Why is it difficult for schools to establish equitable practices in allocating students to attainment 'sets'?

Published in:
British Journal of Educational Studies

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Why is it difficult for schools to establish equitable practices in allocating students to attainment 'sets'?

Becky Taylor, Becky Francis, Nicole Craig, Louise Archer, Jeremy Hodgen, Anna Mazenod, Antonina Tereshchenko & David Pepper

Published online: 16 Jan 2018.

Submit your article to this journal ⧉

Article views: 238

View related articles ⧉

View Crossmark data ⧉

WHY IS IT DIFFICULT FOR SCHOOLS TO ESTABLISH EQUITABLE PRACTICES IN ALLOCATING STUDENTS TO ATTAINMENT 'SETS'?

by BECKY TAYLOR and BECKY FRANCIS , *UCL Institute of Education*, NICOLE CRAIG, *Queen's University Belfast*, LOUISE ARCHER, JEREMY HODGEN, ANNA MAZENOD and ANTONINA TERESHCHENKO, *UCL Institute of Education* and DAVID PEPPER, *King's College London*

ABSTRACT: *Research has consistently shown 'ability' grouping (tracking) to be prey to poor practice, and to perpetuate inequity. A feature of these problems is inequitable and inaccurate practice in allocation to groups or 'tracks'. Yet little research has examined whether such practices might be improved. Here, we examine survey and interview findings from a large-scale intervention study of grouping practices in 126 English secondary schools. We find that when schools are encouraged to allocate students and move them between groups according to equitable principles by participation in a 'best practice' intervention, there is some increased equity of practice (i.e. a reduction in non-attainment factors used in allocation). However, the majority of schools continue to use subjective and potentially biased information to group students. Furthermore, some schools that claim to be using attainment setting appear to be using the inequitable practice of streaming. Our findings show that improvements in equity are constrained by operational and strategic factors, including timetabling, finance, and teachers' values and beliefs relating to student ability and progression. We suggest strategies for encouraging schools to change their grouping practices, drawing on approaches for working with complex organisations.*

Keywords: attainment grouping, setting, English secondary schools, equity

1. INTRODUCTION

Attainment grouping highlights issues of inequity in English secondary schools. There has been substantial investigation of the impact of attainment grouping on achievement and on teachers' and students' attitudes towards it. Outcomes are unequal, with students in high-attainment sets making more progress, achieving higher value credentials and having higher self-confidence than their peers in low-attainment sets (Ireson and Hallam, 2001). Opportunities are unequal, as it has been demonstrated that students in lower-attainment sets are provided with an impoverished curriculum and pedagogy relative to their higher-attaining peers (Hallam and Ireson, 2005).

In this article we address the question of equity, presenting evidence that processes of allocation of students to attainment groups are opaque and inequitable. The factors that influence and constrain practice, perpetuating inequitable attainment grouping, remain under-explored. Without an understanding of how and why inequitable grouping practices endure, educators and policy-makers will struggle to combat their negative effects. We seek to address this gap in this article.

The inequity of setting

The construct of equality in education focuses on equal sharing: equality of outcomes, or, in the case of the more liberal concept of equality of opportunity, equality of starting points (Coleman, 1975). 'Equity', by contrast, encapsulates considerations of fairness and even-handedness (OECD, 2012). Both can incorporate redistributive principles in order to compensate for disadvantage. Setting, a form of attainment grouping commonly practiced in English schools[1], is frequently critiqued from the perspective of equality and equity. Schools report that they use setting because it enables teachers to challenge high-attaining learners and support those who are struggling (Slavin, 1990). However, attainment grouping has been repeatedly demonstrated to be associated with inequitable outcomes in both achievement and self-confidence (Francis *et al.*, 2016, 2017; Higgins *et al.*, 2015; Kutnick *et al.*, 2005; Muijs and Dunne, 2010). Furthermore, low-attainment groups disproportionately represent certain disadvantaged groups, such as students from working class backgrounds and boys of African Caribbean heritage (Dunne *et al.*, 2007; Gillborn and Youdell, 2000), leading to what Francis *et al.* (2016) characterise as 'double disadvantage'.

Setting is a practice where students are separated into different 'ability'[2] groups for each subject, on the basis of attainment in that subject (Ireson and Hallam, 2001). Another strategy is streaming, where students are segregated according to their perceived general 'ability' (Hallam, 2012) and taught within these streams for all subjects. Some schools segregate further by setting by subject-specific attainment within general 'ability' streams. By contrast, some schools group students such that there is a broad range of attainment in mixed-attainment groups (Taylor *et al.*, 2016). It has been found that lower-attaining students do better in mixed-attainment classes where high-attaining peers are also present (Linchevski and Kutscher, 1998). Setting is the most prevalent of these grouping practices, with recent figures suggesting 95% of 15 year olds are taught mathematics in sets (OECD, 2013).

Nevertheless, it might be that the inequitable outcomes of attainment grouping (and inequitable representation of different social groups in different attainment groups) could be improved. And certainly, if practised by state schools, operational applications of attainment grouping should be subject to transparent equality of opportunity.

Setting: allocation and movement of students

It is generally understood that setting is based on the prior attainment of students, however, research demonstrates that this is not so. Jackson (1964) reported that in addition to attainment data, schools used subjective sources such as the 'experienced judgement' of the head teacher and 'teacher recommendations' to allocate students to streams, with over one-third of schools using no 'objective test' at all. Unsurprisingly, Jackson found that stream placement related strongly to social class in his case study school. Hence, in spite of the understanding that these groupings are based on prior attainment (often seen as a proxy for a more fixed notion of 'ability'), Jackson showed that a range of subjective and discriminatory practices were actually applied, resulting in social inequality in grouping allocation.

