British Educational Research Journal
Vol. ●●, No. ●●, ●● 2019, pp. ●●–●●

DOI: 10.1002/berj.3530

The misallocation of students to academic sets in maths: A study of secondary schools in England

Paul Connolly[a,*], Becky Taylor[b], Becky Francis[b], Louise Archer[b], Jeremy Hodgen[b], Anna Mazenod[b] and Antonina Tereshchenko[b]
[a]*Queen's University Belfast, UK;* [b]*University College London, UK*

Drawing upon data gathered from 9301 Year 7 students (12–13 years old) from 46 secondary schools in England, this study represents the first larger-scale attempt to compare their actual set allocations in maths with the counterfactual position where their allocation to sets is based solely on their prior attainment at the end of primary school [using their Key Stage 2 (KS2) fine-grained scores in maths]. Through such an analysis, the study found that nearly a third of students (31.2%) had been misallocated to lower or higher sets than their KS2 results would have warranted. Beyond this, school setting practices were found to exacerbate differences in set allocation in relation to gender and ethnicity, but not socioeconomic background. The odds of girls being misallocated to lower sets in maths than their prior attainment would warrant was found to be 1.5 times higher than that for boys. Similarly, the odds of Black students being misallocated to lower sets was 2.4 times higher than for White students, whilst the odds of Asian students being misallocated to lower maths sets was 1.7 times higher than for White students. The article concludes by reflecting on the significant role that setting by attainment in secondary school can play in exacerbating already established patterns of educational inequalities in gender and ethnicity.

Keywords: academic setting; academic streaming; academic tracking; educational inequalities

Introduction

Whilst grouping students by attainment is a highly favoured practice in English secondary schools, it has attracted criticism in educational research (Oakes, 1985; Ireson & Hallam, 2001; Marks, 2012). Attainment grouping strategies include streaming, setting and within-class grouping (Taylor *et al.*, 2017). Streaming describes the practice of grouping students according to their perceived general 'ability'[1] across all or most subjects, such that students are taught in the same streamed groups for all lessons (Hodgen, 2011). In contrast, setting is grouping based on attainment in individual subjects, so a student might be in a high set for mathematics and a lower set for English. In the US literature, 'tracking' can refer to both streaming and setting

*Corresponding author. Centre for Evidence and Social Innovation, Queen's University Belfast, 2-8 Fitzwilliam Street, Belfast BT9 6AX, UK. E-mail: paul.connolly@qub.ac.uk

(Gamoran & Nystrand, 1994). Within-class 'ability' grouping is most commonly practiced in primary schools, where children are organised at 'ability tables' within a class containing a wide range of prior attainment (Hallam & Parsons, 2013a; Marks, 2013; Bradbury & Roberts-Holmes, 2017).

It is difficult to establish accurately the prevalence of each of these grouping practices, although there are strong indications that these attainment-based strategies predominate. One study, focusing on 44 English schools in areas of disadvantage, found that 100% of the secondary schools taught mathematics to set groups in all years, while 52.7% of the schools taught Year 7/8 students English in set groups, rising to 77.2% in Year 10 (Dunne *et al.*, 2007). Other sources confirm these high figures, particularly for students in Key Stage 4 towards the end of their time in compulsory schooling (OECD, 2013; Jerrim *et al.*, 2018). The practice of setting has also been extending into primary schools, including in Key Stage 1 (Hallam, 2012; Hallam & Parsons, 2013b; Bradbury & Roberts-Holmes, 2017). Furthermore, in 2015 the Department for Education reported that 34.4% of schools sampled had 'introduced/improved setting or streaming' as a strategy for closing the attainment gap between students receiving Pupil Premium and their peers (DfE, 2015b).

The benefits of grouping by attainment are highly contested. Those who favour such strategies argue that it enables teachers to stretch and challenge 'able' learners and provide support to those who are struggling (DfES, 2005). It is claimed that when groups are more homogeneous, leaders can allocate resources and teachers can design learning activities to meet the needs of students and so maximise progress and attainment (Hallinan & Sorenson, 1987; Dar & Resh, 1994; Cahan *et al.*, 1996). However, it is well established that attainment grouping has little or no overall positive impact on student outcomes (Slavin, 1990; Burris & Welner, 2005; Ireson *et al.*, 2005; Nomi, 2009; Rui, 2009; Higgins *et al.*, 2015; Steenbergen-Hu *et al.*, 2016). While students in the highest attainment groups may make small achievement gains, those in lower groups experience a greater negative effect (Boaler & Wiliam, 2001; Wiliam & Bartholomew, 2004; Burris & Welner, 2005; Chambers, 2009; Higgins *et al.*, 2015) on both attainment and affective measures such as self-confidence (Bartholomew, 2000; Francis *et al.*, 2017b). Additionally, there is evidence to suggest that, in primary schools, attainment grouping contributes to the widening of the gap in achievement between students from disadvantaged backgrounds and their peers (Parsons & Hallam, 2014).

This latter point touches on longstanding concerns regarding the potential impact of attainment grouping in exacerbating existing educational inequalities in relation to social class, ethnicity and gender. Since Jackson (1964) found a stark picture of the segregation of students into streams according to their social backgrounds in England, there has been a number of studies suggesting that certain social groups are more likely to be allocated to lower sets and streams. The schools in Jackson's study reported using a range of sources to stream children, including internal and external assessment data as well as teacher judgements or recommendations. More recent research has found a similar range of strategies (Ireson & Hallam, 2001; Muijs & Dunne, 2010; Taylor *et al.*, 2019), but with a greater emphasis on attainment data than found by Jackson.

However, students from lower socioeconomic backgrounds are still consistently found to be concentrated in lower sets and streams. Ball (1981) found that students in the upper stream were more likely to have parents in non-manual occupations, while those in the lower stream were more likely to have parents in manual occupations. In a study of 44 English schools in areas of disadvantage, low socioeconomic status was a strong predictor of lower set membership (Muijs & Dunne, 2010). In a study of schools in North Carolina, fourth and fifth-grade students from lower socioeconomic backgrounds were over-represented in lower sets (Bosworth, 2013).

It has also been found that Black students are frequently found to be more likely to be allocated to lower sets and streams, both in England and the USA (Hallinan, 1996; Muijs & Dunne, 2010; Moller & Stearns, 2012; Modica, 2015). White students, by contrast, are more likely to be allocated to higher sets (Muijs & Dunne, 2010; Moller & Stearns, 2012; Modica, 2015). In England, Bangladeshi students are more likely to be in lower groups (Muijs & Dunne, 2010), while in the USA Asian-American students are more likely to be in college tracks (Moller & Stearns, 2012). Moller and Stearns (2012) note that their findings persist even when attainment is controlled for. Furthermore, Strand (2012) found evidence that Black Caribbean students were systematically under-represented in entry to higher-tier Key Stage 3 attainment tests relative to their White British peers, and suggested that this may be as a result of within-school factors including attainment grouping and teacher perceptions of student 'ability'.

The evidence in relation to gender appears to be more mixed, with some studies finding that boys are more likely to be allocated to lower streams and sets than girls (Jackson, 1964; Van de Gaer *et al.*, 2006; Hallam & Parsons, 2013a), other studies finding no notable differences (Muijs & Dunne, 2010) and yet others suggesting that boys are more likely to be placed in high-'ability' tracks (Moller & Stearns, 2012).

Whilst these existing studies of inequalities associated with set allocation have been important, they have not been able to disentangle the various factors that lead to these patterns of over- and under-representation by social class, ethnicity and gender. For example, there is clear evidence that inequalities in relation to educational attainment emerge early (Waldfogel & Washbrook, 2010) and are already well established by the end of primary school (Connolly, 2004; Lupton, 2014). Thus, at least some of the inequalities in set allocation that have been found consistently in secondary school can be expected and will reflect the different levels of attainment of students when they enter secondary school. However, socioeconomic gaps for attainment widen as students progress through schooling (Clifton & Cook, 2012), and there is also considerable evidence that decisions made by secondary schools contribute directly to these patterns of inequality (Gillborn & Youdell, 2000). One of the most influential early studies of this in relation to ethnicity in England was conducted by Wright (1986). Her ethnographic case studies of two secondary schools included an analysis of the attainment levels gained by Black and Asian students and then their subsequent allocation to differing examination sets. A clear pattern was evident that, regardless of prior levels of attainment, Black and Asian students tended to be allocated to lower examination sets.

