2009, p.5) and a "key workforce skill" (ibid, p. 6). Looking beyond attainment, research on attitudes towards mathematics (and dispositions) is very important and has gained considerable interest over the past 40 years because it may reveal key influences on students' choices and decision-making and hence their future engagement with STEM[1]; this is our emphasis here.

The complex field of attitudes and dispositions towards mathematics

The field of 'attitudes towards mathematics' has seen numerous adaptations of instruments (e.g Hemmings and Kay, 2010; Kleanthous and Williams, 2013; Lim and Chapman, 2013) mainly influenced by the well-known Fennema-Sherman scales (Fennema and Sherman, 1977). Relevant studies and resulting research findings have focused on one or more from many 'dimensions', including beliefs, values, identities, engagement, affect, emotions, motivation, confidence, self-efficacy, and dispositions (Di Martino and Zan, 2011; Hannula, 2002; Ruffell et al., 1998). This complexity, as well as the lack of agreement on the definition of the construct has led researchers

---

[1] There is not space here to detail the literature on students' dispositions and why the study of them is important but see Allexsaht-Snider & Hart (2001), Archer, Halsall, Hollingworth, & Mendick (2005), Brown, Brown, & Bibby (2008) and Pustjens, Van de gaer, Van Damme, & Onghena (2004) for more on this.

2

(e.g. Lim and Chapman, 2013; Watson, 2011) to recently revisit the established instruments of the 1970s and 1980s looking for alternative universal definitions or more parsimonious instruments.

Previous studies had also identified a plethora of socio-cultural factors which are significant in shaping students' dispositions and choice-making in education in general, and in STEM subjects and mathematics in particular: for example class, gender, nationality, ethnicity, parental and peer cultures as well as the intersectionality between these variables (e.g. Cao et al., 2007; Lei and Pampaka, 2018; Mendick, 2005; O'Brien et al., 1999). It is not our intention here to rehearse these discussions, nor to propose an alternative definition of 'attitudes' towards mathematics. Rather, we follow the paradigm introduced and validated in our earlier work with post-secondary students, where instruments were developed for measuring dispositions towards further study [Author citation]; adapting and revalidating these for secondary-school students [Author citation].

The Relationship between Teaching Practices and Dispositions

The quality of mathematics teaching is an important factor in understanding students' dispositions, Within this literature (Askew et al., 1997; Driessen and Sleegers, 2000; McCaffrey et al., 2001; Schuh, 2004; Swan, 2006; Swanson and Stevenson, 2002; Vermunt and Verloop, 1999) it has been widely accepted that effective mathematics teaching should be 'connectionist' in two ways: (a) connecting teaching to students' mathematical understanding, and (b) connecting teaching and learning across mathematics' topics, and between mathematics and other knowledge. Many others have also argued that formative assessment and more dialogical pedagogies are required for the development of deeper mathematics understanding (e.g. Black and Wiliam, 1998; Boaler, 2002; Lapointe et al., 2005; Meece et al., 2003; Ryan and Williams, 2007; Wilkins and Ma, 2003). Such approaches are thought to be the opposite to more 'transmissionist' practices where mathematical knowledge is compartmentalised and treated as separate from other subjects and from students' own context. In previous work we have shown how some institutional and pedagogic practices can encourage the reduction of the learning of mathematics to the procedural level only, thus limiting

3

future educational opportunities). In this work, we have used and 'quantified' the terms

'connectionist' and 'transmissionist' teaching, employing Swan's previously developed instrument

(Swan, 2006), which in turn built on the research findings of Askew et al. (1997) and Ernest (1991).

Pampaka et al. (2012) described in detail how the items of Swan's (2006) 'practice scale' could be

amended to create a unidimensional measure of 'teacher centricism' (or transmissionism) in

teaching mathematics which was successful in associating transmissionist practices with the decline

of students' dispositions over time.

But there is still an unresolved gap in knowledge in regards to the developing students' mathematics

dispositions and how these are shaped in relation to teaching at earlier educational stages. This was

the ultimate aim of the project we draw from here.

Studying and Modelling Dispositions

Quantitative studies of attitudes and dispositions to mathematics tend to employ analytical

techniques such as latent class analysis (Lazarides and Ittel, 2012), and regression modelling (Mata

et al., 2012) as in [authors] previous work. These approaches provide good summaries of whether a

particular attitude is strongly associated with an outcome 'in general' by testing the strength of the

association for all the respondents. This assumes that some general relationship may exist between

attitudes, perceptions of processes/practices and outcomes and also that these variables used as

explanatory variables or predictors in such models are independent from each other, an assumption

hard to meet with social constructs and for social phenomena. In previous work, teaching was

perceived as more transmissionist by female students and by older students but teacher perceived

transmissionism was not always significantly associated with dispositions [Author citation].

Despite finding some statistically significant associations between these variables, the predictive

power of the models overall was not particularly high. One potential explanation of this could be

that subgroups in the model (such as students in different years) had different associative

relationships with the outcome and other variables in the model, which a model for the whole

4

sample failed to capture well. Even though sub-group analysis could probably highlight some of these differing relationships, such partitioning of the whole sample is not always desirable. The same complexity has also been observed with other models and with other similar datasets, especially when it comes to identifying and interpreting interaction effects (Gabriel et al., 2018; Lei and Pampaka, 2019, 2018; Pampaka et al., 2018).

We want thus to investigate further whether different attitudes and practices combine in particular ways to influence a student's dispositions towards mathematics and whether the ways in which they combine may be different for different subgroups of the dataset. An approach that can help us to understand whether these different patterns exist for different subgroups is Qualitative Comparative Analysis (QCA) – a set-theoretic method of data analysis that does not assume linear relationships between explanatory and outcome variables. To our knowledge, there is no previous work reported in modelling (mathematics) dispositions using a QCA approach. In fact, there are only a limited number of papers in education in general which use QCA; see for example the work of Cooper and Glaesser (Cooper and Glaesser, 2010; Glaesser, 2008; Glaesser and Cooper, 2012a, 2012b), Thomson (2011a, 2011b), Lee (2012) and the very recent and timely paper of Bingham et al. (2019). We aim to add to this literature base by demonstrating how QCA combined with interval measures can be used to explore complex relationships between attitudes and other outcomes and what this approach might reveal in addition to findings from other types of analyses. In this paper, we are mainly interested in modelling dispositions as the outcome whereas the majority of studies dealing with maths dispositions or other affective factors usually focus on modelling their effect on mathematics performance and attainment (e.g. Grootenboer and Hemmings, 2007; Hemmings and Kay, 2010).

The paper draws on data from a longitudinal (survey) study of teaching and learning secondary mathematics focusing on the impact of pedagogical practices on important learning outcomes in UK. The particular research questions that guide this paper are as follows:

RQ1: How can QCA be combined with interval scores from Rasch modelling to inform our understanding of complex educational relationships? (methodological)

RQ2: What conditions/combinations of students' characteristics, attitudinal learning outcomes and perceptions of teaching practice can explain whether students will have high or low dispositions towards mathematics during their secondary education? (substantive)

2. Methodology

Instrumentation and Sample

This longitudinal, school-level (for selection purposes) study employed a varied sampling frame to ensure maximum coverage of the schools of England. In total, we approached over 2200 schools and we were able to establish collaboration with 40 of them to participate at least on the first data point[2]. We note here issues around self-selection bias in this type of studies, which could limit the representativeness of the achieved sample[3].

Data collection in these schools involved a student questionnaire (at three data points) with questions about students' attitudes to mathematics, future aspirations, and their perceptions of the teaching they encounter. Teacher perceptions of teaching were also captured through a separate survey administered to respondents' mathematics teachers. Student questionnaires were based on different versions of the same instrument to reflect the age and level of students (i.e. 5 different Year Groups, years 7 to 11). Background variables and information on students' attainment were also collected. The various sections of the questionnaire capturing teaching and learning perceptions were constructed based on a previous project [Project link] framework where instruments for students aged 16 and older were validated and used [Author citations].

---

[2] This was complemented by interview data collected from students and teachers in two case study schools. We focus here on the analysis of the survey data from the 40 schools that participated in the study from the start, and not this interview data.