More recent research has also found that both attainment data and subjective judgements are used by English secondary schools to decide students' set allocation (Ireson and Hallam, 2001; Muijs and Dunne, 2010). Ireson and Hallam (2001) found that for allocating students to English, mathematics and science sets, schools reported using one or more of Key Stage 2 (KS2) test results, standardised tests (e.g. CATs or MidYIS[3]), internal teacher assessments, tests or examinations, teacher judgements and transfer information from primary schools. They, and other researchers, have also found factors such as student gender to be applied in some schools (see e.g. Charlton *et al.*, 2007).

While attainment predicts set placement to a limited extent (Muijs and Dunne, 2010), social inequality in grouping allocation persists. Students from lower socioeconomic backgrounds are consistently found to be concentrated in lower sets and streams (Ball, 1981; Bosworth, 2013; Muijs and Dunne, 2010; Taylor and Sloan, 2016). Black students are more likely to be allocated to lower sets/tracks and White students and those from some Asian backgrounds to higher sets/tracks (Ball, 1981; Gillborn and Youdell, 2000; Hallinan, 1996; Modica, 2015; Moller and Stearns, 2012; Muijs and Dunne, 2010; Shaw *et al.*, 2016; Taylor and Sloan, 2016). There is some evidence that boys are more likely to be placed in lower streams (Campbell, 2014; Hallam and Parsons, 2013; Jackson, 1964; Van De Gaer *et al.*, 2006). However, other studies have failed to find a significant relationship between gender and set placement (Muijs and Dunne, 2010), have found the opposite pattern (Moller and Stearns, 2012), or subject-specific patterns (Taylor and Sloan, 2016).

Also significant is the movement of students between sets after their initial allocation on starting secondary school. In the United States, there has been debate as to the extent of movement between tracks (Hallinan, 1996). However, more recent research in English schools suggests that students tend to stay in the same group, regardless of progress (Dunne *et al.*, 2011, 2007). Furthermore teachers tend to overestimate the amount of movement between groups (Hallam and Ireson, 2005). Static setting serves to perpetuate inequality, because students initially misallocated cannot earn their passage into higher groups.

Furthermore, unequal progression of students in different sets due to differential pace and curriculum coverage can prevent movement (Boaler *et al.*, 2000).

The result is that setting, as currently practised in England, involves the inequitable allocation of students to sets and lack of movement once initial allocations have been made. This combines with differing progress made by students in higher- and lower-attaining sets, to create inequitable, segregated pathways towards academic success or failure (Gillborn and Youdell, 2000).

Towards best practice in setting

We have distilled in detail elsewhere the factors identified by the research literature to create the deleterious effects of setting by attainment on young people in low-attainment groups (Francis *et al.*, 2016). We have also presented our model of 'Best Practice in Setting' (ibid.), which intends to militate against these factors by requiring schools to organise and teach their attainment sets in specific stipulated ways (drawing on the findings of the existing literature). Our hypothesis is that this model could result in a more equitable approach to setting, which mitigates group differences in attainment and self-confidence measures. The Best Practice in Setting intervention is currently being applied in 64 secondary schools, and evaluated through randomised controlled trial (EEF, n.d.).

Included in our Best Practice in Setting intervention are the following requirements relating to allocation of students to, and movement of students between, sets:

(1) Transparent allocation of students to sets, based purely on prior attainment (KS2 test results).
(2) Regular opportunities for students to move between sets on the basis of internal assessment results [4].

These requirements are intended to remediate trends identified in the research literature, which prior studies have suggested contribute to inequity. They attempt to support closer integrity of setting practice with students' prior attainment, reducing subjective bias. In this way it is anticipated that setting practice might be both more transparently and equitably applied, and that this more transparent application would also reduce the extent to which certain social groups are either advantaged or disadvantaged through subsequent setting experience. We proposed that allocating students to sets solely by attainment at KS2 would ensure that, while KS2 test results reflect patterns of inequality already established through early years and primary education including a widening socio-economic gap (DfE, 2015) and are likely to be affected by factors including past teacher bias and ability constructions (Hamilton and O'Hara, 2011; Hart *et al.*, 2004; Oakes, 1985), *current* teacher biases against students from disadvantaged backgrounds would not influence students' initial set position (Campbell, 2015; Timmermans *et al.*, 2015). Focusing on subject-specific attainment allows for more diversity in set

membership and so minimises labelling and self-fulfilling prophecy. Furthermore, assuming that a student is not in lower-attaining sets for all subjects, it reduces the impact of any more impoverished resourcing of low-attainment sets. Additionally, a focus on KS2 attainment means that our setting model depends on a common objective measure, rather than the contentious notion of perceived 'ability'. Moving students between sets regularly, and only on the basis of assessment results, seeks to ensure that students continue to be placed in sets purely on the basis of attainment. Through normalising set movement we also seek to loosen the hold of 'fixed ability' thinking (Dweck, 1986)[5] and emphasise the role of effort in achievement and so to de-stigmatise membership of lower sets. We propose that this could improve impact of setting on self-confidence.

2. RESEARCH QUESTIONS

In this article we seek to address the gap in the current literature, which leaves the factors influencing and constraining school grouping practices unexplored. Through investigating this area, we hope to provide evidence that will help educators and policy-makers increase the equity of grouping in schools. Drawing on extensive data from surveys and interviews with teachers, we address the following questions:

(1) What has been the effect of the Best Practice in Setting intervention on set allocation practices?
(2) How do the techniques and approaches applied by schools in setting secondary school pupils facilitate and/or constrain equitable practice?
(3) What factors influence and constrain equitable set allocation practices?

3. METHODS

The data in this article is drawn from the large-scale, 2-year, mixed methods study, Best Practice in Grouping Students, which includes two randomised controlled trials, student-and-teacher questionnaire surveys and qualitative data collection including teacher interviews and student focus groups (EEF, n.d.; Francis *et al.*, 2016). We have described elsewhere the recruitment and randomisation to intervention and control groups of the schools participating in the study (Taylor *et al.*, 2016). An element of this wider study is the intervention Best Practice in Setting, which was implemented July 2015–June 2017 and is being evaluated via RCT to ascertain the impact of the interventions on attainment and self-confidence. The sample for the Best Practice in Setting trial comprised 126 English secondary schools, with 64 in the intervention group and 62 in the control group. Schools in the intervention group were instructed to allocate students to English and mathematics sets according to best practice principles including allocation by KS2 test results only and regular movement between sets according to internal assessment results only. Schools in the control group were invited to continue with their usual set allocation practices. English and mathematics were

chosen because they are 'core' subjects in the English National Curriculum and also represent very different content and pedagogy.