To date, however, there has been no larger-scale quantitative study that has sought to explore each student's prior attainment and how well their subsequent set allocations match this, and to assess whether any patterns exist in relation to the misallocation of students by social class, ethnicity and gender. Drawing on data gathered from 9301 students in 46 secondary schools in England, the present study seeks for the first time to explore the nature and extent of misallocation to maths sets and how this varies by social class, ethnicity and gender.

Methodology

As summarised above, there is clear evidence of an association between set allocation and a student's social class and ethnic background, whereas the picture appears to be a little more contradictory with regard to gender. What the above studies have not been able to determine, however, is the extent to which these patterns simply reflect the prior differences in educational attainment across social class, ethnicity and gender that students enter secondary schools with, compared to the role that the secondary schools themselves may play in further exacerbating these differences. To address this particular issue, the present study draws on data collected from 9301 Year 7 students from 46 secondary schools in England. The study compares the students' actual levels of attainment in maths at the end of primary schooling [via Key Stage 2 (KS2) test results] and their subsequent allocation to maths sets in the first year of secondary school.

These schools form part of a broader cluster randomised controlled trial (RCT) study of the effectiveness of schools adopting 'best practice' in attainment setting, which is described in detail elsewhere (Francis *et al.*, 2017a; Taylor *et al.*, 2017). A total of 126 schools were recruited through a mixture of volunteer and direct 'cold call' approach sampling, then randomised to the intervention and control groups of the RCT. Volunteer-sampled schools were recruited through a traditional and social media campaign by the authors. Direct approach-sampled schools were identified through a stratified random sample then approached by the National Foundation for Educational Research. In total, across both methods, 1006 schools were approached, of which 234 did not agree to participate, 608 did not respond, 3 had closed and a further 35 dropped out after agreeing to take part and prior to randomisation. The sample frame for the random sample included all state secondary schools with Years 7 and 8 present, and excluded Local Authority areas where the authors were focusing on recruitment, as both strategies took place concurrently. Schools were distributed across England. Schools were eligible for the study if they were state funded, non-selective by attainment and already setting in maths in Years 7 and 8. Schools using streaming or mixed-attainment grouping were not eligible. Schools were made aware of an outline of the intervention under evaluation during the recruitment process, as implementation would need to be feasible if randomised to the intervention group. The subset of students and schools that provides the focus for this article represented the control group for the trial and thus were schools that continued 'business as usual' in relation to their existing practices with regard to attainment setting. The data reported here draws on the students' KS2 maths scores achieved in the final year of their primary school (May 2015), as derived from the National Pupil Database (NPD), and then their subsequent allocation to maths sets, as reported by the participating schools. Maths set allocation data were collected from schools in November 2015, by which point all schools had made their initial set allocations.

The overall characteristics of the 9301 students in the sample are summarised in Table 1. The sample is broadly reflective of the national population. In particular, it can be seen that the sample is well balanced in terms of gender and also broadly representative of the national population in relation to ethnicity [where it is reported that, nationally, 76% are White, 10% Asian, 6% Black and 5% mixed; see DfE (2015a, p. 15)]. The sample is also broadly representative in relation to the proportion of disadvantaged students, with 30.6% of the present sample having been eligible for free school meals (FSM) at some point, compared to the nationally reported figure of 32% (DfE, 2015a, p. 14). In addition, the overall characteristics of the 46 schools are summarised in Table 2. The sample is broadly reflective of the national sample of schools. This sample of schools was recruited for the purpose of the trial and, as such, had expressed some interest in adopting ‘best practice’ in attainment grouping. Hence, whilst not fully representative, the sample may be considered a ‘telling case’ in that these schools might be expected, if anything, to be more interested than other schools in allocating students equitably to attainment groups (Mitchell, 1984). Further details on the 46 schools and their approaches to setting are provided in Table 3. As can be seen, the number of maths sets per school varies from two to eight. However, and within this, most schools tended to use between three and five sets for maths (82.6% of schools).

The analysis that follows represents the first attempt to compare the actual allocation of students to sets in maths with the counterfactual case where allocations are based strictly on the students’ prior attainment. Table 4 illustrates how the counterfactual case has been derived using data from one of the sample schools. The table shows the students’ KS2 (fine-grading) scores for maths at the end of their primary school and then their subsequent allocation to one of three maths sets during the first year of their secondary school. As can be seen, of the 190 Year 7 students in this particular school, 90 were allocated to the top set, 54 to the middle set and 46 to the bottom set. It can also be seen that whilst there is a notable relationship between KS2 scores and set allocation, denoted by the grey shading, this is not perfect. There are instances where students have attained the same KS2 scores but were subsequently allocated to different sets. For example, although there were 17 students who had achieved a KS2 score of 32.46, only 11 of these were subsequently allocated to the top set and the remaining 6 were allocated to the middle set.

Based on the existing number of sets within the school, and the numbers of students within each set, the counterfactual case has been used to distinguish between those students in Table 4 who could be considered to have been correctly allocated to a set based solely on their KS2 scores (highlighted by grey shading) and those who could be considered to have been misallocated (the remaining students). The classification of students in this way was achieved using the following method. First, it was noted that for this school, 90 students were allocated to the top set. If allocation were based solely on prior attainment, then the 90 top-scoring students should have been allocated to this top set. However, and because some students achieved the exact same scores in their KS2 tests, it is not possible to identify the top 90 students precisely. Rather, and as can be seen from Table 1, it is only possible to identify the 88

Table 1. Characteristics of the students in the sample

	No.	Valid %
Gender		
Boy	4737	51.3
Girl	4501	48.7
Missing	63	
Total	9301	100.0
Household socioeconomic background		
Higher	2546	48.6
Intermediate	1818	34.7
Lower	880	16.8
Missing	4057	
Total	9301	100.0
Ever eligible for free school meals		
No	6142	69.4
Yes	2711	30.6
Missing	448	
Total	9301	100.0
Ethnicity		
White	4385	74.6
Mixed	481	8.2
Black African	247	4.2
Bangladeshi	144	2.5
Pakistani	144	2.5
Asian Other	115	2.0
Indian	101	1.7
Black Other	68	1.2
Black Caribbean	58	1.0
Chinese	23	0.4
Missing	3535	
Total	9301	100.0

Column percentages may not sum to 100.0% due to rounding.

top-scoring students (i.e. the first six rows of students scoring from 39.00 to 31.92). Of these 88, 76 were allocated correctly to the top set and 12 were misallocated to the middle set.

Beyond these first 88 students, the next six highest-scoring students all achieved the same score of 31.38. This presents a difficulty in relation to relying solely on KS2 scores to identify the remaining two students for the top set. In such cases, and because not all six can be allocated to the top set, these students can be considered 'borderline'. In the absence of any further information on their prior attainment, it is therefore assumed that their scores could lead them to be either allocated to the top set or the middle set and that both options can be considered correct. However, this is a critical assumption and its validity rests on whether each subgroup of students (defined either in terms of gender, social class or ethnicity) classified as borderline is then found to be equally likely to be allocated to the nearest set upwards or downwards. This is an assumption that will be tested in the analysis to follow.

Beyond this top set, it can be seen from Table 1 that 54 students were allocated to the middle set and that, cumulatively, 144 students were allocated either to the top or

Table 2. Characteristics of the schools in the sample

	No.	Valid %
OFSTED grade		
Outstanding	10	21.7
Good	27	58.7
Requires improvement	7	15.2
Inadequate	2	4.3
Total	46	100.0
Student characteristics		
> 10 Black students	13	28.3
≤ 10 Black students	33	71.7
Total	46	100.0
> 10 Asian students	15	32.6
≤ 10 Asian students	31	67.4
Total	46	100.0

Column percentages may not sum to 100.0% due to rounding.