[3] A recent MSc dissertation had focused on the implications of missing data on national comparability and did not find evidence of bias (Brown, 2017).

6

This analysis draws on the first data point (DP1, out of three) of our longitudinal survey of students and their mathematics teachers. This data collection took place from October to December 2011. During this period we collected data from a sample of 13,643 students in 40 schools. As per our usual measurement practice the construction of the measures, and their consequent calibration for the purposes of this analysis, has been conducted with all available data. From this DP1 we then analyse the responses on a complete case analysis (i.e. those with responses in all relevant variables). The resulting sample consists of 5022 students: 2988 girls (59%) and 2034 boys (41%). This was a pragmatic decision to facilitate the demonstration of this novel methodological approach and also comparisons between previous work which was based on the same sub-sample. We acknowledge the potential limitations of this decision, but dealing with missing data as we would have done with traditional modelling approaches [Author citation] goes beyond the scope of this paper.

Measurement Methodology

Before any substantial analysis takes place, the first step we follow in our methodological approach is the validation of our constructed measures. This measurement step is based on the assumption that there are various underlying constructs behind the groups of items in the questionnaires, which were brought together after studying previous research literature and looking at other researchers' instruments.

Given the students' responses to the relevant questions we then attempt to validate these aforementioned constructs: in other words to check whether they exist as "measures" (or scales), and if not if there are other dimensions relevant and useful. Our validation process, thus, refers to the accumulation of evidence to support validity arguments regarding the students' reported measures (Messick, 1988). We employed a psychometric analysis for this purpose, conducted within the Rasch measurement framework, following relevant guidelines (Wolfe and Smith Jr., 2007a, 2007b) and with the Rasch rating scale model, in particular, which is the most appropriate

7

model for Likert type items, as in this project. Our decisions about the validity of the measures are based on different statistical indices, such as item fit statistics, category statistics, differential item functioning and person-item maps (Bond and Fox, 2015). More information about the validation of the "mathematics disposition" construct is presented elsewhere [Author citation]. As also indicated earlier we have validated measures for both teachers' and students' perceptions of (the same) teaching practice. In particular we use here a measure of students' perception of how transmissionist teaching is (which we call "student perceived transmissionism" throughout) and of teacher's perception of how transmissionist teaching is (which we call "teacher perceived transmissionism" throughout) [Author citation].

In the next section we introduce the QCA approach using generic/artificial examples before we move on to discuss how we used it within the context of this study.

Analytical Approach for this paper - QCA

Rationale
Fuzzy-set qualitative comparative analysis (fs/QCA) is a case-based (rather than variable-oriented), set-theoretic method of data analysis that can be used with any size of dataset but is less commonly used with large-size datasets (Rihoux and Ragin, 2009). It is used here to sort student data into types, based on combinations (called configurations) of student characteristics and then to assess which of these configurations are typically associated with high/low dispositions and the strength of the relationship. In our fs/QCA, the student characteristics, constructed from variables/measures in the original dataset, are conceived of as sets. Each case (i.e. student) has a membership score for every set, to indicate the degree to which that case 'belongs' to that set.

Using fs/QCA allows us to assume a situation where the relationships between student characteristics and dispositions may be complex and asymmetric and where there may be more than one configuration of characteristics showing a strong relationship with high (or low) maths dispositions (Ragin, 2008a, 2008b). As we discussed earlier, there is every reason to expect our

8

data to display such asymmetry and complexity as we are considering the conjunction of several

related, but not identical, aspects of mathematical attitudes [Author citations]. We predict that there

may only be a small number of students for whom the various aspects of mathematical attitudes do

not line up but we are just as interested in the relationship between the particular configuration of

these measures of attitudes and practices and its impact on their dispositions. Such detail can be

lost when employing a linear modelling approach as the fitted model is trying to show the best

estimate for the effect of any variable (or interaction of variables) for all cases – i.e. assuming that it

is sensible to calculate such a thing. Where we expect, as we do here, that attitudes and practices

may have a different degree (or direction) of influence depending on what they are in conjunction

with, then fs/QCA is an appropriate choice of analytic method. We assume that the interplay of

characteristics may be crucial for the formation of mathematical dispositions and not just the levels

or presence of the individual characteristics themselves. Regression-based approaches to analysis

of this kind, however sophisticated, assume each factor/variable can have a unique effect on an

outcome but, actually, what is established by models of this kind is an average, or 'net-effect', of a

variable for all cases or a group of cases with random effects multi-level models (Ragin, 2006a, p.

13).

QCA, according to Byrne (2013) allows us to "get a handle on what works where (in what context),

when (in what temporal context), and in what order" (p. 217). Regression-based approaches, even

those which include interaction terms in an attempt to account for complexity, assume that variables

in a model are capable of 'influencing the level or probability of an outcome' on their own (Ragin,

2006a p 14). Standard multiplicative interaction terms, e.g. between variables A and B, can hide

complexity behind a single value. For example, if variable A has value 1 and variable B has value

0.5, A*B = 0.5; but A*B would also equal 0.5 if A had value 0.5 and B had value 1. In fs/QCA,

these two situations are treated as qualitatively different (see Vaisey, 2007). In addition, we want to

consider 3- and 4-way interactions between characteristics and attempting this within a linear-

regression framework leads to a great deal of interpretative difficultly. Of course, it is possible that

9

some characteristics may dominate to the extent that other characteristics matter very little and a fs/QCA approach allows us to see when this is the case.

Terminology
We will talk, throughout, about our measures of attitudes and practices as 'characteristics' of the cases (students) rather than 'variables' or 'factors' (partially, to avoid confusion with other analytic methods). We avoid the use of 'variables' to emphasise that we are, as Ragin (2000 p65) describes, 'assessing scores in context' rather than automatically treating cases with similar scores on one characteristic as similar. We avoid using 'factors' so as to make clearer that our approach is not a form of factor analysis. In explanatory factor analysis, researchers are concerned with reducing the number of explanatory factors believing that the values in some small set of factors may be reflective of those in a wider set and thus that those values in the extra factors are redundant. In fs/QCA, although a process of simplification does occur, if a characteristic is minimised out of a solution term, this is because it, empirically, makes no difference to the outcome, in the presence of the other model characteristics. We can use this approach to see which particular combinations of students characteristics are necessary and/or sufficient for high dispositions in mathematics.

Constructing sets and assigning set membership scores for cases
As mentioned, QCA relies on conceiving our characteristics as sets and assigning a set-membership score, to each case, for that set. Sets can be crisp to indicate the presence (score membership = 1) or absence (score membership = 0) of a characteristic. This is more appropriate for binary categorical variables. Most usefully for the study of dispositions (and other similar continuous measures) sets can also be fuzzy and then cases can take any score between 0 and 1. For fuzzy sets, the score can be anything between 0 and 1 where 0 indicates complete non-membership in the set, 1 indicates complete membership and a value of 0.5 indicates that the case is neither in nor out of the set in question. In other words, the degree of set-membership gets larger as scores approach 1 and smaller as scores approach 0. The 0.5 value marks a qualitative threshold between cases that are definitely in the set (those with scores above 0.5) and those that are definitely out of it (those with

10

scores below 0.5).   If a case has a 0.5 value, it has not been possible to determine whether it definitely sits within the set or outside of it.

It is important to note here that calibration is not a simple re-coding.   The researcher must develop a method and rationale for assigning ~~these~~ set-membership scores to cases and this should be driven by information about the cases themselves. For this, researchers must decide which values in the original dataset correspond to the boundaries of the set (the 0, 0.5 and 1 values). They may choose to use meaningful values or summaries of the original variables to do this, depending on what makes sense in the context of the data and research aims. For example, if we were creating a calibrated set from an original set of test scores, we may decide that the 'pass mark' constitutes a reasonable value to equate to a set-membership score of 0.5. In this case, the created set would constitute the set of passes and set-membership scores approaching 1 would be those where the original scores were almost full marks. Another approach would be to equate the 0.5 set-membership score to a particular percentile in the original distribution of test scores. Equating, for example, the 80th percentile score to the 0.5 value would create a slightly different set – the set of high scores. The purpose of the analysis and properties of the data should drive the decision about calibration.