This article draws on the baseline teacher survey, conducted in autumn 2015, and interviews with teachers in intervention and control schools, conducted in the second year of the study.

Teacher questionnaire survey

English and mathematics teachers from the participating schools were invited to complete an online survey between November and December 2015. Responses were collected from 597 teachers from 82 schools participating as intervention or control schools within the Best Practice in Setting trial. It is not possible to state a precise response rate as the exact size of the target population is not known. However, assuming that an average-sized secondary school is likely to have around eight teachers per English/mathematics department teaching Year 7 students, with a sample of 81 English and 120 mathematics departments the response rate is estimated to be approximately 37%.

This article analyses responses to two questions that asked teachers to identify which sources of information were used to allocate Year 7 attainment groups in English and Mathematics, respectively. Teachers were only asked to respond regarding their own subject. Teachers were asked:

> Which sources of information are used to allocate students to Year 7 English [mathematics] classes/sets/streams this year and how does this compare with last year? Please tick all that apply, but exclude exceptional cases such as students arriving from overseas.

Teachers were provided with the following options: National Curriculum KS2 test results; National Curriculum KS2 teacher assessments; a commercially available test (e.g. CATS/MidYIS); the school's own test of student attainment; teacher judgements of students' abilities; teacher observations of student behaviour; a parent's judgement of their child's ability; information about students' feeder schools; the results of random allocation of students; I'm not sure which sources. The intention was to investigate whether methods of allocation were simply based transparently on prior attainment or whether other factors had an influence; and whether the intervention had improved this practice in comparison with the schools in the control group.

Teacher interviews

Teacher interviews were carried out in 10 schools: five from the intervention group and five from the control group. These schools were purposefully sampled from the 126 secondary schools participating in the trial with the intention of ensuring a variety of geographical contexts, student populations, grouping practices and Ofsted [6] judgements (see Table 1).

The 10 schools were each asked to nominate two English teachers and two maths teachers currently teaching Year 8 groups to be interviewed. The individual teacher interviews were semi-structured and typically lasted 40 min. A total of 34 teachers (15 English, 19 maths) were interviewed. The sample includes more maths than English teachers because two of the schools are participating for maths only. There are equal numbers of women ($n = 17$) and men ($n = 17$) in the sample. Teachers self-categorised their ethnicity, with 23 identifying as White, two as Black and nine as Asian.

The interviews were audio-recorded, transcribed professionally and pseudonyms assigned. Transcripts were coded for key themes in Nvivo. The coding work was divided so that one researcher coded several transcripts to draw up a preliminary coding framework consisting of umbrella nodes and sub-nodes that were both descriptive and more interpretive in nature. The framework was then discussed, amended and agreed by other team members. Nodes were defined to ensure the categories are understood by all and would make sense to the external observer.

Transcripts were coded by one member of the research team. To ensure the consistency in coding and to enhance the trustworthiness of the process, a coder comparison coding exercise was undertaken mid-way through data coding (Bazeley and Jackson, 2013). Two additional members of the research team, who were involved in the data collection, each coded the same randomly selected transcript, using the established coding structure. The coding of additional coders was then compared against the 'lead' coder with the assistance of a coding comparison query in NVivo. The query allowed the identification of potential problem areas, such as the lower agreement scores between different coders on individual nodes across the document. The coders achieved a relatively high agreement score: 90% or more on half of the nodes coded, with disagreement in coding no more than 30% on any node. Disagreements were then resolved before the final analysis.

For the present analysis, we have used extracts coded at the node 'Year 7 grouping arrangements'. Extracts were read and re-read by the first author and emergent themes identified. The extracts were then re-read and re-coded for these emergent themes, which were then discussed and clarified with the second author.

4. RESULTS

Teacher questionnaire survey

RQ1. What has been the effect of the Best Practice in Setting intervention on set allocation practices?

Teachers were asked to identify which sources of information were used to allocate students to Year 7 groups in English and Mathematics. The following analysis includes only responses from teachers who reported that their school grouped students

TABLE 1. Summary of schools participating in teacher interviews

School	Location	Trial group	Free school meals	Ofsted rating (at time of interviews)	Set levels	English teachers	Maths teachers
Y	Outer London	Intervention	22%	Outstanding	3 streams, set within streams	2	2
X	Outer London	Intervention	17%	Good	5 – mathematics mixed attainment – English	n/a	2
R	Market town, north west England	Intervention	12%	Good	4	2	2
T	City, Yorkshire	Intervention	24%	Inadequate	5	2	2
Q	City, south coast	Intervention	16%	Good	4	2	2
V	Outer London	Control	21%	Good	7 – mathematics mixed attainment – English	n/a	2
W	Semi-rural, Midlands	Control	7%	Requires improvement	6	2	2
S	Outer London	Control	21%	Good	6	2	2
U	Rural, south east England	Control	28%	Requires improvement	5	2	2
P	Market town, Midlands	Control	5%	Outstanding	6 – mathematics 3 – English	1	1

into sets. Results show the numbers of teachers who reported using the specified sources of information to allocate students to sets, by trial group (intervention and control). This question was answered by 66 English and 134 mathematics teachers in the intervention group and 56 English and 206 mathematics teachers in the control group. Only teachers whose department was participating in the trial were included in the analysis.

Teachers could select more than one source of information, therefore the total number of responses was greater than the number of teachers who responded. More than one teacher may have responded from each school, so these results need to be interpreted with care. Findings are summarised in Table 2.