Table 3. Number of maths sets in each of the sample schools

	Number of schools	
Number of sets in the school	No.	%
2	1	2.2
3	11	23.9
4	19	41.3
5	8	17.4
6	3	6.5
7	3	6.5
8	1	2.2
Total	46	100.0

middle sets. However, again, in this case it is not possible to identify the top 144 students precisely. Rather, and as can be seen, it is possible to identify the 141 top-scoring students (i.e. the top 13 rows) and use this to determine whether these 141 students have been correctly allocated or misallocated. However, and beyond this, the same issue arises in that whilst there are three more places remaining to be filled in the middle set, there are six eligible students who all attained the same KS2 score (27.84). Following the same procedure as previously, these six were also considered as 'borderline' and, in the absence of any other information, it was concluded that their scores could have led them to be allocated to either the middle or the bottom set.

Finally, and with only one set remaining, it can be assumed that all the remaining students who attained the lowest KS2 scores would have been allocated to the bottom set if the decision was made based solely on their prior attainment. Overall, and with regard to this school, it can be seen from Table 1 that in relation to the counterfactual case, whilst most students (151 or 79.5%) can be considered to have been allocated

Table 4. An illustration of the relationship between students' KS2 scores for maths and the subsequent maths sets they were allocated to for one secondary school

KS2 maths point score	Maths set allocated to			Total	Cumulative total
	Top set	Middle set	Bottom set		
39.00	19			19	19
35.16	5			5	24
34.38	16			16	40
33.54	18			18	58
32.46	11	6		17	75
31.92	7	6		13	88
31.38	3	3		6	94
31.08	2	1		3	97
30.84	6	9		15	112
29.82	3	11		14	126
28.92		4		4	130
28.74		6	3	9	139
28.02		1	1	2	141
27.84		2	4	6	147
27.66		5	6	11	158
27.12			7	7	165
26.34			11	11	176
25.08			12	12	188
19.80			1	1	189
15.00			1	1	190
Total	90	54	46	190	
Cumulative total	90	144	190		

For this school, there were 58 unique KS2 scores found amongst the 190 students. These categories have been conflated and reduced to 20 for the purposes of illustration and due to the confines of space.
Fine-grading scores used, ranging from 3.00 to 39.00. Taken from the National Pupil Database.
"Note: Grey shading denotes those students that were correctly allocated to a maths set based solely on their prior KS2 scores."

correctly, there were 27 students (14.2%) identified as misallocated and a further 12 (6.3%) who can be considered 'borderline'.

This process has been applied to students in each of the 46 schools in the sample for maths, and regardless of how many sets each school had for each subject. Using this method, each student was then assigned to one of three categories: those who had been correctly allocated (coded '0'); those who had been allocated to a set below that to which they should have been allocated based solely on their prior attainment (coded '−1'); and those who had been allocated to a set above that to which they should have been allocated based solely on their prior attainment (coded '+1').

It is important to reflect on the use of the term 'misallocation' in this context. There are two issues here. First, the term assumes that the creation of academic sets is based, predominantly, on each student's prior academic attainment. As set out above, the analysis to follow is based on KS2 test scores as the measure of prior attainment. In England, these are made freely available to secondary schools, so it is reasonable to regard these as a comprehensive and nationally benchmarked measure for use in secondary schools in assessing pupils' prior attainment. However, this is not without controversy, as many secondary school teachers and leaders do not trust KS2 results; for example, believing that primary schools teach pupils to the test, and/or otherwise manipulate results (Taylor *et al.*, 2019). To this end, while KS2 tests are said to be the predominant measure used by schools in set allocation (Taylor *et al.*, 2019), many schools additionally purchase alternative tests. However, in analysing the accuracy of these tests as predictors of Key Stage 4 outcomes, Treadaway (2013) has shown that KS2 test results remain among the most accurate (and more so than some paid-for tests).

Second, whilst the use of prior attainment, as measured through KS2 test scores, may be the predominant method of allocating students to academic sets, it is accepted that, pragmatically, some minor discrepancies are likely to exist. Such discrepancies may be due, in a small number of cases, to other factors in relation to the specific personal needs and circumstances that a school may take into account when deciding on the allocation of a particular student. Moreover, if particular schools have used their own additional tests then this could also introduce some level of discrepancy.

However, it would be expected that such cases would be small proportionately and also be found to occur randomly across particular subgroups of students. The use of the term 'misallocation' for the purposes of this article is therefore not primarily applied to the analysis of individual students, but to broader systemic trends and patterns. Thus, should there be found to be a large proportion of students whose set allocation does not reflect their prior attainment, then this would indicate that there is a notable level of misallocation occurring in the system. Similarly, if the incidence of misallocation is not randomly spread across the student population but is associated disproportionately with particular subgroups of students, then this again would indicate misallocation at a systematic level.

Whilst the focus of this article is on studying broader trends and patterns of potential misallocation, further details on differences in the schools' approaches to set allocation and their perspectives on where there may be divergence between academic set allocated to and prior KS2 attainment are discussed elsewhere (Taylor *et al.*, 2019).

Attainment at Key Stage 2 and set allocation

Table 5 summarises the students' attainment in KS2 maths. As can be seen, there is a clear correlation between a student's social class background and their levels of attainment in maths, as indicated by proxy measures of both household socioeconomic background and FSM eligibility. There is also a clear pattern in relation to gender, with boys achieving better than girls, on average, in maths. With regard to ethnicity, the picture is more complex but overall differences are evident that tend to reflect the findings from other studies, with the three Black student groups attaining the lowest scores in maths, on average, and students of Chinese and Indian heritage achieving especially highly (Archer & Francis, 2007; Gillborn, 2008; DfE, 2015a).

In relation to set allocation, because the number of sets in each school differs, it is not possible to compare the set levels of students directly. For example, a student who is allocated to a level two set may reflect the fact that they are in the bottom set at

Table 5. Attainment in KS2 maths (fine-grained scores) for students in the sample, by gender, household socioeconomic background, FSM eligibility and ethnicity

	Mean (SD)	No.
All students	29.3 (4.9)	8883
Gender		
Boy	29.9 (5.1)	4462
Missing	29.5 (3.7)	26
Girl	28.7 (4.7)	4395
Household socioeconomic background		
Higher	30.6 (4.5)	2409
Intermediate	29.6 (4.6)	1724
Missing	28.6 (5.2)	3903
Lower	28.4 (4.6)	847
Ever eligible for free school meals		
No	30.0 (4.7)	6136
Yes	27.8 (5.0)	2705
Missing	26.2 (6.9)	42
Ethnicity		
Chinese	32.3 (5.9)	21
Indian	31.6 (4.5)	98
Bangladeshi	30.0 (4.5)	141
Asian Other	29.7 (5.3)	105
White	29.7 (4.7)	4152
Mixed	29.5 (4.8)	462
Pakistani	29.5 (4.3)	137
Other	29.3 (5.2)	150
Missing	28.7 (5.2)	3266
Black Other	28.6 (4.2)	63
Black African	28.3 (5.0)	235
Black Caribbean	28.1 (4.3)	53

one school (if that school has only two set levels) but, conversely, may reflect the fact that they are in one of the top sets at another school (for those schools having six or more set levels). As such, and to gain some sense of how set allocation varies by gender, socioeconomic background and ethnicity, students were organised into three categories: those in the top-level set in their respective school; those in the bottom-level set in their respective school; and those in one of the middle sets in their school. The actual set allocations in maths, broken down by gender, socioeconomic background, FSM eligibility and ethnicity, are summarised in Table 6.

It can be seen that the clear pattern of KS2 attainment in maths by social class is also reflected in set allocations in secondary schools in relation to both household socioeconomic background and FSM eligibility. Similarly, the picture in relation to KS2 attainment by gender is also reflected in subsequent set allocations, with boys being slightly more likely than girls to be allocated to top sets in maths. With respect to ethnicity, it can be seen that the broad rankings of ethnic groups in relation to KS2 attainment in maths, on average, are also reflected in the allocations to sets in secondary school.