In our particular example, we can take advantage of the properties of the Rasch model to help us make decisions about how to calibrate because we have measures of mathematics dispositions created by the Rasch model. The examples above show, however, that other calibration methods are possible and sensible, even for large-n datasets, depending on the particular context. We explain more about our calibration procedures in section 3 but first explain some other important concepts in QCA.

Necessary and sufficient conditions
The concepts of sufficiency and necessity are central in the QCA approach: the membership scores, described earlier, can be used to examine whether certain configurations of characteristics can be

11

considered necessary and/or sufficient for a given outcome. A necessary condition[4] for an outcome

must be present for the outcome to be achieved/observed but may not be enough on its own to

produce the outcome (Ragin, 2008b, 2000). Similarly, if every instance of a characteristic, or

configuration of characteristics, is followed by the outcome, we can say that that

characteristic/configuration of characteristics is a sufficient condition for the outcome. In some

cases, a condition may be necessary and sufficient and, in others, it may be neither necessary nor

sufficient. A configuration is perfectly sufficient for an outcome if the set membership scores in it

are less than the scores of the outcome set, for all cases – or, in other words, when the condition is a

subset of the outcome set. When this relationship holds, every case that is in the configuration set

is, by default, also in the outcome set. Similarly, a condition is perfectly necessary for an outcome

if the set membership scores in the condition are greater than the scores of the outcome set, for all

cases – or when the outcome is a subset of the condition set. When examining the social world,

particularly, though not exclusively, when using large datasets, it is unusual to encounter such

examples of perfect necessity or sufficiency. Instead, we allow the relationships between

configuration and outcome sets to be less than perfectly necessary/sufficient because they allow us

to determine whether a characteristic is, for example, 'almost always' sufficient, or quasi-sufficient

(Ragin, 2008b, 2000). In this way, we are looking for empirical evidence of relationships that *may*

be causal in reality. Evidence of strong necessary and/or sufficient relationships (i.e. those

approaching perfect necessity or sufficiency) between outcome and configurations *could* be

evidence of a wider causal pattern though this analysis alone cannot provide definitive evidence of

causality and any claims of causality should also make appropriate appeals to theory, research

design, and other sources of evidence, as appropriate. In this paper, we will focus on relationships

of sufficiency as these fit well with our substantive questions about which combinations of

characteristics may be leading to high or low dispositions.

---

[4] A condition may be one characteristic or a configuration of characteristics.

Using truth tables to understand the data

The dataset, then, must be calibrated to translate original values into crisp or fuzzy-set scores before

QCA analysis can begin[5]. Once the data has been calibrated, it can be analysed to determine

whether there are any patterns indicating a relationship of quasi-necessity or quasi-sufficiency

between characteristics.

Table 1 is an example of a truth table, with made-up data, where "Male", "Maths ability", and

"Maths confidence" are characteristics being drawn upon to explain an outcome, "high Maths

disposition"[6]. Table 1 has 8 (or 2^3) rows and shows all the possible combinations of "Male",

"Maths ability", and "Maths confidence" and each row represents a configuration. In the case of

the set 'Male', where the characteristic is a crisp set, the 1's and 0's in the table simply indicate the

presence or absence of this characteristic. So, for example, rows 1, 3, 4, and 6 represent boys and

rows 2, 5, 7, and 8 represent girls. In Table 1, 'Maths ability' and 'Maths confidence' are fuzzy –

so each row includes cases with a range of membership scores in these sets[7]. The 1's in Table 1 in

the 'Maths ability' and 'Maths confidence' columns indicate cases where the fuzzy membership

score is greater than 0.5 and the 0's are cases with a score of less than 0.5. So, row 1 represents

boys with scores of greater than 0.5 in 'Maths ability' and 'Maths confidence'. The rows of Table 1

then, are intersections of sets or configurations and we write row 1 as MALE*MATHS

ABILITY*MATHS CONFIDENCE with the operator '*' indicating set intersection and the capital

letters indicating the presence of a characteristic[8]. Similarly, we use small letters to indicate the

absence of a characteristic; row 2 could, thus, be written as male*MATHS ABILITY*MATHS

CONFIDENCE.

---

[5] More details about the specific calibration process for this analysis are provided later in the paper, in the results section.

[6] A corresponding analysis could be done for 'low maths disposition'. A case's score in the set 'low maths disposition' can be calculated from the formula: 1-(score in 'high maths disposition').

[7] Note here that QCA will allocate cases to rows in the truth table (or to analytic types) based on their set membership scores in different sets. This will produce 'clusters' of cases of a similar type but the allocation of cases to types is not affected by the distribution of set-membership scores as it would be if conducting a cluster analysis. See Cooper and Glaesser (2011) for a detailed exploration of the differences between fsQCA and fuzzy cluster analysis.

[8] This is standard QCA notation and will be adopted throughout.

[Table 1 about here]

The rows of Table 1, representing particular, separate configurations of characteristics, can be thought of as new sets and a case's membership score in them can be calculated as the minimum of its membership scores in the component sets[9]. The score of a case in row 1, MALE*MATHS ABILITY*MATHS CONFIDENCE, for example, would be the minimum of its scores in the sets 'Male', 'Maths ability', and 'Maths confidence'. The consistency measure for Maths disposition tells us the extent to which these new sets, formed from intersections of the original sets, coincide with the outcome: here, a measure of maths disposition. Remember that each case has set membership scores for 'Male', 'Maths ability', 'Maths confidence', and 'Maths Disposition'. If all of these were crisp sets, then the consistency of row 1 for 'Maths Disposition' would simply be the proportion of all cases in row 1 that also have a score of '1' for 'Maths disposition'. When considering fuzzy sets, the consistency formula:

$$\text{consistency } (X_i \leq Y_i) = \Sigma \min(X_i, Y_i) / \Sigma(X_i)$$

calculates, for each truth table row, the score of a case in that row and assesses this against the case's score for the outcome set, for all cases in the dataset. A consistency score of 1 would indicate that, all cases of that configuration, have smaller membership scores for the configuration than those for the outcome. A consistency score of 0 indicates that no cases of a configuration are smaller than that of the outcome. A consistency score between 0 and 1 indicates that some cases of a configuration have smaller score than that of the outcome – where the closer the score to 1, the higher the number of cases fulfilling this criterion.

As analysts, we must make a judgement about what level of consistency threshold is acceptable to indicate quasi-sufficiency and this affects the results we produce[10]. These ideas are summarised in

---

[9] As explained earlier, this is not the same as a creating a multiplicative interaction term in a regression model.
[10] Throughout, we will refer to all situations where the sufficiency relationships deduced are not perfect as quasi-sufficiency

14

Figure 1 where X can be any configuration of characteristics and Y is an outcome set., The situation in Figure 1a cannot be thought of as quasi-sufficient as more of X is out of Y than in it (which would be indicated by a consistency score of less than 0.5). Choosing what score *is* acceptable to indicate quasi-sufficiency, is not always straightforward, however (Ragin, 2000). We discuss the implications of this for our data in a later section.

[Figure 1 about here]

Generating and interpreting solutions
As mentioned earlier, one of the benefits of using fs/QCA is that it can highlight asymmetric relationships between characteristics and outcomes. This, however, means that, when using fuzzy sets, the analysis for one outcome (e.g. high dispositions) cannot be used to derive a solution for the absence of the outcome (e.g. low dispositions). For example, Table 1 is a truth table showing the degree to which configurations are sufficient for high maths dispositions. It would be wrong to immediately conclude from Table 1 that row 8 is an example of a configuration leading to low dispositions. This is because, if we run an analysis for the absence of the outcome, i.e. low maths dispositions, we could find that the configuration in row 8 of Table 1 does not have a consistency score high enough to warrant being described as sufficient for that outcome[11].