Overall, KS2 tests were the most frequently reported source of information about students. However, many teachers reported that they used other data to establish groupings.

There was some evidence of the impact of our Best Practice in Setting intervention, although this was not indicative of high fidelity (complete fidelity would show 100% reported application of KS2 tests from intervention schools, without use of other data). More English teachers in the intervention group reported using KS2 tests than in the control group ($X^2 = 4.78$, $p < 0.05$). However, both intervention and control group English teachers reported using other sources of information. No other differences reached significance. This suggests that the intervention instructions have encouraged application of KS2 test results.

The pattern was somewhat different for mathematics with the percentage of teachers reporting use of KS2 test results very similar for the intervention and control groups. This may reflect differences in the KS2 assessment regime between mathematics and English, where all of mathematics is assessed through a test, while only

TABLE 2. *Sources of information for setting in English and mathematics, 2015/16 (teachers in participating departments only)*

Sources of information	English				Mathematics			
	Intervention		Control		Intervention		Control	
Total teachers	66	100%	56	100%	134	100%	206	100%
KS2 test	53	89%	35	73%	99	74%	151	75%
KS2 teacher assessment	24	42%	25	54%	45	34%	78	38%
Commercial test	26	42%	18	48%	35	26%	66	32%
School's test	20	35%	19	38%	61	46%	118	59%
Teacher judgement	18	30%	19	39%	34	25%	71	35%
Teacher observation	8	14%	6	11%	10	7%	40	20%
Parental judgement	0	0%	0	0%	0	0%	2	1%
Feeder school information	10	17%	5	9%	8	6%	19	9%
Random allocation	0	0%	1	2%	1	1%	3	1%
I'm not sure	9	14%	3	7%	4	3%	14	7%
Total responses	168		131		297		562	

the reading and spelling, punctuation and grammar components of English are assessed by test with writing being teacher-assessed only. This may lead English teachers to favour combining test marks with teacher assessments. Alternatively, it may reflect differences between mathematics and English teachers in preference for testing over teacher assessment. The latter interpretation is supported by the higher percentage of mathematics teachers who reported using the school's own test to allocate students to sets ($X^2 = 4.50$, $p < 0.05$).

In a separate analysis, the number of teachers stating that *only* KS2 test results were used was established. This is summarised in Table 3. This supports the view that participation in the intervention group makes it more likely to use KS2 alone to allocate students to sets, as mathematics teachers in the intervention group were more likely to report use of KS2 tests only ($X^2 = 5.60$, $p < 0.05$) and the difference approached significance for English teachers ($X^2 = 3.24$, $p = 0.07$).

The results show some encouraging indications of modified set allocation practices in intervention schools. However, these modifications are limited and there was substantial use of sources of information other than national KS2 results to allocate students to sets. The qualitative data help us shed light on why this is, when 'ability' grouping is supposed to be based on prior attainment.

Teacher interviews

RQ2. How do the techniques and approaches applied by schools in setting secondary school pupils facilitate and/or constrain equitable practice?

Theme: Use of data to assign students to sets

All teachers interviewed reported using KS2 tests to allocate students to sets, confirming findings from the questionnaire, but there was great variation in how test results were used. Some teachers reported that KS2 test results were used alone: 'we rank order them [by KS2 result] and then apportion the groups in terms of that way' (Clyde, School W, English). Others reported that KS2 results were used in combination with other sources:

TABLE 3. *Use of KS2 test data as the sole source of information for setting in English and mathematics, 2015/16 (teachers in participating departments only)*

Sources of information	English				Mathematics			
	Intervention		Control		Intervention		Control	
Total teachers	66	100%	56	100%	134	100%	206	100%
KS2 test **only**	12	18%	4	8%	34	28%	31	15%

[groups were initially] based on SATS and teacher-assessed work [...] And then after October half-term [...] we used the NFER reading tests and the graded 100 scores to allocate them into sets. (Ellie, School T, English)

Sources were combined in different ways, for example calculating an average score, or refining or triangulating decisions. Andrew (School P, mathematics) described how a number of sources of information about students are combined in an ongoing process:

[T]his year we took the standardised score and we rank ordered those. [...] If there's any tie with the standardised score we look at the raw score. And then what we've done this year as well, we've baseline tested the students, so again if there's any tie between the raw score we then look at the baseline test. For a small number of students we also looked at the baseline tests and if there were any anomalies we made a judgement call, but given that we assess the students quite regularly there is lots of room for manoeuvre.

While some schools used minimal additional evidence, it is clear that some approaches were highly complex. Teachers may have felt that they were executing a sophisticated and accurate grouping strategy, but these approaches risked reducing equity by introducing opportunities for biased judgements to influence the process.

Some schools appeared to be streaming ('they're in the same [literacy-based] group for everything', Anthony, School U, mathematics), or setting within streams ('[the bands are across] all subjects', Patrick, School X, English). Both of these practices represented poor practice and risked compounding inequity as streaming is typically associated with greater inflexibility of movement and fixed ability thinking.

Theme: Movement between sets

In our conversations with teachers, it became apparent that their practices of moving students between sets would also have a significant impact on equity. As with initial allocation to sets, teachers reported that they used both assessment data and teacher perceptions. Jamie (School W, mathematics) described how his department's approach allowed movement based on both objective assessment and potentially biased teacher perceptions:

It tends to be after a summative assessment. [...] But again, if a teacher feels that there is a need for a student to be moved, they can be moved at any time.