Table 6. Allocation of Year 7 students to maths sets by gender, household socioeconomic background and ethnicity (48 schools)

	% Allocated			
	Top set	Middle set(s)	Bottom set	Total (No.)
All	29.6	57.0	13.4	9301
Gender				
Boy	32.1	55.8	12.1	4737
Girl	27.1	58.3	14.6	4501
Missing	23.8	58.7	17.5	63
Household socioeconomic background				
Higher	35.6	54.8	9.7	2546
Intermediate	29.8	56.8	13.4	1818
Missing	27.2	57.3	15.5	4057
Lower	23.2	62.5	14.3	880
Ever eligible for free school meals				
No	33.1	54.9	12.0	6142
Yes	23.5	59.9	16.5	2711
Missing	18.7	68.1	13.2	448
Ethnicity				
Chinese	60.9	39.1	0	23
Indian	55.5	38.6	5.9	101
Asian Other	40.9	50.4	8.7	115
Pakistani	37.5	54.9	7.6	144
Other	35.8	54.1	10.1	159
White	30.1	56.8	13.1	4385
Mixed	28.1	58.2	13.7	481
Missing	28.0	56.8	15.2	3376
Bangladeshi	27.8	65.3	6.9	144
Black African	25.9	64.8	9.3	247
Black Other	20.6	75.0	4.4	68
Black Caribbean	15.5	65.5	19.0	58

Set misallocation

The above descriptive analysis demonstrates that there is a degree of consistency between overall levels of attainment at KS2 in maths at the end of primary school and subsequent allocations to maths sets by secondary schools in Year 7. However, and as noted earlier, such an analysis in itself provides no insight into the potential role that secondary schools might be playing in reducing or exacerbating these differences at KS2 through their own set-allocation practices. To assess this, there is a need to compare the actual set allocations of students in maths with the counterfactual case where they are allocated solely on the basis of their KS2 levels of attainment in maths. Given the specific case of students classified as borderline, as discussed earlier, the analysis to follow has been undertaken in two stages: first through an analysis of all students except those found to be borderline (n = 8271, 94.2%); and then a specific analysis of the subsample of borderline students (n = 508, 5.8%).

All students except those classified as borderline

Regarding the allocation of all students except those who were borderline, Table 7 summarises the proportions of students either allocated correctly or misallocated based solely on their attainment in KS2 maths.

Overall, and based solely on performance in KS2 tests, it can be seen that nearly a third of students were misallocated to maths sets (31.2%). Beyond this, it would appear that boys are slightly more likely to be misallocated upwards than downwards, and conversely, girls are slightly more likely to be misallocated downwards than upwards. The pattern is less clear in relation to socioeconomic background (or FSM eligibility), with the proportions of students from different subgroups being misallocated upwards or downwards tending to be similar. As regards ethnicity, whilst a slightly higher proportion of White students would appear to be misallocated upwards

Table 7. Proportions of Year 7 students misallocated to maths sets by gender, household socioeconomic background and ethnicity (48 schools)

	% Allocated			
	Misallocated upwards	Correctly allocated	Misallocated downwards	Total (No.)
All	15.7	68.9	15.5	8271
Gender				
Missing	20.0	60.0	20.0	25
Boy	16.7	70.3	13.0	4143
Girl	14.7	67.4	17.9	4103
Household socioeconomic background				
Lower	18.0	64.6	17.4	789
Missing	17.0	67.8	15.3	3641
Intermediate	15.6	69.1	15.4	1601
Higher	12.9	72.0	15.1	2240
Ever eligible for free school meals				
Yes	18.8	62.5	18.7	2507
No	14.4	71.7	13.9	5727
Missing	8.1	59.5	32.4	37
Ethnicity				
Black Caribbean	21.6	54.9	23.5	51
Other	18.8	59.0	22.2	144
Mixed	18.7	64.9	16.4	427
Pakistani	18.0	60.2	21.9	128
Bangladeshi	17.5	65.1	17.5	126
Missing	16.5	68.6	14.9	3067
White	15.1	71.0	13.9	3839
Black African	11.6	59.1	29.3	215
Asian Other	11.1	64.7	24.2	99
Indian	8.5	77.7	13.8	94
Black Other	6.7	63.3	30.0	60
Chinese	4.8	76.2	19.1	21

Row percentages may not sum to 100.0% due to rounding.

than downwards, the opposite trend appears to be the case for many, but not all, of the Asian and Black subgroups.

Unfortunately, this type of visual inspection of descriptive statistics can only take us so far. Firstly, and in relation to ethnicity, socioeconomic background profiles tend to differ across subgroups and thus we are not comparing like with like. For example, whilst 14.9% of White students in the present sample come from lower socioeconomic backgrounds, this rises to 30.7% of Bangladeshi students and 31.1% of Pakistani students. As such, it is not possible to determine from the current descriptive comparisons of ethnic groups whether any differences in their tendencies to be misallocated to sets is due solely to their ethnicity and/or is influenced by their socioeconomic background. In this respect, it is important to attempt to disentangle the respective influences of socioeconomic background and ethnicity.

Secondly, it is difficult to make direct comparisons between different subgroups based solely on the percentage figures reported in Table 6. For example, it would appear that those eligible for FSM are more likely to fare better in relation to being misallocated upwards than those who are not eligible for FSM (18.8% compared to 14.4%). However, this simple comparison does not take into account the converse position that those eligible for FSM are also more likely to be misallocated downwards compared to those not eligible for FSM (18.7% compared to 13.9%). As such, there is a need to apply a more sophisticated statistical technique which can identify overall potential differences and tendencies by considering all the data together. Also, and thirdly, where we are dealing with smaller subgroups (particularly applicable for many of the ethnic groups) and where the percentage differences between subgroups are smaller, it would be helpful to assess whether such differences could have occurred randomly, by chance, rather than being reflective of underlying differences between the two groups. Since the ethnic groups are small, it is necessary to collapse these subgroups into broader categories in order to improve the precision of the percentage differences.

Given these limitations, it is necessary to analyse the data more formally with the use of an appropriate statistical model. In this case, a multilevel regression model would be most appropriate, given the clustered nature of the data (students clustered within schools), and also because we are considering the relationships of several independent variables (gender, ethnicity and socioeconomic background) on the allocation of students to maths sets. Moreover, given that the dependent variable in this case is ordinal, with three categories (misallocated upwards, correctly allocated or misallocated downwards), the most parsimonious model would be an ordered logistic multilevel regression model. Such a model was fitted to the data and the results are reported fully in Table 8. As can be seen, four independent variables were included in the model: male (coded '1' for male students and '0' for female students); FSM eligibility (coded '1' for those eligible for FSM and '0' otherwise)—selected in preference to socioeconomic background due to the much higher proportion of missing data for the latter; and two dummy variables representing Asian students (n = 527; drawing together Bangladeshi, Chinese, Indian, Pakistani and Asian Other students) and Black students (n = 373; drawing together Black African, Black Caribbean and Black Other students), with White students (n = 4385) representing the reference category.

Table 8. Ordered logistic multilevel regression model predicting student misallocation to maths sets[a]

Independent variables	Odds ratios[b]	Standard error	Sig.
Boy[c]	1.387	0.088	<0.0005
FSM[d]	0.925	0.069	0.291
Asian[e]	0.773	0.082	0.015
Black[e]	0.505	0.064	<0.0005
Cutpoint 1	−1.649	0.056	
Cutpoint 2	1.841	0.058	
Variance (school)	4.63×10^{-32}	9.78×10^{-17}	
No. schools	38		
No. students	4609		
Model fit (−2LL)	7525.4		

[a]Dependent variable has three categories: misallocated to higher set (coded '1'); allocated correctly (coded '0'); and misallocated to lower set (coded '−1').
[b]These estimated odds ratios should be treated with caution as there is evidence that the proportionality of odds varied significantly across the differing response categories ($p < 0.0005$, chi-square = 38.32, d.f. = 4).
[c]Coded: 1 = boy, 0 = girl.
[d]Coded: 1 = eligible for FSM at some time; 0 = never eligible.
[e]Reference category = White students.