If we select a consistency threshold of 0.85, we find that there are 2 rows which have consistency scores higher than this and therefore can be considered quasi-sufficient for Maths disposition. These are the rows representing the configurations MALE*MATHS ABILITY*MATHS CONFIDENCE and male*MATHS ABILITY*MATHS CONFIDENCE (rows 1 and 2 in Table 1). To simplify, we can first produce a solution of the form

MALE*MATHS ABILITY*MATHS CONFIDENCE + male*MATHS ABILITY*MATHS CONFIDENCE → MATHS DISPOSITION, (where '+' represents logical OR);

---

[11] To do this, fuzzy set scores in the 'negated outcome' are calculated from the formula: (1-score in outcome). Then analysis proceeds as above.

which shows that either boys or girls with high maths ability and high mathematical confidence consistently have high mathematical dispositions. It is possible to simplify this further by noticing that a case's score in the set 'MALE' is not making a difference here – both 'MALE' and 'male' (i.e. female) are, in conjunction with high mathematical ability and high mathematical confidence (MATHS ABILITY*MATHS CONFIDENCE), is quasi-sufficient for high maths disposition.

So, our simplified solution is just MATHS ABILITY*MATHS CONFIDENCE → MATHS DISPOSITION. If we set the threshold at 0.8 instead, we would produce the simplified solution MALE*MATHS ABILITY + MATHS ABILITY*MATHS CONFIDENCE → MATHS DISPOSITION – a very different solution obtained by the decision to include only one more row than previously. Software packages for fs/QCA perform this act of Boolean simplification for more complicated truth tables by considering pairs of quasi-sufficient truth table rows together and eliminating redundant parts of terms until the simplest expression is found. This simple example reminds us, however, that it is our decisions about consistency thresholds that will determine which solution is generated.

How should we, then, evaluate which solution should be chosen as a good representation of the data? Each term (e.g each configuration) in a solution can be thought of as one route to the outcome. In situations such as the example above, it is helpful to know the 'empirical importance' of each of these routes as this can help us determine which of the solutions best explains the outcome (Ragin, 2006b p299). This can be found by calculating the 'coverage' of a configuration – a measure of how much of the outcome set coincides with the configuration in question. The coverage measure in QCA for the whole solution can be thought of as analogous to R^2 measures in linear (regression) models as it gives a measure of how much of the outcome set is being explained by the configurations in the model. QCA solutions will also contain details of raw and unique coverage of each part of the solution or solution term. The unique coverage is the coverage specific

16

to that term and so solution terms which have an element in common, such as MALE*MATHS

ABILITY + MATHS ABILITY*MATHS CONFIDENCE from earlier, will have a unique coverage

of 0. The raw coverage for solution terms is the amount of the outcome covered by that term.

Solution terms which do not overlap with others will have equal raw and unique coverage scores.

Figure 2 shows two instances of perfect sufficiency – one where coverage is low and another

where it is high. Both these instances would have consistency scores of 1 but Y is being explained

better by X in the right-hand diagram where the coverage score is higher and more of the set X is

'covered' by Y.

[Figure 2 about here]

3. Results

As already mentioned this analysis is focused on data from a complete case analysis of 5022

students in Years 7 to 11, with mathematics dispositions as the main outcome. In order to conduct

QCA each case must have information in all of the possible characteristics – otherwise it is not

possible to create a set-membership score in all sets and assign the case to its proper row in the truth

table. Before the presentation of the main results we illustrate in more detail our calibration

approach which was performed with the full DP1 dataset (N=13,643).

Calibration of variables/measures

The first step in QCA is to create sets (i.e. calibrate the measures). We outline our method and

rationale below for each set in turn.

Gender is coded as a crisp set called 'MALE' where a score of 0 indicates girls and a score of 1

indicates boys. Our sample has a higher percentage of girls (59%) to boys (41%) but by creating

this set we are able to see if relationships between transmissionist teaching practices and

dispositions are different for girls and boys.

17

We use crisp sets as year dummies to investigate whether the results differ by year group and include four of these in the model (for years 7-10) since a score of zero in all of these will indicate that cases are in year 11[12]. Ordinal variables cannot be translated into fuzzy sets in QCA as doing so would assign a distance between levels that does not necessarily exist in the original variable.

Fuzzy-sets are then created for each of the attitudinal measures of mathematics identity, mathematical self-efficacy and mathematics disposition using the QCA package in R (Dusa, 2019). As noted earlier, for continuous measures the researcher must choose an approach to calibration that is meaningful in their context. In our dataset, these measures have been constructed using the Rasch model, which produces interval measures in logits, with scores ranging from negative values through to positive values centred around zero. If a student in our dataset, for example, scores 0 on the measure of mathematics disposition, this can be taken to be indicative that that student has neither high nor low mathematics dispositions.

As the starting point of our calibration process we argue that the concept of a zero score in a Rasch measure (i.e. the expected centre of its distribution) is analogous to the concept of a set membership score of 0.5 in fuzzy set theory and so, for all our measures, we translate scores of zero in the original measure to 0.5 in the resulting fuzzy-set. The next step in calibration is to decide the upper and lower thresholds in the original measures that will correspond to complete set-membership or non-membership of the sets. We set these boundaries at less than/more than the mean plus 1 standard deviation – indicating that these cases are far enough away from the actual centre value to indicate complete set non-membership or full set-membership respectively This calibration decision was the result of an iterative process where different calibration thresholds were tried and discussed by the authors using graphical methods to help understand their differences. It is this researcher

---

[12] The alternative is to include all years as crisp sets but then have many redundant rows which contain non-possible scenarios e.g. being in years 7 and 8.

18

control over calibration that provides an important qualitative dimension to QCA and is often the part that researchers find most challenging (Russo and Confente, 2019).

Our aim, when choosing the final thresholds was 1) to minimise the effect of very high and low original measures on the calibration, 2) to use the information contained in the original Rasch scores to determine what Rasch scores equate to 'very high' or 'very low' set-membership scores and 3) to account for the Rasch scores in our complete cases sample dataset not being (perfectly) normally distributed (i.e. with scales where the mean was not the expected 0). Our decision to set the thresholds for full set-membership and non-membership to the mean + 1 standard deviation fulfils these criteria. It is well known that the calibration chosen can effect results (Glaesser and Cooper, 2014; Skaaning, 2011) and so we also checked the effect of changing this slightly on the results generated to ensure they were stable. This transformation of the Rasch scores into fuzzy sets is demonstrated in Figure 3 for the maths disposition measure.

Practically, once the thresholds were chosen we used the QCA package in R (Dusa, 2019) to calibrate our fuzzy sets. This function requires 4 inputs – the data to be calibrated and the values corresponding to 0, 0.5 and 1 in the new set. With these inputs, it uses the empirical cumulative distribution function to create the new set membership scores. It is important to note here that this function will not automatically assign a score of exactly 0 or 1 so some variation is still preserved at the extremes of the distribution of set membership scores. For our purposes, we were content to retain some distinction between the extreme values.

[Figure 3 about here]

Figure 3 illustrates clearly that set calibration is not merely a straightforward translation of the data to the (0,1) interval. In our calibration, for example, several values at the extremes of the original distribution have been calibrated to have set membership scores very close to 0 or 1. We do this because we think, for example, that very small values on the original disposition measure indicate roughly the same thing and it would be unwise to exaggerate the differences between them in our

19

calibrated measure. However, because we considered 0 as the expected centre of the scores and thus denoted values above or below zero as qualitatively different (i.e. high and low dispositions respectively), we want the differences between these to be reflected in our final calibration.

Solutions

Next, we organise the cases in the dataset into a truth table (not shown) according to their set-membership scores in each of our created sets and then solutions are generated from the truth tables using the approach outlined earlier. For both high dispositions and low dispositions, solutions were generated with a consistency threshold of 0.68. Choosing a consistency threshold is another qualitative decision by the researchers and we note here that ours is lower than is usual in the (predominantly small-n) QCA literature which is reflected in the strength of the relationships detected. As discussed earlier, there is a trade-off between consistency and coverage with high consistency often yielding very low coverage scores. This was the case in our analysis where higher consistency thresholds either yielded no solutions or solutions about one year group only.

Both these analytic processes were carried out using the QCA package in R (Dusa, 2019). The results for both the outcome (high disposition) and the negated outcome (low dispositions) are presented in Tables 2 and 3, respectively.