In other departments, movement was in response to longer-term patterns: 'it wouldn't be based on just one assessment, it would be based on numerous assessments' (Clyde, School W, English). In fact, Clyde regarded set movement as 'based on the teacher's opinion, rather than data-driven, that that is informed by the data that they have collected from that student'. Chloe (School R, English) felt that summative assessments didn't give the 'full picture of everything [...] they might just have

had a bad day' and stressed the importance of 'build[ing] up that picture over time' before moving a student. Worryingly, teachers mentioned factors such as 'attitude to learning' (Larry, School Q, mathematics) and 'work ethic' (Charleen, School W, English), perceptions of which are known to be influenced by student characteristics (e.g. Gillborn and Youdell, 2000). In other words, combining assessment data with teacher perceptions over the long term (rather than acting on summative assessment) is a subjective activity where, we would argue, equity could be reduced as teacher biases influence the interpretation of achievement patterns.

Theme: Lack of clarity of grouping practices

As a final note in this discussion, some teachers – even heads of subject – seemed very unclear about the actual grouping processes being used in their schools. At School U, teachers contradicted each other to such an extent that we sought to clarify the setting structure at a second visit to the school. At School Q there was also a high degree of disagreement between staff about how bands and sets were structured. In both of these schools, staff turnover was high – all the teachers interviewed at School Q had been in post less than 3 years and there had been nearly 100% turnover in English and maths teachers at School U within the past year. This level of instability may have prevented classroom teachers and subject leaders from engaging in the decision-making process about set allocation.

RQ3. What factors influence and constrain set allocation practices?

Theme: Practical and financial factors

Confirming previous research, we found that setting is constrained by practical and financial factors. Where there were bigger cohorts or where schools expanded, schools responded by having more teaching groups and/or larger class sizes: Patrick (School Y, English) described how 'class sizes have jumped from twenty-six, or twenty-four, twenty-six to like thirty', while Andrew (School P, mathematics) claimed that the increased number of students in higher sets was because 'money's a little bit tight and then the numbers of teachers is also a bit tight'. However, provided that allocation of students to sets was transparent, this should not have had a negative impact on equity.

Nevertheless, practical and financial factors did influence equity of allocation. In some larger schools, timetable flexibility was achieved through banding: dividing a year group into two or more subgroups that were then timetabled as a block. Banding permitted teachers to teach more than one class within a year group and was therefore an economical way of employing specialist teachers. Furthermore, in some schools the bands were set up at different 'ability' levels and operate as streams. Samira (School T, English) described the system within her school: '7A1, 7A2, I have 7A2 which is the A band, and then the B band, B1 and B2'. Students could not move between bands A and B during the school year. At School Y, Patrick told us he 'cannot take them from [middle band] and put them into [lower

band], that's down to pupil progress' (the middle leader responsible for pupil progress within a school year group) – in other words, movement between streams was highly restricted and had to be negotiated with colleagues.

In some cases, the structure of the school timetable required English or mathematics to share groupings with one or more other subjects. For example, at School R, English was timetabled against RE. Dawn (School R, English) had not found this a problem for the English department, 'if we wanted to make a change in English, RE have always been happy to oblige with it', although this partial streaming approach risked creating self-fulfilling prophecies in RE attainment.

By contrast, at School S, English and science were set against each other causing English teacher Sanjiv concern. Class numbers were capped because of health and safety requirements, meaning 'a block of 84 students in the top three sets needs to be the same 84 students in science'. Sanjiv was frustrated because his view was that 'I have got 90 spaces in my top three sets' and 'six pupils [...] are being failed by a setting system and [have] to stay in Set 4, 5 or 6'. This situation illustrates how set allocation can be influenced by entirely unrelated practical factors.

Theme: Attitudes and beliefs

Beyond financial and practical constraints, attitudes held by teachers further influenced and constrained set allocation and movement. Teachers frequently expressed mistrust of KS2 tests and of primary schools more generally. Danielle (School U, mathematics) believed that 'sometimes the [KS2] data isn't 100 percent reliable'. Sophie (School Q, mathematics) gave the specific example of a Year 8 student who 'apparently got a level 5 at Key Stage 2' but 'he can't times, he can't add, he can't subtract. He doesn't know his times tables and we do kind of wonder how much he was supported in his Key Stage 2 SATs paper'. This scepticism about primary school practices led teachers to bring additional sources of information to bear on set allocation processes, thus potentially reducing equity.

As noted before, ease of movement between sets was reduced if schools used a banding/streaming system. However, ease of movement was also influenced by teacher attitudes towards the desirability of moving students between sets. Candice (School V, mathematics) reported that she had seen 'more [movement] than in other schools that I've worked in', while Dawn (School R, English) stated that 'we try to keep it to a minimum'. While some teachers were willing to move students up a set, they were often more reluctant to move them down. Hayley (School R, mathematics) describes strategies her department used before moving set as 'the last option': 'if it's behaviour, they can go on report. If it's achievement, we have a discussion with them about what they can do'. Richard (School Q, English) argued that 'I think actually having the consistency in the setting is as important, in some ways, as being sure someone is in the right group', feeling that 'a good relationship with a teacher [...] can be far more powerful'. Karen (School W, mathematics) had a powerful expectation that the amount of movement between sets would reduce as students go up the school: 'as we go up the school we get fewer and fewer moves,

really, as the kind of sets start to settle'. Teacher beliefs and values were clearly a strong influence on set movement.

Teacher beliefs about student 'ability' also constrained movement. Labelling of students by 'ability' was endemic, for example teachers at School Y labelled students by band membership. In three schools, set allocation was according to a future target, rather than past attainment, for example: 'it's grouped by their prediction, based on KS2, for what they will achieve at the end of GCSE' (Dawn, School R, English). Once a trajectory had been assigned to a child, there was an understanding that they should continue on that path and not slip from it.

5. DISCUSSION: WHY IS IT DIFFICULT FOR SCHOOLS TO ESTABLISH EQUITABLE SET ALLOCATION PRACTICES?

As stated earlier, our contention was that more equitable grouping could be established through a set of prescribed practices that sought to mitigate factors identified in the research literature as contributing to inequitable outcomes; including two relating to practices of set allocation:

(1) Transparent allocation of students to sets, based purely on prior attainment (KS2 test results).
(2) Regular opportunities for students to move between sets on the basis of internal assessment results.