The odds ratios reported in Table 8 would appear to suggest that, overall, male students do tend to fare better than female students whilst, conversely, Asian and Black students tend to fare less well than their White counterparts. Interestingly, the findings would also appear to suggest that FSM eligibility has no notable impact on set allocation. However, and as also noted, these odds ratios need to be treated with caution as there is evidence that one of the key assumptions underpinning the model (the parallel lines assumption) has been violated (approximate likelihood-ratio test of proportionality of odds across response categories: $p < 0.0005$, chi-square = 38.32, d.f. = 4). Given this, and to ensure that comparisons between subgroups are estimated accurately, there is a need to fit two separate multilevel binary logistic regression models instead, representing the two different cutpoints. For these two models, two dependent variables were thus created: Model 1, where the dependent variable categorised students as either being misallocated to a lower set (coded '1') or allocated correctly or misallocated to a higher set (coded '0'); and Model 2, where students were categorised as either being misallocated to a higher set (coded '1') or allocated correctly or misallocated to a lower set (coded '0'). For both models, the same independent variables were included as previously. Both models are reported fully in Table 9.

As can be seen, FSM eligibility was found not to have a statistically significant influence on the odds of a student being misallocated to a higher set, but there was evidence that the odds of being misallocated to a lower set were 1.22 higher for those eligible for FSM compared to those not eligible. There is also evidence that gender has an impact on a student's odds of being misallocated. Thus, it can be seen that the odds of male students being misallocated to a higher set were 1.32 times higher than for female students. Similarly, the odds of male students being misallocated to a lower set were 0.65 times lower than for female students.

Table 9. Binary logistic multilevel regression models predicting student misallocation to maths sets

Independent variables	Model 1[a] (misallocated to lower sets)			Model 2[b] (misallocated to higher sets)		
	Odds ratio	Standard error	Sig.	Odds ratio	Standard error	Sig.
Boy[c]	0.653	0.058	<0.0005	1.318	0.119	0.002
FSM[d]	1.218	0.122	0.050	0.951	0.100	0.632
Asian[e]	1.653	0.256	0.001	0.580	0.100	0.002
Black[e]	2.434	0.386	<0.0005	0.478	0.094	<0.0005
Constant	0.150	0.023	<0.0005	0.123	0.025	<0.0005
Variance (school)	0.624	0.182		1.258	0.383	
Estimated ICC[f]			27.0% (27.3% for null model)			
No. schools	38			38		
No. students	4609			4609		
Model fit (−2LL)	3684.5			3585.9		

[a]Dependent variable coded: 1 (misallocated to lower set); 0 (allocated correctly or misallocated to higher set).
[b]Dependent variable coded: 1 (misallocated to higher set); 0 (allocated correctly or misallocated to lower set).
[c]Coded: 1 = boy, 0 = girl.
[d]Coded: 1 = eligible for FSM at some time; 0 = never eligible.
[e]Reference category = White students.
[f]The estimated intra-class correlations (ICCs) are based on fitting a third model that included the same four independent variables but that used a combined dependent variable coded as: '1' for students who were misallocated (either to a higher or a lower set) and '0' for those allocated correctly. To allow for direct comparisons, the ICC estimated for the null model is based on a sample where cases were deleted if there were any missing data in relation to the dependent or independent variables in the associated full model. The ICCs were estimated from the school-level variance ('var') using: var/(var + $\pi^2/3$).

There is also clear evidence of patterns of misallocation in relation to ethnicity. Thus, the odds of Black students being misallocated to a lower set were 2.43 times higher than for White students. Similarly, the odds of Asian students being misallocated downwards were 1.65 times higher than for their White counterparts. Conversely, the odds of Black students and Asian students being misallocated to a higher set were 0.48 and 0.58 times lower than for White students, respectively. These findings in relation to ethnicity need to be treated with some caution, as the two categories of Asian and Black both combine a number of different ethnic groups whose particular chances of being misallocated may vary. The descriptive statistics summarised earlier (Table 6) would indeed suggest that such differences may exist. However, further research would be required, with larger subsamples, to explore these differences further.

One final point to note is the role that schools play in determining the level of misallocation. It can be seen from Table 9 that it is estimated that approximately 27.3% of the variation in the tendency for students to be misallocated can be attributed to variations in school-level factors (what is commonly termed 'intra-class correlation'). This is a notable proportion, and would suggest that there is significant potential for schools to have an impact on reducing levels of misallocation. However, and beyond this, it is also interesting to note that once gender, socioeconomic background and ethnicity have been controlled for, the proportion of this variance that remains associated with schools only drops marginally (to 27.0%). This would indicate that there is no evidence to suggest that the association between gender and ethnicity and the tendency to misallocate students is related to variations between schools in relation to the composition of their student intake. This suggests that the tendency to misallocate Black and Asian students may not be associated with the proportions of Black and Asian students in particular schools, although further research would be needed to investigate this fully.

Overall, and in relation to these two models summarised in Table 8, there are two limitations that are important to note. First, there is a high proportion of missing data (with both models based on a sample of only 4609 students from a total sample of 9301). Second, the measure of FSM eligibility is widely recognised as a relatively poor proxy measure for socioeconomic background. With these limitations in mind, further sensitivity analyses were undertaken by comparing the estimated odds ratio for both models with two separate models in turn: one that included a further dummy variable for ethnicity which included all other students (whether they were coded as 'Mixed' or where their ethnic data were missing, resulting in a much higher sample size of n = 8209); and a second model that used socioeconomic background instead of FSM eligibility (including dummy variables for those categorised as being from families in a 'higher' or 'intermediate' socioeconomic background, with those from 'lower' backgrounds acting as the reference category). The findings of these sensitivity analyses are summarised in Table 10.

It can be seen that the estimated odds ratios for male, Asian and Black students remain fairly consistent across models, suggesting that the findings reported in Table 9 can be regarded as sufficiently robust. The one exception relates to the finding reported in Table 9 that those eligible for FSM are more likely to be misallocated to lower sets. This finding became non-significant when the additional ethnicity dummy variable was included (p = 0.070) and, moreover, it was not replicated at all when using the alternative measure for socioeconomic background, where both dummy variables for higher SES (socioeconomic status) and intermediate SES were also found to be non-significant (p = 0.546 and 0.501, respectively). Given this, and given the inconsistency in the main findings reported in Table 9 (which appeared to suggest some evidence that those eligible for FSM were more likely to be allocated to lower sets but no evidence that, conversely, they were less likely to be allocated to higher sets), it would be reasonable to conclude that there is insufficient evidence generated through this present analysis to conclude that socioeconomic background is associated with the misallocation of students to maths sets.

With all this in mind, and for ease of reference, the key findings from these analyses are summarised in Table 11.

Borderline students

The above analyses did not include the small proportion of students (5.8%, n = 508) classified by the research team as borderline for the purpose of the present analysis. This category was not one used by the school but was based solely on the students' KS2 results and derived from the present authors' attempts to create a counterfactual condition where students were allocated to sets based solely on these results. As

Table 10. Sensitivity analyses (estimated parameters with their associated standard errors in parentheses)

	Model 1[a] (misallocated to lower sets)			Model 2[b] (misallocated to higher sets)		
Independent variables	Model 1a	Model 1b	Model 1c	Model 2a	Model 2b	Model 2c
Boy[c]	0.653 (0.058)	0.687 (0.045)	0.625 (0.055)	1.318 (0.119)	1.186 (0.075)	1.376 (0.125)
FSM[d]	1.218 (0.122)	1.144 (0.082)		0.951 (0.100)	1.089 (0.078)	
Higher SES[e]			0.956 (0.114)			0.710 (0.087)
Intermediate SES[e]			0.976 (0.121)			0.886 (0.110)
Asian[f]	1.653 (0.256)	1.452 (0.213)	1.708 (0.273)	0.580 (0.100)	0.669 (0.108)	0.566 (0.105)
Black[f]	2.434 (0.386)	2.175 (0.324)	2.602 (0.425)	0.478 (0.094)	0.529 (0.099)	0.526 (0.111)
Mixed[f]		1.068 (0.158)	1.311 (0.206)		1.060 (0.149)	0.880 (0.142)
Missing[f,g]		1.031 (0.099)			1.115 (0.104)	
Other[f]		1.607 (0.357)	2.067 (0.473)		1.067 (0.244)	0.998 (0.251)
All Other[f]		0.157 (0.020)	0.170 (0.028)		0.128 (0.019)	0.138 (0.030)
Constant	0.150 (0.023)	0.541 (0.135)	0.537 (0.159)	0.123 (0.025)	1.186 (0.075)	1.376 (0.125)
Variance	0.624 (0.182)	0.687 (0.045)	0.625 (0.055)	1.258 (0.383)	0.790 (0.206)	1.231 (0.384)
No. schools	38	46	38	38	46	38
No. students	4609	8234	4630	4609	8234	4630
−2LL	3684.5	6616.0	3756.5	3585.9	6712.7	3561.3