[Table 2 about here]

Table 2 shows that there are different solutions for different year groups – confirming our suspicion that the associations found in earlier work [Author citation] provide better explanations for some year groups than others. For those in years 7 and 8, low scores in the 'student perceived transmissionist' measure are quasi-sufficient for high mathematical dispositions though the consistency score for year 7 is higher than that of year 8 – suggesting that the relationship between low student perceived transmissionist scores and high dispositions is weaker for the year 8 students.

20

This also suggests that if girls perceive their mathematics teaching to be transmissionist at this age, they will not want to keep studying it (as reflected in lower dispositions).

Another similarity between the results of those in year 7 and 8 is that being male is quasi-sufficient for high dispositions – i.e. regardless of their scores in the transmissionist measures, boys generally have high mathematical dispositions at this educational stage. Again, as with the student perceived transmissionist measure, the strength of the relationship is weaker for those in year 8 than in year 7. This should not be taken to mean that there are not girls in these year groups with high mathematical dispositions – merely that there is no overall relationship between the terms in the model that holds in general for girls. In other words, the girls in the group show sufficiently different relationships between their perceptions of teaching practices and their dispositions that a good summary for the group of girls in these years **overall** is not possible.

Similarly, there are no quasi-sufficient relationships in the solution for students in years 9 and 11. Again, this should not be taken to mean that there are not students in years 9 and 11 with high dispositions but that the factors we have modelled have an inconclusive relationship with mathematical dispositions for these year groups.

For students in year 10, there is a more complex relationship between the measures of perceived transmissionism, gender and dispositions. For male students in this year, it is quasi-sufficient for high dispositions to have either low scores in the student transmissionist measure or the teacher one. This is the only year group where the teacher measure shows any sort of relationship with dispositions which could explain why it appeared to be not statistically significant in our earlier work using regression modelling [Author citation]. Considering the sample as a whole in a regression model would show such a small association between this teacher perceived transmissionism and dispositions because there is no clear relationship except for students in year 10.

21

Considering the solution as a whole, the coverage score of 0.58 means that most of the individual cases have been explained using one (or more) of the individual solution terms in Table 2[13]. However, many cases still have not and this will be largely due to there being no terms in the solution that explain high dispositions for students in years 9 and 11.

[Table 3 about here]

Table 3 shows the results for an analysis of low dispositions. Again, we find results that differ by year group but that, unlike for high dispositions, there are some relationships for each year group in the sample.

For female students in year 7, having high scores in the teacher perceived transmissionist and the student perceived transmissionist measure is quasi-sufficient for low dispositions. This is interesting because it suggests that high scores in either is not quasi-sufficient for low dispositions but high scores in both. In earlier work, we found that the younger students in the sample generally had higher dispositions towards mathematics than those in years 10 and 11 [Author citation]. Therefore it makes sense that those with low dispositions are those with a particular set of circumstances – high teacher and student perceptions of transmissionist teaching. In this case, it is boys who have an inconclusive relationship between their transmissionist scores and their dispositions.

For female students in year 8, having a high score only in the student perceived transmissionist measure is quasi-sufficient for low dispositions. Considering the results for high dispositions for year 8 students along with these results we see that, for girls in year 8, there is an almost symmetrical relationship between their perception of whether teaching is transmissionist and their dispositions. In other words, girls in year 8 with perceptions that the teaching they receive is not very transmissionist have high dispositions but those who perceive their teaching to be

---

[13] Note that this does not equate to 58% of cases.

transmissionist have low dispositions. For boys, the relationship between their perceptions of transmissionist teaching and dispositions is asymmetrical: when they perceive the teaching they encounter to have low levels of tranmissionism, they have high dispositions but perceived high level of transmissionism does not necessarily lead to low dispositions for them.

For female students in years 9 and 10, the transmissionist or otherwise nature of teaching has no relationship with their dispositions – these are low regardless. This is also the case for male students in year 11 though the strength of the consistency relationship is weaker than it is for the girls in years 9 and 10. Further work is needed to explore these results as they show a possible relationship between curriculum stage, gender and dispositions.

In Table 3, we see that students' perceptions of high student transmissionism is quasi-sufficient for low dispositions for both boys and girls in years 9 and 11. For students in year 11, high levels of teacher transmissionism alone are quasi-sufficient for low dispositions. We suggest possible reasons for this in the discussion section.

4.   Discussion

Methodologically, our contribution in this paper has thus been two-fold: (a) we demonstrated how a set-theoretic approach such as QCA can be used in education and in particular with dispositional measures to inform about their configuration/formation in students' secondary education, and (b) we have also showed for the first time how this approach can be combined with the results of measurement models (i.e. Rasch). One might ask regarding the former: what's the extra benefit of using this approach? Our earlier work using only linear models (as well as the model presented in Appendix 1), did not alert us to the asymmetry shown earlier. In particular, relationships between, for example, gender and dispositions for different year groups were averaged out in the previous

23

linear model – leading us to conclude that girls had lower dispositions **in general**.   Table 3 shows us that, for students in year 11, the opposite relationship holds.

In our QCA, the process of calibration has been aided by the use of measurement theory which was employed to construct the original measures used in the model. Thus, we can proceed to calibrate the measures into fuzzy sets using some external criteria knowing that this will be meaningful. Without this, we would have to consider a relative calibration of the measures – obtained by equating 0.5 in the fuzzy set with some centre-point in each measure (e.g. the mean). Any interpretation of the results would have to reflect this – for example, we would talk of 'high mathematical disposition *for this group'* rather than 'high mathematical disposition' in general. It would then be hard to say what the applicability of the results would be to other populations of students. The use of the Rasch scores addresses thus a key methodological issue that arises when QCA to large sample size datasets, which is how to replicate the kind of case-level familiarity that is possible with small sample size datasets and avoid error in the coding of cases (see Fiss et al., 2013). As scores constructed through Rasch measurement have already undergone a process of validation, the conversion of item responses into a score makes this score robust quantitatively and also meaningful qualitatively.

Given the novelty of the approach and its implementation in educational research there are some further methodological issues to be noted.

The first point is the unavoidable comparison of the method with more straightforward approaches like regression models. For the interested reader we provide the results of twoa regression models with the same variables and with the same sample in Appendix 1 – one with interaction terms and one without - but opt out of direct comparisons as this cancels out our initial intention (i.e. to show the potential of combining Rasch measurement with QCA to explain complicated relationships between attitudes to mathematics) and these models are not directly comparable with a QCA model. It is worth mentioning however that interaction terms in regression models behave slightly

24

differently to the configurations in QCA as interaction terms assume a symmetric relationship whereas conjunctions of characteristics do not (see Thiem et al., 2016 for a very detailed and extended comparison of the two).

Debate is also ongoing between QCA users as to whether multilevel data, such as this, needs to be analysed differently or can be analysed using a standard QCA approach. Denk (2010) was the first, to our knowledge, to propose an adapted QCA procedure for multilevel data but, as Rohlfing (2012) argues persuasively, this can often lead to unnecessarily complex solutions which offer little added insight balanced against their more challenging interpretation. Rohlfing (2012) also argues, as would we, that standard QCA shows if any configurations across systems (in this case, schools) typically have a relationship with the outcome in question.

Substantively, one could ask 'what this novel analytical approach adds to our knowledge about students' mathematics dispositions at secondary school?' Our configurational analysis indicates interesting substantial findings regarding the formation of maths disposition at the different stages of formal secondary education. First, we see that transmissionist teaching does not necessarily lead to high or low dispositions **in general** – but that in conjunction with other characteristics and particular school stages, both transmissionist and non-transmissionist teaching practices are associated with both high and low dispositions.

As with previous results, student perceptions of whether teaching is transmissionist or not seems to be more strongly connected with dispositions [Author citation]. However, the approach taken here allows us to understand the complexity of this relationship. In earlier work, we found that the year 7 students in the sample generally had higher dispositions towards mathematics than those in years 10 and 11 [Author citation]. Our results here show that year 7 students with low dispositions are typically female students with high student and teacher perceptions of transmissionist teaching and whilst they may not make up the majority of the year 7 cohort, their experiences can be pinpointed using this approach.

