

Teacher ‘quality’ and attainment grouping: the role of within-school teacher deployment in social and educational inequality

Professor Becky Francis¹, Professor Louise Archer¹, Dr Nicole Craig², Dr Anna Mazonod¹, Dr Becky Taylor¹, Dr Antonina Tereshchenko¹, Professor Jeremy Hodgen¹, Professor Paul Connolly²,

1 UCL Institute of Education, UK; 2 Queens University Belfast, UK

Corresponding author:

Professor Becky Francis

Director, UCL Institute of Education

20 Bedford Way

London, UK

B.Francis@ucl.ac.uk

Teacher ‘quality’ and attainment grouping: the role of within-school teacher deployment in social and educational inequality

Abstract

Prior research suggests that where pupils are 'tracked', better qualified, more experienced teachers tend to be deployed to higher attainment groups, at the expense of pupils in lower tracks. This is especially pertinent from a social justice perspective, given consistent findings in the UK that pupils from socially-disadvantaged backgrounds are over-represented in low attainment groups. This article draws on data from 380 teachers, drawn from 126 secondary schools in England, and interviews with 118 Year 7 students, to examine whether these findings from prior research in the US and elsewhere extend to the case of England in the present day. Findings show *some* evidence of these inequitable tendencies: those teachers highly qualified in their taught subject were less likely to be allocated to low sets. We also examine whether an intervention designed to encourage more equitable distribution had any impact on practice, and find tentative evidence that deployment in intervention schools had been impacted in relation to teacher subject qualifications. Pupils believed that teachers of higher sets had higher expectations and standards of behaviour, whereas those for low sets were seen to be unhelpfully indulgent, indicating a need for research attention to pedagogy and tracking. Findings are analysed from a social justice perspective, with interest in the consequences of inequitable distribution of teachers for the reproduction of social inequality.

Teacher ‘quality’ and attainment grouping: the role of within-school teacher deployment in social and educational inequality

While research on segregation by attainmentⁱ remains a contested areaⁱⁱ, nevertheless is widely established that practices of segregation by attainment tend to disadvantage those pupils allocated to low attainment groups, who make less progress than their peers in higher attainment groups (Slavin, 1990; Suknandan & Lee, 1998; Ireson & Hallam, 2001; Wiliam & Bartholomew, 2004; Kutnick et al, 2005; Higgins/EEF, 2016). It is also consistently shown that pupils from socio-economically disadvantaged backgrounds (and those from certain minority ethnic groups) are over-represented in low attainment groups (Jackson, 1968; Kutnick et al, 2005; Dunne et al, 2007; Taylor et al, 2018). We have argued that this means pupils from disadvantaged backgrounds in low attainment groups are subject to a double disadvantage (Francis et al, 2017a), as they enter the education system disadvantaged in relation to their more affluent peers (Waldfoegel & Washbrook, 2010), and are then subject to practices known to have a detrimental impact on their performance.

The various potential explanations for the sub-optimal progress and outcomes for those placed in low attainment groups have also been subject to extensive research and debate. Our review of prior research identified seven explanations reflected in the research literature (Francis et al, 2017a). An especially controversial explanation among these is the *quality of teaching* experienced by pupils in low ‘ability’ⁱⁱⁱ groups. Although the findings of prior studies on this topic are not straight-forward, in their reviews of the literature both Slavin (1990) and Ireson et al (2001) maintain there is some evidence that teachers perceived as lower quality tend to be placed with lower ‘ability’ groups. Such findings are supported by more recent studies such as those of Papay and Kraft (2014) and Kelly (2004).

Evidently, the notion of ‘teacher quality’ is nebulous as well as controversial. In the United States, a substantial literature on ‘teacher effectiveness’ refers to teacher ‘quality’ and ‘effectiveness’ reasonably interchangeably (Boyd et al 2008). The status, validity and impact of ‘observable qualifications’ of effectiveness continue to be debated (Boyd et al, 2008; Hanushek & Rivkin, 2010), but conventional measures include indicators such as level and type of qualification, level of curriculum subject expertise, and length/type of teaching experience. Clearly, none of these indicators can be taken as guarantees of effective teaching or otherwise (indeed inconsistency in prior study findings indicate their relative weakness in this regard), and teacher ‘quality’ should not be seen as fixed or static. Nevertheless, quantitative and experimental research has often shown significance of these indicators across teacher populations. Some studies have found length of experience to be especially associated with effectiveness (Rockoff et al, 2011), with the thrust of findings on this topic