[a]Dependent variable coded: 1 (misallocated to lower set); 0 (allocated correctly or misallocated to higher set).
[b]Dependent variable coded: 1 (misallocated to higher set); 0 (allocated correctly or misallocated to lower set).
[c]Coded: 1 = boy, 0 = girl.
[d]Coded: 1 = eligible for FSM at some time; 0 = never eligible.
[e]Reference category = lower SES.
[f]Reference category = White students.
[g]Included in Models 1c and 2c but this variable was subsequently omitted due to collinearity.

Table 11. Summary of the findings in relation to misallocation of students to maths sets by gender, ethnicity and socioeconomic background[a]

Nature of misallocation	Difference in odds of being misallocated
To lower set in maths	• Black students **2.43** times more likely than White students. • Asian students **1.65** times more likely than White students. • Female students **1.53**[b] times more likely than males. • Insufficient evidence of differences between students from differing socioeconomic backgrounds.
To higher set in maths	• White students **2.09**[b] times more likely than Black students. • White students **1.72**[b] times more likely than Asian students. • Male students **1.32** times more likely than females. • Insufficient evidence of differences between students from differing socioeconomic backgrounds.

[a]All figures based on odds ratios estimated through binary logistic regression models reported in Table 9.
[b]To aid in interpretation of the findings, these odds ratios represent the inverse of those reported in Table 9.

explained earlier and illustrated in Table 1, allocating students solely on their KS2 scores creates difficulties at the boundaries between set levels, where there will be a small number of students with exactly the same KS2 results who, given the constraints of school class size, cannot all fit into one set and could thus be 'legitimately' allocated to the respective higher or lower set above and below the boundary on which they sit. One key remaining question to test in relation to this group is whether they were, indeed, equally likely to be allocated to either the higher or the lower set or, conversely, whether any patterns existed in the schools' allocation practices that differentially impacted on students in relation to their gender, social class and ethnicity. Whilst this is a much smaller subsample, it is of particular interest as students are completely matched in relation to their KS2 results.

Overall, the actual allocations of the 508 borderline students are summarised in Table 12. As can be seen, similar proportions of borderline students were allocated upwards and downwards, and most (85.8%) were allocated to the immediate set

Table 12. The allocation of students classified as 'borderline'

The number of sets above or below the border that the student was allocated to		No.	%
+4		2	0.4
+3		10	2.0
+2		20	3.9
+1		211	41.5
	−1	225	44.3
	−2	28	5.5
	−3	9	1.8
	−4	2	0.4
	−5	1	0.2
Total		508	100.0

above or below their position, as determined by their KS2 results. Much smaller proportions of borderline students were allocated to sets beyond those that were immediately above or below their position. Moreover, it is difficult to compare such students directly, as some will have been attending schools with just two or three sets and others will have been attending schools with up to eight sets (see Table 3). Thus, the potential for some students to be allocated beyond the sets immediately above and below their position, as determined by their KS2 results, will have varied. Moreover, and in relation to the analysis to follow, these problems of comparison become exacerbated due to the small sizes of some of the subsamples being compared. It is with this in mind that borderline students have simply been organised into two categories: those allocated to a higher set and those allocated to a lower set. Similarly, and in relation to ethnicity, the three main categories of White, Asian and Black will be compared in the subsequent analysis as the numbers of students in most of the more defined ethnic group categories were too small (typically less than 10).

The proportions of borderline students allocated upwards and downwards with regard to their position as determined by the KS2 results are summarised in Table 13. It can be seen that whilst there is some variation across the subgroups of students, it is not clear whether these are likely to reflect real underlying patterns or, given some of the low subsample sizes, could have occurred by chance. To test this, two multilevel binary logistic regression models were fitted with either household socioeconomic background (Model 1) or FSM eligibility (Model 2) included as proxy measures of social class. In both cases, the dependent variable was binary and coded as either '1' (for those allocated to a higher set) or '0' (for those allocated to a lower set). Both models are summarised in Table 14 and, as can be seen, none of the differences between the subgroups was found to be statistically significant. There is thus no evidence, from this current analysis, that there are systematic differences in the

Table 13. The allocation of borderline students by schools to maths sets by gender, social class and ethnicity (%)

	Allocated upwards	Allocated downwards	Total (No.)
All	47.8	52.2	508
Gender			
Boy	48.7	51.3	265
Girl	47.1	52.9	242
Household socioeconomic background			
Lower	47.9	52.1	48
Intermediate	38.3	61.7	94
Higher	54.8	45.2	135
Ever eligible for free school meals			
No	49.3	50.7	345
Yes	45.3	54.7	159
Ethnicity			
Asian	48.5	51.5	33
Black	63.6	36.4	22
White	46.0	54.0	237

Table 14. Binary logistic multilevel regression models predicting borderline students' likelihood of being allocated upwards[a]

	Model 1			Model 2		
Independent variables	Odds ratio	Standard error	Sig.	Odds ratio	Standard error	Sig.
Male[b]	1.552	0.422	0.106	1.060	0.197	0.755
Higher SES[c]	1.069	0.402	0.858			
Intermediate SES[c]	0.618	0.242	0.219			
Eligible for FSM[d]				0.849	0.174	0.423
Asian[e]	1.420	0.670	0.457	1.097	0.425	0.810
Black[e]	2.577	1.552	0.116	2.132	1.008	0.109
Constant	0.739	0.265	0.400	0.887	0.147	0.470
Variance (school)	0.396	0.285		0.076	1.498	
No. schools	34			43		
No. students	277			503		
Model fit (−2LL)	367.0			691.8		

[a]Dependent variable for both models is coded: 0 (allocated to lower set); 1 (allocated to higher set).
[b]Reference category = female student.
[c]Reference category = lower SES.
[d]Reference category = never eligible for FSM.
[e]Reference category = White students.

specific allocations of the small group of borderline students by gender, social class or ethnicity.

Discussion

Overall, this is the first study to provide robust and relatively large-scale evidence regarding the set allocation practices of secondary schools. Based on the counterfactual case whereby students are allocated solely on the basis of their prior attainment, this study has found that significant proportions of students have been misallocated, representing nearly a third (31.2%) of all students. This, in turn, is in line with the findings of previous studies where some teachers have reported that their setting decisions are based on their assessments of the students' 'ability' and wider behaviour and attitudes as well as on their actual prior attainment (Muijs & Dunne, 2010). The present study has not only provided confirmatory evidence of such practices, but also the extent of misallocation. The findings have also provided some evidence of the extent to which differences between schools have an impact on levels of misallocation. In particular, just over a quarter (27.3%) of the variation in levels of student misallocations to maths sets was found to be associated with school-level factors.

Beyond this, the present study has, for the first time, generated evidence on whether the setting practices of secondary schools in Year 7 exacerbate existing educational inequalities that are evident amongst the students as they transfer from primary school. In this regard, in relation to socioeconomic background and ethnicity, clear patterns were evident in relation to the attainment of students at KS2, reflecting longstanding trends (Archer & Francis, 2007), with those from higher socioeconomic backgrounds tending to achieve higher scores, and Chinese and Indian-heritage students outperforming the White majority, while those from Black minority ethnic groups underperform the White majority. Moreover, these patterns were also clearly reflected in the patterns of set allocation in maths and English by socioeconomic background, with those from higher socioeconomic backgrounds tending to be allocated to high sets in maths, and the same being true for ethnicity (with Chinese and Indian-heritage students more likely to be placed in high sets, and so on). The most notable finding from the present study, however, is that the setting practices of schools did not tend to exacerbate these differences with regard to socioeconomic background. However, and within this, it was also found that those from lower socioeconomic backgrounds were both more likely to be misallocated downwards and also more likely to be misallocated upwards. The higher proportions of those misallocated downwards fits existing research findings that stereotyping and labelling lead working-class students to be misallocated downwards (e.g. Jackson, 1964). However, and in contrast, the simultaneous tendency to be slightly more likely also to be misallocated upwards is previously undocumented, and intriguing. It is possible that there could be some application of a 'deserving scholarship' impetus operating here—further research would be required to explore this hypothesis.