25

Second, we can also identify that, for particular sub-groups of students, teacher perceptions of transmissionist practices may explain dispositions. The association of teaching related variables with students' disposition has been recently ~~reported~~ explored ~~by others, our team's work~~ in post-secondary maths [Author citation] and in regards to its complex effect on mathematics self-efficacy (Lei and Pampaka, 2018;2019). Riconscente (2014) also recently found that teacher-related variables contribute to students' interest, self-efficacy and achievement after controlling for demographics and initial levels of interest and self-efficacy, with a sample of Latino students in USA.

Third, the relationship between the teacher measures and dispositions seems to be strongest for year 11 and may be related to the increased emphasis on formal assessment in the later stages of secondary school. In England, GCSEs are taken by most students in year 11 but students will study towards them from year 10. For students in year 11, high levels of teacher transmissionism alone are quasi-sufficient for low dispositions and this may be a result of the pressure on teachers to prepare students for high-stakes examinations in this year.

5. Conclusion

Our paper has shown that the relationship between student ~~attitudes~~perceptions of teaching/learning experience, teaching practices and mathematical dispositions is complex. Such complexity is common in educational research and a challenge for researchers to explore; nevertheless, this complexity can be understood by employing QCA as an analytical technique especially when used in combination with other qualitative and quantitative methods. This paper is seminal in that it introduces the fs/QCA with a practical real example with a large dataset and also in combination with interval measures. As the measures themselves have been validated using the Rasch model, our use of them here allows us to say something in general about student dispositions and the relationships with perceptions of teaching/learning experience ~~attitudes~~ and perceived practices.

We have discussed in depth how this method works and drawn attention to a particular stage of analysis – the calibration stage – that usually requires the researcher to make a judgement about how the raw data should be transformed into a collection of set-membership scores. In our case, this decision is informed by measurement theory as the measures of attitudes and practices in our original dataset have a meaningful zero value and upper and lower bounds that produce sensible analogues for the set-membership scores of 0 and 1.

By exploring the solutions for both high and low dispositions, we find that these relationships are asymmetric and differ by year group and gender. Identifying these differences allows for interventions to be more precisely targeted and suggests that for some groups of students, transmissionist teaching may be more damaging to their dispositions than for others.

27

References

ACME, 2012. Raising the bar: developing able young mathematicians. ACME, London.

ACME, 2009. The Mathematics Education Landscape in 2009. Advisory Committe on Mathematics Education - Royal Society, London.

Allexsaht-Snider, M., Hart, Laurie.E., 2001. "Mathematics for All": How Do We Get There? Theory Into Practice 40, 93–101.

Archer, L., Halsall, A., Hollingworth, S., Mendick, H., 2005. "Dropping out and drifting away": an investigation of factors affecting inner-city pupils' identifies, aspirations and post-16 routes. Institute for Policy Studies in Education, London.

Askew, M., Brown, M., Rhodes, V., Wiliam, D., Johnson, D., 1997. Effective Teachers of Numeracy - Report of a study carried out for the Teacher Training Agency. King's College London, London.

Bingham, A.J., Dean, S., Castillo, J., 2019. Qualitative comparative analysis in educational policy research: Procedures, processes, and possibilities. Methodological Innovations 12, 2059799119840982. https://doi.org/10.1177/2059799119840982

Black, P., Wiliam, D., 1998. Inside the black box: raising standards through classroom assessment. King's College, London.

Boaler, J., 2002. Experiencing School Mathematics: Traditional and Reform Approaches to Teaching and Their Impact on Student Learning. Lawrence Erlbaum Associates, New Jersey.

Bond, T.G., Fox, C.M., 2015. Applying the Rasch Model: Fundamental measurement in the human sciences, 3rd ed. Lawrence Erlbaum Associates Inc., NJ.

Brown, D., 2017. An analysis of the effect of missing data on the results and sample composition of a longitudinal study into maths teaching and learning (MSc). University of Manchester.

Brown, M., Brown, P., Bibby, T., 2008. "I would rather die": reasons given by 16-year-olds for not continuing their study of mathematics:. Research in Mathematics Education 10, 3–18. https://doi.org/10.1080/14794800801915814

Cao, Z., Bishop, A., Forgasz, H., 2007. Perceived Parental Influence on Mathematics Learning: A Comparison Among Students in China and Australia. Educational Studies in Mathematics 64, 85–106.

Cooper, B., Glaesser, J., 2010. Contrasting variable-analytic and case-based approaches to the analysis of survey datasets: exploring how achievement varies by ability across configurations of social class and sex. Methodological Innovations Online 5, 4–23.

Denk, T., 2010. Comparative multilevel analysis: proposal for a methodology. International Journal of Social Research Methodology 13, 29–39. https://doi.org/10.1080/13645570802583779

Di Martino, P., Zan, R., 2011. Attitude towards mathematics: A bridge between beliefs and emotions. ZDM - International Journal on Mathematics Education 43, 471–482. https://doi.org/10.1007/s11858-011-0309-6

Driessen, G., Sleegers, P., 2000. Consistency of Teaching Approach and Student Achievement: An Empirical Test. School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice 11, 57–79.

Dusa, A., 2019. Package 'QCA.

Ernest, P., 1991. The philosophy of mathematics education. Falmer, Basingstoke.

Fennema, E., Sherman, J.A., 1977. Sex-related differences in mathematics achievement, spatial visualisation and affective factors. American Educational Research Journal 14,.

Fiss, P.C., Sharapov, D., Cronqvist, L., 2013. Opposites Attract? Opportunities and Challenges for Integrating Large-N QCA and Econometric Analysis. Political Research Quarterly 66, 191–198.

Gabriel, F., Signolet, J., Westwell, M., 2018. A machine learning approach to investigating the effects of mathematics dispositions on mathematical literacy. International Journal of

Research & Method in Education 41, 306–327.
https://doi.org/10.1080/1743727X.2017.1301916

Glaesser, J., 2008. Just how flexible is the German selective secondary school system? A configurational analysis. International Journal of Research & Method in Education 31, 193–209.

Glaesser, J., Cooper, B., 2014. Exploring the consequences of a recalibration of causal conditions when assessing sufficiency with fuzzy set QCA. International Journal of Social Research Methodology 17, 387–401. https://doi.org/10.1080/13645579.2013.769782

Glaesser, J., Cooper, B., 2012a. Educational achievement in selective and comprehensive local education authorities: a configurational analysis. British Journal of Sociology of Education 33, 223–244.

Glaesser, J., Cooper, B., 2012b. Gender, parental education and ability: their interacting roles in predicting GCSE success. Cambridge Journal of Education 42, 463–480.

Grootenboer, P., Hemmings, B., 2007. Mathematics performance and the role played by affective and background factors peter grootenboer and brian hemmings. Math Ed Res J 19, 3–20. https://doi.org/10.1007/BF03217459

Hannula, M.S., 2002. Attitude towards mathematics: emotions, expectations and values. Educational Studies in Mathematics 49, 25–46. https://doi.org/10.1023/A:1016048823497

Hemmings, B., Kay, R., 2010. Prior achievement, effort, and mathematics attitude as predictors of current achievement. The Australian Educational Researcher 37, 41–58.

Kleanthous, I., Williams, J., 2013. Perceived parental influence and students' dispositions to study mathematically-demanding courses in Higher Education. Research in Mathematics Education 15, 50–69. https://doi.org/10.1080/14794802.2013.763608

Lapointe, J.M., Legault, F., Batiste, S.J., 2005. Teacher interpersonal behavior and adolescents' motivation in mathematics: A comparison of learning disabled, average, and talented students. International Journal of Educational Research 43, 39–54.

Lazarides, R., Ittel, A., 2012. Instructional Quality and Attitudes toward Mathematics: Do Self-Concept and Interest Differ across Students' Patterns of Perceived Instructional Quality in Mathematics Classrooms? Child Development Research 2012, 11. https://doi.org/10.1155/2012/813920

Lei, K.H., Pampaka, M., 2019. Mathematics Self-efficacy of secondary school students in the UK: The role of parental support and perceived teaching practices, in: Proceedings of the PME & Yandex Russian Conference. Presented at the PME, PME, Moscow, Russia.