We envisaged that these stipulations would facilitate equality of opportunity via encouraging a focus on attainment rather than 'ability', transparent practice accordingly with the removal of biases that support inequitable trends, and opportunities to move between sets if/when attainment outcomes change.

In this article we have provided substantial evidence that teachers make use of a range of information when assigning students to sets, including KS2 tests, but also using sources such as teacher judgements that are liable to bias (Campbell, 2015; Timmermans *et al.*, 2015). We found that even in our intervention group only 18% of English teachers and 28% of mathematics teachers reported that KS2 test results were the sole data source. It was encouraging to note that this represents more rigorous practice than in our control group, in line with the expectations of the intervention. Nevertheless, it was much more frequent for teachers to report combining KS2 data with other sources, illustrating how counterintuitive our stipulated practices were for schools, and how other factors may militate to undermine best practice. We have also shown that there is also a significant amount of 'tweaking' of groups – a process that may admit implicit biases into the setting process.

We have provided evidence that equity may be constrained by partially (or fully) streamed classes, resulting from practical constraints such as size of intake, staffing and finance. Streaming risks restricted set movement and increased labelling (and thence, self-fulfilling prophecy). Labelling by 'ability' was evident in a number of schools. Many teachers are averse to moving students between sets or are prevented

from doing so by complex timetabling arrangements. Finally, equity may be constrained in some schools because teachers are not fully aware of what their practices are. Constraints such as these may make the enactment of a structural intervention impossible in some schools, without complete overhaul of the timetabling process – likely to be a significant task (albeit a very important one).

These difficulties likely combine and reinforce one another, given the complex nature of schools as organisations. The findings from teacher respondents in our intervention group illustrates the scale of challenge to change practice: the data presented here were collected as part of a high-profile research project with the explicit goal of improving the impact of setting on attainment and self-confidence, to make it more equitable. Schools in the intervention group were given instructions as to the structures and practices required. While we see evidence of some differences in practices between intervention and control groups, with the intervention group generally conforming more closely to transparent practice in attainment grouping, it is still a minority of schools that have followed our instructions assiduously.

Cuban (2013) suggests that policymakers make three errors of thinking in relation to policy implementation, believing that: (1) structural changes will change teaching practices and outcomes; (2) schools and classrooms are complicated rather than complex systems; and (3) the policymaker worldview is the same as that of teachers. We have not focused in this article on changes in teaching practices, but points (2) and (3) may help us to understand why our intervention had limited success in increasing the equity of setting practices.

Addressing point (2), we observe that schools are complex organisations, made up of a multiplicity of interdependent and diverse elements, all potentially interacting (Sargut and McGrath, 2011). The nature of a complex organisation is that it is difficult to isolate a discrete part of the organisation to apply the intervention. The same set of instructions (our proposed intervention) results in unpredictably different outcomes in different schools, depending on the nature of the interactions within the school. Figure 1 attempts to capture some of the interacting influences on setting in a typical school, to illustrate the difficult terrain that such instructions must navigate on the way to enactment. The process of enactment is further confounded by the lack of a common vocabulary of attainment grouping (see Taylor *et al*., 2016), making it very difficult to give unambiguous instructions or make reliable changes to structures and practices. We have noted earlier the confusion even within schools about their own grouping practices.

Finally, and addressing point (3), Cuban observes that teachers and policymakers inhabit separate worlds, differing in focus, aims and purposes. A similar distinction can be drawn in the case of our study, between the separate priorities and motivations of teachers, school leaders, and researchers, such that the intervention does not get implemented as the research team intended or necessarily achieve the hoped-for outcomes. Schools are often able to make some superficial changes, but frequently unwilling or unable to disrupt their accustomed habits.

Nevertheless, we also see cause for optimism in our findings. Our data suggests that integrity of allocation practice in setting by attainment was improved in the
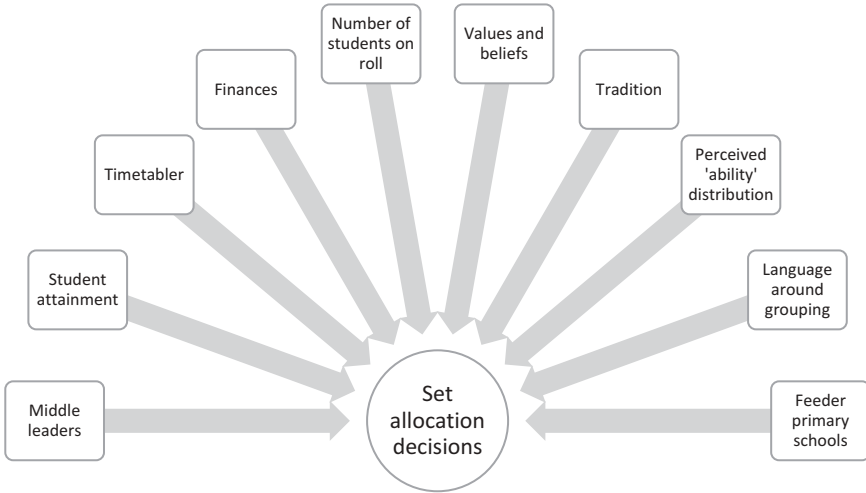
Figure 1. Potential influences on set allocation decisions in a typical school

intervention schools, demonstrating that improvement is possible. Even small improvements in practice may have important impacts on the young people experiencing it. This improvement might also be developed and extended if schools embrace the importance of this issue.

Achieving equitable setting

We have shown previously that our intervention has successfully improved set allocation practices in some schools. However, significant challenges remain. We suggest now that strategies specifically designed for achieving change in complex organisations will need to be employed in order to achieve even more equitable practices.