This overall finding that set allocation does not tend to exacerbate differences according to socioeconomic background is an important point. However, it does require that two caveats are made. First, the finding does not suggest that socioeconomic background has no bearing on a student's set allocation. On the contrary, it is clear from the findings above that inequalities with regard to socioeconomic background and educational attainment are evident at the end of KS2 in the test scores reported (see Table 5). It is this influence of socioeconomic background, already internalised by students before the end of primary school, that is then carried forward into secondary school and reflected in their set allocations. The key point from the present study is simply that no evidence has been found that secondary schools exacerbate these patterns of inequality further through their setting processes. The second caveat to this main finding is that this does not imply that setting does not then have a subsequent impact on these inequalities with regard to socioeconomic background. Indeed, and as summarised earlier, there is a significant body of work to suggest that setting of students is likely to exacerbate existing attainment gaps over time. Given that students from low socioeconomic backgrounds are over-represented in low sets, any inequality of resources or pedagogy will disproportionately impact this group of students, with implications for the reproduction of social inequality (see e.g. Francis *et al.*, 2017b; Mazenod *et al.*, 2019).

On the contrary, the findings in relation to gender and ethnicity both suggest that schools *do* have a role to play in exacerbating existing inequalities, specifically through their setting practices. With regard to gender, there is evidence of differences between boys and girls as they enter secondary school, with boys tending to attain slightly higher KS2 scores in maths, on average, compared to girls. Given this, if the setting of students were based solely on prior attainment, then one would expect to find some gender differences in set allocations in maths. However, the findings reported above provide clear evidence that, even after prior attainment is taken into account, boys are still more likely to be misallocated to higher sets in maths compared to girls and, conversely, girls are more likely to be misallocated to lower sets in maths compared to boys. Given the ongoing concerns about the under-representation of girls in STEM (science, technology, engineering and maths) subjects and careers, within which maths often plays a fundamental part, this finding is especially alarming. It also stands as a corrective to pervasive concerns that boys are always disadvantaged by school practices in comparison to girls.[2]

Finally, the most notable effects of schools in relation to exacerbating existing educational inequalities through set allocation practices are evident in relation to ethnicity, with Black and Asian students being more likely to be misallocated to lower sets in maths than White students. The differences appear to be most stark for Black students, who are over 2.4 times more likely to be misallocated to lower maths sets than their prior attainment would warrant compared to White students. Whilst these findings tend to confirm a large body of (mainly qualitative) evidence that exists to date with regard to the schooling of Black students, this is the first reliable estimate of the extent of these practices, and a demonstration of how school decisions in relation to setting practices tend to exacerbate inequalities. However, three caveats are necessary with regard to this finding. First, the level of missing data about students' ethnicity is high, although this does not appear to significantly affect the results. Second, as is clear from the raw data presented in Table 7, this broad finding may obscure some complexity across different ethnic subgroups. Indeed, one interesting issue relates to Black African students. A relatively high proportion of these students are misallocated downwards (29.3% compared to 11.6% misallocated upwards). Yet, in general, Black African students achieve much better GCSE outcomes at age 16 than White British pupils, even without any adjustment for the much higher levels of disadvantage they experience (see Strand, 2015). So, the systematic misallocation that we have demonstrated does not seem to impact negatively on this subgroup of students. However, as we have already noted, the number of Black African students is small, and the percentage differences observed in the raw data need to be treated with some caution. The third caveat relates to school size. Our model does not take account of school size. Larger schools are likely to have more sets and thus may tend to have a larger absolute number of misallocated students. Since many larger schools are in urban areas with higher proportions of ethnic minority students, this may be a confounding factor.

Aside from these important findings concerning the extent of misallocation, and the impact or otherwise of pupil identity variables (ethnicity, gender and socioeconomic background) on allocation of pupils to maths sets, the study has also served to identify a practical/methodological problem with setting practice which has not been identified previously. This is the case of 'borderline' pupils. Set group size is often constrained by optimal (or maximum) class size decisions, as well as by timetabling [see Taylor *et al.* (2019) for a discussion of the latter], rather than simply being governed by the number of pupils achieving particular benchmarks of prior attainment. As such, inevitably, there tends to be a group of 'borderline' pupils who have achieved the same results but cannot all fit into the higher set. As such, an arbitrary decision is necessitated about which pupils should go where, automatically precipitating injustice (given that the decision is arbitrary) and opening up the possibility of bias according to pupil characteristics. In relation to the present study, no evidence was found of any systematic bias in allocation practices with respect to gender, socioeconomic background or ethnicity for students identified as 'borderline'. However, further research is required in relation to this subgroup of students, and especially whether schools have identified them as borderline and, if so, how decisions are made regarding these.

Our findings, then, require urgent reflection and action on the part of schools. They lend credence to the argument that, if setting is to be used, it must be applied purely on the basis of prior attainment, to avoid misallocation and the creeping prejudice that our findings suggest applies in misallocated cases. Albeit, given the illustration of how social inequalities are reflected in the prior attainment patterns underpinning 'correct' allocation into set groups, and the literature showing how set allocation precipitates distinctions in access to resources and other outcomes, we argue that the decision to adopt setting should never be taken lightly. Moreover, we suggest that schools should prioritise a review of the issue of 'borderline' students and the practice applied to decision-making in their case, again reflecting on the implications for school organisation and practice, and ways to ensure parity. We hope that in this way our findings can support schools to improve practice and equity.

NOTES

[1] While it is commonplace to refer to 'ability' grouping, this conflates current attainment with the notion of an objective measure of general ability. This article therefore refers to *grouping by attainment* and uses 'ability' in inverted commas throughout to emphasise its contested nature.

[2] However, preliminary analysis of our equivalent data for allocation to sets in English indicates that boys are disadvantaged relative to girls.

References

Archer, L. & Francis, B. (2007) *Understanding ethnic minority achievement* (London, Routledge).

Ball, S. J. (1981) *Beachside Comprehensive: A case-study of secondary schooling* (Cambridge, Cambridge University Press).

Bartholomew, H. (2000) Negotiating identity in the community of the mathematics classroom, paper presented at the *British Educational Research Association Annual Conference*, Cardiff, 7–10 September.

Boaler, J. & Wiliam, D. (2001) Setting, streaming and mixed-ability teaching, in: J. Dillon & M. Maguire (Eds) *Becoming a teacher* (Maidenhead, Open University Press), 173–181.

Bosworth, R. (2013) What sort of school sorts students?, *International Journal of Quantitative Research in Education*, 1(1), 20–38.

Bradbury, A. & Roberts-Holmes, G. (2017) *Grouping in Early Years and Key Stage 1: 'A necessary evil'?* (London, National Education Union).

Burris, C. C. & Welner, K. G. (2005) Closing the achievement gap by detracking, *Phi Delta Kappan*, 86(8), 594–598.

Cahan, S., Linchevski, L., Ygra, N. & Danziger, I. (1996) The cumulative effect of ability grouping on mathematical achievement: A longitudinal perspective, *Studies in Educational Evaluation*, 22 (1), 29–40.

Chambers, T. V. (2009) The "receivement gap": School tracking policies and the fallacy of the "achievement gap", *Journal of Negro Education*, 78(4), 417–431.

Clifton, J. & Cook, W. (2012) *A long division: Closing the attainment gap in England's secondary schools* (London, Institute for Public Policy Research).

Connolly, P. (2004) *Boys and schooling in the early years* (London, RoutledgeFalmer).