Lei, K.H., Pampaka, M., 2018. Mathematics self-efficacy and its association with parental support and teaching practices in UK secondary school mathematics, in: Curtis, F. (Ed.), Proceedings of the British Society for Research into Learning Mathematics 38 (3) November 2018. Presented at the BSRLM, BSRLM, pp. 1–6.

Lim, S.Y., Chapman, E., 2013. Development of a short form of the attitudes toward mathematics inventory. Educ Stud Math 82, 145–164. https://doi.org/10.1007/s10649-012-9414-x

Mata, M. de L., Monteiro, V., Peixoto, F., 2012. Attitudes towards Mathematics: Effects of Individual, Motivational, and Social Support Factors. Child Development Research 2012, 10. https://doi.org/10.1155/2012/876028

McCaffrey, D.F., Hamilton, L.S., Stecher, B.M., Klein, S.P., Bugliari, D., Robyn, A., 2001. Interactions among Instructional Practices, Curriculum, and Student Achievement: The Case of Standards-Based High School Mathematics. Journal for Research in Mathematics Education 32, 493–517.

Meece, J.L., Herman, P., McCombs, B.L., 2003. Relations of learner-centered teaching practices to adolescents' achievement goals. International Journal of Educational Research 39, 457–475.

Mendick, H., 2005. Mathematical stories: why do more boys than girls choose to study mathematics at AS-level in England? British Journal of Sociology of Education 26, 235–251. https://doi.org/10.1080/0142569042000294192

29

Messick, S., 1988. The once and future issues of validity: Assessing the meaning and consequences of measurement, in: H. Wainer, H. I. Braun (Eds.), Test Validity. Lawrence Erlbaum Associates, Publishers, London, pp. 33–45.

Mevarech, Z., Kramarski, B., 2014. Critical Maths for Innovative Societies.

O'Brien, V., Martinez-Pons, M., Kopala, M., 1999. Mathematics self-efficacy, ethnic identity, gender, and career interests related to mathematics and science. Journal of Educational Research 92, 231–235.

OFSTED, 2011. Tackling the challenge of low numeracy skills in young people and adults. OFSTED, London.

Pampaka, M., Swain, D., Jones, S., Williams, J., Edwards, M., Wo, L., 2018. Validating constructs of learners' academic self-efficacy for measuring learning gain. Higher Education Pedagogies 3, 118–144. https://doi.org/10.1080/23752696.2018.1454264

Pampaka, M., Williams, J., Hutcheson, G., Wake, G., Black, L., Davis, P., Hernandez-Martinez, P., 2012. The association between mathematics pedagogy and learners' dispositions for university study. British Educational Research Journal 38, 473–496. https://doi.org/10.1080/01411926.2011.555518

Pustjens, H., Van de gaer, E., Van Damme, J., Onghena, P., 2004. Effect of Secondary Schools on Academic Choices and on Success in Higher Education. School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice 15, 281–311.

Ragin, C.C., 2008a. Fuzzy Sets: Calibration versus Measurement, in: Box-Steffensmeier, J., Brady, H., Collier, D. (Eds.), The Oxford Handbook of Political Methodology. Oxford University Press, Oxford.

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

Ragin, C.C., 2006a. The Limitations of Net-Effects Thinking, in: Rihoux, B., Grimm, H. (Eds.), Innovative Comparative Methods for Policy Analysis. Springer, New York, pp. 13–41.

Ragin, C.C., 2006b. Set Relations in Social Research: Evaluating Their Consistency and Coverage. Political Analysis 14, 291–310. https://doi.org/10.1093/pan/mpj019

Ragin, C.C., 2000. Fuzzy-set Social Science. University of Chicago Press, Chicago.

Riconscente, M.M., 2014. Effects of Perceived Teacher Practices on Latino High School Students' Interest, Self-Efficacy, and Achievement in Mathematics. The Journal of Experimental Education 82, 51–73. https://doi.org/10.1080/00220973.2013.813358

Rihoux, B., Ragin, C.C., 2009. Configurational Comparative Methods, Applied Social Research Methods Series. Sage, London.

Rohlfing, I., 2012. Analyzing multilevel data with QCA: a straightforward procedure. International Journal of Social Research Methodology 15, 497–506.

Ruffell, M., Mason, J., Allen, B., 1998. Studying attitude to mathematics. Educational Studies in Mathematics 35, 1–18. https://doi.org/10.1023/A:1003019020131

Russo, I., Confente, I., 2019. From dataset to qualitative comparative analysis (QCA)—Challenges and tricky points: A research note on contrarian case analysis and data calibration. Australasian Marketing Journal (AMJ) 27, 129–135. https://doi.org/10.1016/j.ausmj.2018.11.001

Ryan, J., Williams, J. (Eds.), 2007. Children's mathematics 4 - 15. Open University Press, Milton Keynes.

Schuh, K.L., 2004. Learner-centered principles in teacher-centered practices? Teaching and Teacher Education 20, 833–846.

Skaaning, S.-E., 2011. Assessing the Robustness of Crisp-set and Fuzzy-set QCA Results. Sociological Methods & Research 40, 391–408. https://doi.org/10.1177/0049124111404818

Smith, A., 2017. Report of Professor Sir Adrian Smith's review of post-16 mathematics. DfE, London.

Swan, M., 2006. Designing and using research instruments to describe the beliefs and practices of mathematics teachers. Research in Education 58–70.

Swanson, C.B., Stevenson, D.L., 2002. Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP State Assessments. Educational Evaluation and Policy Analysis 24, 1–27.

Thiem, A., Baumgartner, M., Bol, D., 2016. Still Lost in Translation! A Correction of Three Misunderstandings Between Configurational Comparativists and Regressional Analysts. Comparative Political Studies 49, 742–774. https://doi.org/10.1177/0010414014565892

Thomson, S., 2011a. The problem of limited diversity in Qualitative Comparative Analysis: a discussion of two proposed solutions. Journal of Multiple Research Approaches 254–268.

Thomson, S., 2011b. Exploring when maternal interest is sufficient for high attainment in mathematics: A configurational analysis using longitudinal data. Methodological Innovations Online 6, 142–163.

Vaisey, S., 2007. Structure, Culture, and Community: The Search for Belonging in 50 Urban Communes. Am Sociol Rev 72, 851–873. https://doi.org/10.1177/000312240707200601

Vermunt, J.D., Verloop, N., 1999. Congruence and friction between learning and teaching. Learning and Instruction 9, 257–280.

Watson, P., 2011. Disposition towards engagement in mathematics, in: Proceedings of the British Society for Research into Learning Mathematics.

Wilkins, J.L.M., Ma, X., 2003. Modeling Change in Student Attitude Toward and Beliefs about Mathematics. The Journal of Educational Research 97, 52–63.

Wolfe, E.W., Smith Jr., E.V., 2007a. Instrument development tools and activities for measure validation Using Rasch Models: Part I - Instrument development tools. Journal of Applied Measurement 8, 97–123.

Wolfe, E.W., Smith Jr., E.V., 2007b. Instrument development tools and activities for measure validation using Rasch Models: Part II - Validation activities. Journal of Applied Measurement 8, 204–234.