Sargut and McGrath (2011) suggest a number of strategies for working with complex systems. One strategy is to 'ensure diversity of thought' (p. 76). An institutional culture that values diversity and challenge may be more likely to allow the questioning of unintentionally unfair consequences. A second strategy is 'decoupling and redundancy': essentially this reduces the complexity. For example, a simpler timetable with fewer interdependencies will result in fewer unintended outcomes. A third strategy is the 'real options approach' of making mistakes 'cheaply and early, learning from them and increasing your resilience as you go' (p. 76). A reflective and questioning organisation will be watchful for the outcomes of decisions and adjust practices swiftly if unjust and unintended consequences arise. Sargut and McGrath recommend 'post mortems' and the triangulation of hard and soft data – an approach likely to facilitate rich evaluative opportunities. However, as

we have shown, schools may have limited capacity to be aware of and respond to their practices, so this may not be suitable in all schools.

In terms of improving interventions in complex organisations, there may be key lessons from two aspects of Sargut and McGrath's analysis. They emphasise the importance of outliers as more relevant than the average case. In designing our intervention we worked with a number of schools with a good track record of successful practice. By seeking out schools with an exceptional record of equitable outcomes and examining their practices in detailed case studies we could develop an improved intervention with potential for more equitable outcomes. Another recommendation from Sargut and McGrath is the role of storytelling and the use of counterfactuals. We have noted elsewhere (Taylor *et al.*, 2016) that research evidence alone is insufficient to persuade many schools to more equitable practices. In this article we have explored some of the structural factors that inhibit equity and that showed only moderate improvement through our intervention. Case studies as suggested formerly would not only provide understanding of how to achieve equity, but could also provide compelling narratives that might advocate for the necessity and possibility of reform.

The question remains as to whether, in the current social and educational landscape, *any* attainment grouping practice can be 'equitable'. Our focus has been on improved practice in application of setting. The findings of our comparison between intervention and control show *some* potential here. However, as we have also demonstrated, the complex relationships and arrangements within schools frequently result in good intentions being frustrated and change. It seems likely therefore that attainment grouping will inevitably be socially based (and therefore inequitable, Kelly, 2004). Professional dialogue with educators needs to focus on the extent to which commitments to segregation by attainment can realistically be underpinned by transparent practices such as those specified within our intervention; or whether our findings concerning the challenges of improving practice indicate that educators would be better investing in alternative practices such as mixed-attainment grouping.

8. DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

9. NOTES

[1] The term 'tracking', found in the US literature, sometimes refers to practices analogous to setting and sometimes to streaming (Gamoran and Nystrand, 1994).

[2] We dislike the term 'ability' as it implies that each child has a fixed and measurable, general level of ability. Instead we use the term attainment grouping to reflect that, at least in principle, schools are grouping by some kind of measured attainment.

[3] Cognitive Ability Tests (CATs, https://www.gl-assessment.co.uk/products/cognitive-abilities-test-cat4/) and MidYIS tests (http://www.cem.org/midyis) are commercially available tests of 'ability' designed for use with school-age children.

[4] Schools in the intervention group were required to test students at three intervals within the 2-year period of the intervention, and move students between sets accordingly, on the basis of the test results.

[5] We also challenged the concept of 'ability' in our professional development with teachers, drawing on the work of Dweck (1986) and of Hart *et al*. (2004). Our aim was to challenge teachers' practice of labelling students as possessing a fixed level of ability and to encourage teachers to have high expectations of all students regardless of prior attainment. In this way we hoped to reduce the influence of ability labelling in students' school experiences, at least in the subjects involved in the intervention.

[6] Ofsted is the school inspectorate in England.

ORCID

Becky Taylor http://orcid.org/0000-0002-7257-4463
Becky Francis http://orcid.org/0000-0002-7257-4463
Anna Mazenod http://orcid.org/0000-0003-0175-1634

10. REFERENCES

Ball, S. J. 1981 *Beachside Comprehensive: A Case-Study of Secondary Schooling* (Cambridge, CUP Archive)

Bazeley, P. and Jackson, K. 2013 *Qualitative Data Analysis with Nvivo* (London, Sage)

Boaler, J., Wiliam, D. and Brown, M. 2000 Students' experiences of ability grouping-disaffection, polarisation and the construction of failure, *British Educational Research Journal*, 26(5), 631–648. doi:10.1080/713651583

Bosworth, R. 2013 What sort of school sorts students? *International Journal of Quantitative Research in Education*, 1(1), 20–38. doi:10.1504/IJQRE.2013.055639

Campbell, T. (2014). *Selected at Seven: The Relationship between Teachers' Judgments and Assessments of Pupils, and Pupils' Stream Placements*. Department of Quantitative Social Science Working Papers.

Campbell, T. 2015 Stereotyped at seven? Biases in teacher judgement of pupils' ability and attainment, *Journal of Social Policy*, 44(03), 517–547. doi:10.1017/S0047279415000227

Charlton, E., Mills, M., Martino, W. and Beckett, L. 2007 Sacrificial girls: A case study of the impact of streaming and setting on gender reform, *British Educational Research Journal*, 33(4), 459–478. doi:10.1080/01411920701434011

Coleman, J. S. 1975 What is meant by 'an equal educational opportunity? *Oxford Review of Education*, 1(1), 27–29

Cuban, L. 2013 Why so many structural changes in schools and so little reform in teaching practice? *Journal of Educational Administration*, 51(2), 109–125. doi:10.1108/09578231311304661

DfE 2015 *National Curriculum Assessments at Key Stage 2 in England, 2015 (Revised): Statistical First Release* (London, DfE)

Dunne, M., Humphreys, S., Dyson, A., Sebba, J., Gallannaugh, F. and Muijs, D. 2011 The teaching and learning of pupils in low-attainment sets, *Curriculum Journal*, 22(4), 485–513. doi:10.1080/09585176.2011.627206

Dunne, M., Humphreys, S., Sebba, J., Dyson, A., Gallannaugh, F. and Muijs, D. 2007 *Effective Teaching and Learning for Pupils in Low Attaining Groups* (London, DFES)

Dweck, C. S. 1986 Motivational processes affecting learning, *American Psychologist*, 41 (10), 1040–1048. doi:10.1037/0003-066X.41.10.1040

EEF. (n.d.) Best practice in grouping students. Available at https://educationendowment foundation.org.uk/our-work/projects/best-practice-in-grouping-students/(Accessed 20 June 2017).