Dar, Y. & Resh, N. (1994) Separating and mixing students for learning: Concepts and research, *Pedagogisch Tijdschrift*, 19(2), 109–126.

DfE (2015a) *National curriculum assessments at Key Stage 2 in England, 2015 (revised): Statistical first release* (London, Department for Education).

DfE (2015b) Supporting the attainment of disadvantaged pupils: Articulating success and good practice. Available online at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/473974/DFE-RR411_Supporting_the_attainment_of_disadvantaged_pupils.pdf (accessed 1 December 2015).

DfES (2005) *Higher standards, better schools for all: More choice for parents and pupils* (London, The Stationery Office).

Dunne, M., Humphreys, S., Sebba, J., Dyson, A., Gallannaugh, F. & Muijs, D. (2007) Effective teaching and learning for pupils in low attaining groups. Research Report (London, Department for Children, Schools and Families).

Francis, B., Archer, L., Hodgen, J., Pepper, D., Taylor, B. & Travers, M.-C. (2017a) Exploring the relative lack of impact of research on 'ability grouping' in England: A discourse analytic account, *Cambridge Journal of Education*, 47(1), 1–17.

Francis, B., Connolly, P., Archer, L., Hodgen, J., Mazenod, A., Pepper, D. *et al.* (2017b) Attainment grouping as self-fulfilling prophecy? A mixed methods exploration of self confidence and set level among Year 7 students, *International Journal of Educational Research*, 86, 96–108.

Gamoran, A. & Nystrand, M. (1994) Tracking, instruction and achievement, *International Journal of Educational Research*, 21(2), 217–231.

Gillborn, D. (2008) *Racism and education: Coincidence or conspiracy?* (London, Routledge).

Gillborn, D. & Youdell, D. (2000) *Rationing education: Policy, practice, reform, and equity* (Maidenhead, Open University Press).

Hallam, S. (2012) Streaming and setting in UK primary schools: Evidence from the Millennium Cohort Study, *FORUM: For Promoting 3–19 Comprehensive Education*, 54, 57–64.

Hallam, S. & Parsons, S. (2013a) The incidence and make up of ability grouped sets in the UK primary school, *Research Papers in Education*, 28(4), 393–420.

Hallam, S. & Parsons, S. (2013b) Prevalence of streaming in UK primary schools: Evidence from the Millennium Cohort Study, *British Educational Research Journal*, 39(3), 514–544.

Hallinan, M. T. (1996) Track mobility in secondary school, *Social Forces*, 74(3), 983–1002.

Hallinan, M. T. & Sorenson, A. B. (1987) Ability grouping and sex differences in mathematics achievement, *Sociology of Education*, 60(2), 63–72.

Higgins, S., Katsipataki, M., Coleman, R., Henderson, P., Major, L., Coe, R. *et al*. (2015) *The Sutton Trust–Education Endowment Foundation teaching and learning toolkit* (London, Education Endowment Foundation).

Hodgen, J. (2011) Setting, streaming and mixed ability teaching, in: J. Dillon & M. Maguire (Eds) *Becoming a teacher: Issues in secondary education* (Maidenhead, Open University Press), 210–221.

Ireson, J. & Hallam, S. (2001) *Ability grouping in education* (London, Paul Chapman).

Ireson, J., Hallam, S. & Hurley, C. (2005) What are the effects of ability grouping on GCSE attainment?, *British Educational Research Journal*, 31(4), 443–458.

Jackson, B. (1964) *Streaming: An education system in miniature* (London, Routledge & Kegan Paul).

Jerrim, J., Greany, T. & Perera, N. (2018) *Educational disadvantage: How does England compare?* (London, Education Policy Institute).

Lupton, R. (2014) Raising teachers' voice on achievement in urban schools in England: An afterword, *The Urban Review*, 46(5), 919–923.

Marks, R. (2012) Discourses of ability and primary school mathematics: Production, reproduction and transformation (London, King's College London).

Marks, R. (2013) "The blue table means you don't have a clue": The persistence of fixed-ability thinking and practices in primary mathematics in English schools, *FORUM: For Promoting 3–19 Comprehensive Education*, 55(1), 31–44.

Mazenod, A., Francis, B., Archer, L., Hodgen, J., Taylor, B., Tereshchenko, A. *et al.* (2019) Nurturing learning or encouraging dependency? Teacher constructions of students in lower attainment groups in English secondary schools, *Cambridge Journal of Education*, 49, 53–68.

Mitchell, J. C. (1984) Typicality and the case study, in: R. F. Ellen (Ed.) *Ethnographic research: A guide to general conduct* (London, Academic Press), 238–241.

Modica, M. (2015) *Race among friends: Exploring race at a suburban school* (New Brunswick, NJ, Rutgers University Press).

Moller, S. & Stearns, E. (2012) Tracking success: High school curricula and labor market outcomes by race and gender, *Urban Education*, 47(6), 1025–1054.

Muijs, D. & Dunne, M. (2010) Setting by ability – or is it? A quantitative study of determinants of set placement in English secondary schools, *Educational Research*, 52(4), 391–407.

Nomi, T. (2009) The effects of within-class ability grouping on academic achievement in early elementary years, *Journal of Research on Educational Effectiveness*, 3(1), 56–92.

Oakes, J. (1985) *How schools structure inequality* (New Haven, CT, Yale University Press).

OECD (2013) Selecting and grouping students, in: *PISA 2012 results: What makes schools successful?* (vol. IV: *Resources, policies and practices*) (Paris, OECD Publishing), 71–92.

Parsons, S. & Hallam, S. (2014) The impact of streaming on attainment at age seven: Evidence from the Millennium Cohort Study, *Oxford Review of Education*, 40(5), 567–589.

Rui, N. (2009) Four decades of research on the effects of detracking reform: Where do we stand? A systematic review of the evidence, *Journal of Evidence-Based Medicine*, 2(3), 164–183.

Slavin, R. E. (1990) Achievement effects of ability grouping in secondary schools: A best-evidence synthesis, *Review of Educational Research*, 60(3), 471–499.

Steenbergen-Hu, S., Makel, M. C. & Olszewski-Kubilius, P. (2016) What one hundred years of research says about the effects of ability grouping and acceleration on K-12 students' academic achievement: Findings of two second-order meta-analyses, *Review of Educational Research*, 86 (4), 849–899.

Strand, S. (2012) The White British–Black Caribbean achievement gap: Tests, tiers and teacher expectations, *British Educational Research Journal*, 38(1), 75–101.

Strand, S. (2015) Ethnicity, deprivation and educational achievement at age 16 in England: Trends over time. Research Report (London, Department for Education).

Taylor, B., Francis, B., Archer, L., Hodgen, J., Pepper, D., Tereshchenko, A. *et al.* (2017) Factors deterring schools from mixed attainment teaching practice, *Pedagogy, Culture & Society*, 25(3), 327–345.

Taylor, B., Francis, B., Craig, N., Archer, L., Hodgen, J., Mazenod, A. *et al.* (2019) Why is it difficult for schools to establish equitable practices in allocating students to attainment 'sets'?, *British Journal of Educational Studies*, 67, 5–24.

Treadaway, M. (2013) An analysis of Key Stage 2 reliability and validity. Research Report (Cowbridge, FFT Education Ltd).

Van de Gaer, E., Pustjens, H., Van Damme, J. & De Munter, A. (2006) Tracking and the effects of school-related attitudes on the language achievement of boys and girls, *British Journal of Sociology of Education*, 27(3), 293–309.

Waldfogel, J. & Washbrook, E. (2010) *Low income and early cognitive development in the UK: A report for the Sutton Trust* (London, Sutton Trust).

Wiliam, D. & Bartholomew, H. (2004) It's not which school but which set you're in that matters: The influence of ability grouping practices on student progress in mathematics, *British Educational Research Journal*, 30(2), 279–293.

Wright, C. (1986) School processes: An ethnographic study, in: J. Eggleston, D. Dunn & M. Anjali (Eds) *Education for some: The educational & vocational experiences of 15–18 year old members of minority ethnic groups* (Stoke-on-Trent, Trentham), 127–179.