31

Table 1: Example of a Truth Table

Row number (configurations)	Male	Maths ability	Maths Confidence	Number of cases	Consistency for high Maths disposition
1	1	1	1	453	0.89
2	0	1	1	224	0.85
3	1	1	0	130	0.81
4	1	0	0	194	0.63
5	0	0	1	12	0.52
6	1	0	1	219	0.48
7	0	0	0	98	0.47
8	0	1	0	484	0.12

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Results for high dispositions

	Raw coverage	Unique coverage	consistency
student transmissionist*YEAR7	0.27	0.15	0.73
MALE*YEAR7	0.16	0.03	0.63
student transmissionist *YEAR8	0.20	0.10	0.66
MALE*YEAR8	0.13	0.04	0.56
MALE* student transmissionist * YEAR10	0.04	0.01	0.66
MALE*teacher transmissionist*YEAR10	0.04	0.01	0.67
Whole solution	0.58	n/a	0.63

Table 1: Results for low dispositions

	Raw coverage	Unique coverage	consistency
male*TEACHER TRANSMISSIONIST*STUDENT TRANSMISSIONIST*YEAR7	0.06	0.06	0.72
male* STUDENT TRANSMISSIONIST *YEAR8	0.10	0.10	0.71
male *YEAR9	0.14	0.05	0.61
STUDENT TRANSMISSIONIST *YEAR9	0.13	0.04	0.74
male*YEAR10	0.11	0.11	0.67
MALE *YEAR11	0.06	0.01	0.58
TEACHER TRANSMISSIONIST *YEAR11	0.10	0.01	0.74
STUDENT TRANSMISSIONIST *YEAR11	0.10	0.01	0.75
Whole solution	0.58	n/a	0.67

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

a. Low consistency score
➔ X not sufficient for Y

b. High consistency score
➔ X quasi-sufficient for Y

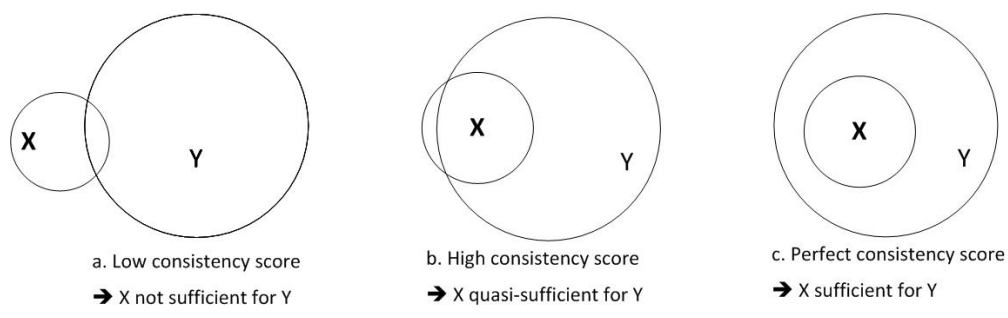c. Perfect consistency score
➔ X sufficient for Y

Figure 1 – Venn diagrams representing various degrees of sufficiency of X for Y
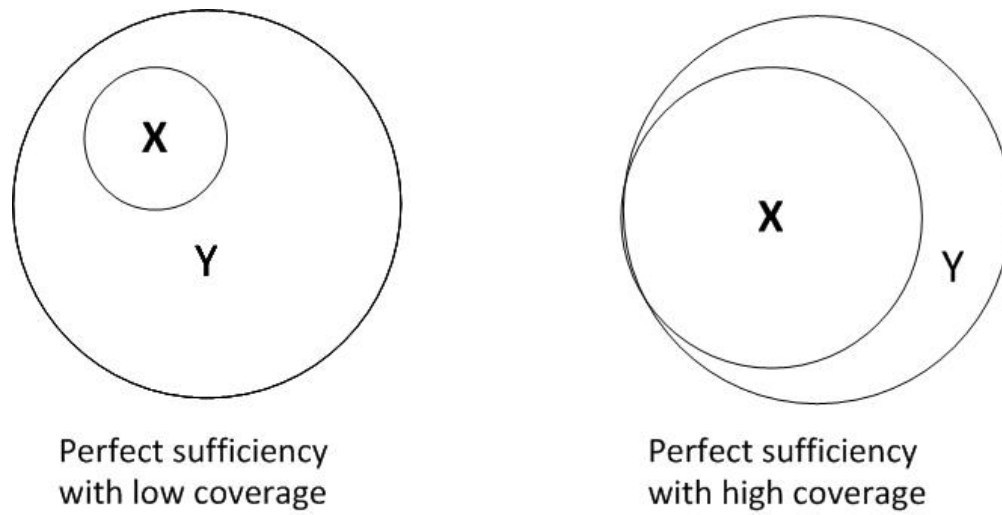
304x90mm (96 x 96 DPI)

Figure 2: Venn diagrams showing low and high coverage

175x88mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
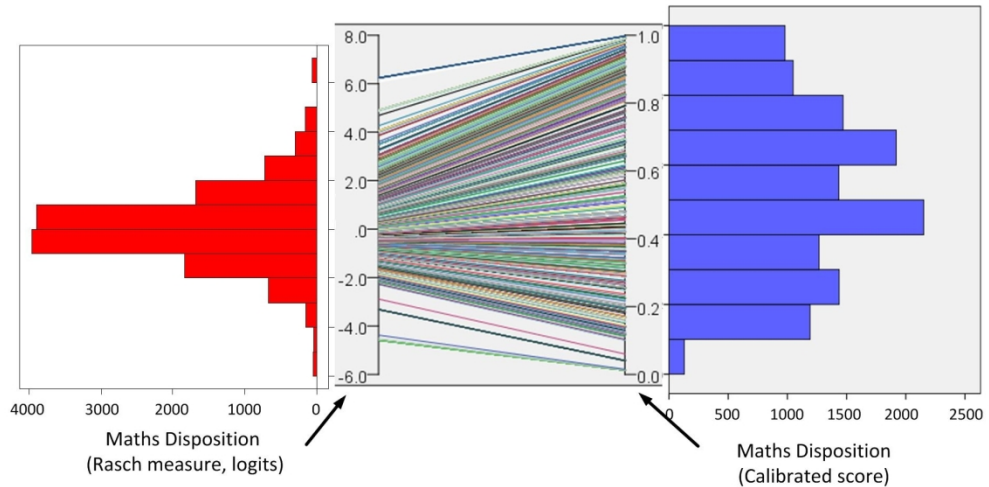49
50
51
52
53
54
55
56
57
58
59
60

Figure 3: Distribution of original Rasch score and calibrated fuzzy set for maths disposition

502x250mm (96 x 96 DPI)

Appendix 1

	Coefficient	Standard error	t	p
Gender (Ref: Male)	-0.289	0.043	-6.7	<0.001
Year Group (Ref: Year 7)				
Year 8	-0.240	0.056	-4.26	<0.001
Year 9	-0.619	0.062	-9.92	<0.001
Year 10	-0.486	0.069	-7.05	<0.001
Year 11	-0.668	0.074	-9.09	<0.001
Students Transmissionism	-0.680	0.041	-16.51	<0.001
Teachers Transmissionism	-0.124	0.038	-3.23	0.001
Constant	0.947	0.047	20.18	<0.001

F(7,5014)=90.13, p<0.001, R-square=0.112, Adjusted R-square=0.111, n=5022

	Model 1				Model 2			
	Coef.	s.e.	t	p	Coef.	s.e.	t	p
Gender (Ref: Male)	-0.289	0.043	-6.7	<0.001	-0.201	0.079	-2.56	0.01
Year Group (Ref: Year 7)								
Year 8	-0.240	0.056	-4.26	<0.001	-0.224	0.087	-2.57	0.01
Year 9	-0.619	0.062	-9.92	<0.001	-0.634	0.103	-6.18	<0.001
Year 10	-0.486	0.069	-7.05	<0.001	-0.159	0.106	-1.5	0.135
Year 11	-0.668	0.074	-9.09	<0.001	-0.648	0.111	-5.85	<0.001
Students Transmissionism	-0.680	0.041	-16.51	<0.001	-0.677	0.041	-16.44	<0.001
Teachers Transmissionism	-0.124	0.038	-3.23	0.001	-0.12	0.038	-3.12	0.002
Constant	0.947	0.047	20.18	<0.001	0.893	0.062	14.51	<0.001
Gender*Year Group (Ref: Male@Year7)								
Female-Year8					-0.023	0.114	-0.20	0.839
Female-Year9					0.015	0.129	0.12	0.907
Female-Year10					-0.561	0.137	-4.08	<0.001
Female-Year11					-0.033	0.145	-0.22	0.823

Model 1: F(7,5014)=90.13, p<0.001, R-square=0.112, Adjusted R-square=0.111, n=5022
Model 2: F(11,5010)=59.43, p<0.001, R-square=0.115, Adjusted R-square=0.114, n=5022

1