Francis, B., Archer, L., Hodgen, J., Pepper, D., Taylor, B. and Travers, M.-C. 2016 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. doi:10.1080/0305764X.2015.1093095

Francis, B., Connolly, P., Archer, L., Hodgen, J., Mazenod, A., Taylor, B., Tereshchenko, A., Pepper, D., and Travers, M. 2017 Attainment grouping as self-fulfilling prophesy? a mixed methods exploration of self confidence and set level among year 7 students. *International Journal Of Educational Research*, 86, 96-108. doi: 10.1016/j. ijer.2017.09.001

Gamoran, A. and Nystrand, M. 1994 Tracking, instruction and achievement, *International Journal of Educational Research*, 21(2), 217–231. doi:10.1016/0883-0355(94)90033-7

Gillborn, D. and Youdell, D. 2000 *Rationing Education: Policy, Practice, Reform, and Equity* (Buckingham, Open University Press)

Hallam, S. (2012). *Streaming and Setting in UK Primary Schools: Evidence from the Millennium Cohort Study.* Paper presented at the FORUM: for Promoting 3-19 Comprehensive Education.

Hallam, S. and Ireson, J. 2005 Secondary school teachers' pedagogic practices when teaching mixed and structured ability classes, *Research Papers in Education*, 20(1), 3–24. doi:10.1080/0267152052000341318

Hallam, S. and Parsons, S. 2013 The incidence and make up of ability grouped sets in the UK primary school, *Research Papers in Education*, 28(4), 393–420. doi:10.1080/ 02671522.2012.729079

Hallinan, M. T. 1996 Track mobility in secondary school, *Social Forces*, 74(3), 983–1002. doi:10.1093/sf/74.3.983

Hamilton, L. and O'Hara, P. 2011 The tyranny of setting (ability grouping): challenges to inclusion in Scottish primary schools, *Teaching and Teacher Education*, 27(4), 712–721. doi:10.1016/j.tate.2010.11.009

Hart, S., Dixon, A., Drummond, M. and McIntyre, D. 2004 *Learning without Limits* (Maidenhead, Open University Press)

Higgins, S., Katsipataki, M., Coleman, R., Henderson, P., Major, L., Coe, R. and Mason, D. 2015 *The Sutton Trust - Education Endowment Foundation Teaching and Learning Toolkit* (London, Education Endowment Foundation)

Ireson, J. and Hallam, S. 2001 *Ability Grouping in Education* (London, Paul Chapman)

Jackson, B. 1964 *Streaming: An Education System in Minature* (London, Routledge and Kegan Paul)

Kelly, S. 2004 Are teachers tracked? On what basis and with what consequences, *Social Psychology of Education*, 7(1), 55–72. doi:10.1023/B:SPOE.0000010673.78910.f1

Kutnick, P., Sebba, J., Blatchford, P., Galton, M., Thorp, J., MacIntyre, H. and Berdondini, L. 2005 *The Effects of Pupil Grouping: Literature Review* (London, DFES)

Linchevski, L. and Kutscher, B. 1998 Tell me with whom you're learning, and I'll tell you how much you've learned: mixed-ability versus same-ability grouping in mathematics, *Journal for Research in Mathematics Education*, 29(5), 533–554. doi:10.2307/749732

Modica, M. 2015 *Race among Friends: Exploring Race at a Suburban School* (New Brunswick, NJ, Rutgers University Press)

Moller, S. and Stearns, E. 2012 Tracking success: high school curricula and labor market outcomes by race and gender, *Urban Education*, 47(6), 1025–1054. doi:10.1177/0042085912454440

Muijs, D. and 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. doi:10.1080/00131881.2010.524750

Oakes, J. 1985 *How Schools Structure Inequality* (New Haven, Yale University Press)

OECD 2012 *Equity and Quality in Education: Supporting Disadvantaged Students and Schools* (Paris, OECD Publishing)

OECD 2013 *Selecting and Grouping Students PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices* Vol. IV71–92. (Paris, OECD Publishing)

Sargut, G. and McGrath, R. G. 2011 Learning to live with complexity, *Harvard Business Review*, 89(9), 68–76

Shaw, B., Menzies, L., Bernardes, E., Baars, S., Nye, P. and Allen, R. 2016 *Ethnicity, Gender and Social Mobility* (London, Social Mobility Commission)

Slavin, R. E. 1990 Achievement effects of ability grouping in secondary schools: a best-evidence synthesis, *Review of Educational Research*, 60(3), 471–499. doi:10.3102/00346543060003471

Taylor, B., Francis, B., Archer, L., Hodgen, J., Pepper, D., Tereshchenko, A. and Travers, M.-C. 2016 Factors deterring schools from mixed attainment teaching practice, *Pedagogy, Culture & Society*, 25(3), 327–345. doi:10.1080/14681366.2016.1256908

Taylor, B. and Sloan, S. (2016). *Best Practice in Grouping Students? Characteristics of Students in English and Mathematics 'Ability' Set Groups in English Secondary Schools*. Paper presented at the National Pupil Database User Group, University of Bristol, 7 September 2016.

Timmermans, A. C., Kuyper, H. and Van der Werf, G. 2015 Accurate, inaccurate, or biased teacher expectations: do Dutch teachers differ in their expectations at the end of primary education? *British Journal of Educational Psychology*, 85(4), 459–478. doi:10.1111/bjep.2015.85.issue-4

Van de Gaer, E., Pustjens, H., Van Damme, J. and 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. doi:10.1080/01425690600750478

Correspondence
BECKY TAYLOR
Department of Education, Practice and Society
UCL Institute of Education
London, UK
Email: becky.taylor@ucl.ac.uk