suggesting that the largest teacher gains from experience occur in the first five years of teaching (Harris & Sass, 2006; Clotfelter et al, 2006; Papay & Kraft, 2014) – albeit, Wiswall (2013) and Papay & Kraft (2014) show that, especially for maths achievement, there continues to be a return from experience throughout a teaching career. Clotfelter et al’s (2006) study shows that teacher licensure test scores correlate with pupil achievement in maths. Likewise, while the impact of certification levels remains debated, with qualifications being a relatively weak proxy for teacher quality (Shulman, 1986), Kane et al (2007) find that initial certification status does impact student test performance; and Rockoff et al (2011) also show that teacher maths knowledge significantly impacts pupil attainment. Coe et al (2014) likewise argue that teachers with strong subject knowledge make a greater impact on pupils’ learning. Boyd et al (2008) find that recruiting teachers with stronger credentials – e.g. test scores or ‘certification status’ – “could substantially improve student achievement” (p. 794).

Indeed, researchers such as Coe et al (2014) and Hattie (2013) maintain that – in terms of in-school factors - it is quality of teaching that makes the strongest impact on pupil outcomes across the board. Research by Sanders and Rivers (1996), Aaronson et al (2007) and Kane et al (2007) demonstrates the differential in student achievement gains attributed to differences in teacher effectiveness. Moreover, teacher quality is found to be especially impactful to the attainment of pupils from socio-economically disadvantaged backgrounds (Ainscow et al, 2012; Sutton Trust, 2011), and for those with low prior attainment (Black and Wiliam, 1998). It has also been asserted that teacher quality ranges widely within schools, in England (Husbands, 2012). This diversity in effectiveness has however been shown to map on to school and pupil demographic factors: extensive research in the United States has found show that more highly qualified, and more experienced, teachers tend to be matched with more socially advantaged students; the converse being the case for disadvantaged students (Lankford et al, 2002; Clotfelter et al, 2006). And in the UK, research by Sims and Allen (2017) demonstrates that pupils in the most disadvantaged quintile of schools are around twice as likely to have an unqualified teacher, and more likely to have an inexperienced teacher.

Relating these measures of teacher effectiveness or quality to the literature on attainment grouping, it has been suggested that higher groups are more likely to be allocated highly qualified and experienced teachers (Sukhnandan & Lee, 1998; Kelly, 2004). Whereas lower attainment groups have been found to be more often assigned to the least well-prepared teachers (Good and Marshall 1984, Oakes 1985); less likely to be taught by a subject specialist (Kelly, 2004), and to experience more changes of teacher (Boaler et al, 2000).

Papay & Kraft (2014) likewise find that novice teachers tend to be assigned to lower attaining students. Kelly's (2004) research additionally found that teachers designated to higher tracks had "much higher levels of perceived efficacy and satisfaction with teaching" (p. 69).

Hence, the existing research highlights: i) the importance of high quality teaching, especially for pupils from socio-economically disadvantaged backgrounds; ii) a tendency for less effective teaching to be provided to socially disadvantaged pupils, and to lower attaining groups; and iii) the over-concentration of socially-disadvantaged students *in* low attainment groups. We can deduce, then, the socially retrogressive impact of these trends.

As well as the 'quality' of the *teacher*, there is extensive evidence in the literature on segregation by attainment that the level of the 'ability' group to which a teacher is assigned influences the quality of the *pedagogy* provided, due to application of different expectations that in turn impact pace and quality of pedagogy (Murphy & Hallenger, 1989; Ireson et al, 2005; Mazenod et al, 2018). For example, Gamoran's (1992) review suggests that teachers of higher tracks were more enthusiastic and devoted more time to preparing lessons (see also Hallinan, 1984, for the converse being the case with low tracks). Ireson et al (2005) and Boaler et al (2000) show that teachers of high sets tend to provide fast-paced and challenging work, whereas pupils in low sets are often subject to slow-paced lessons which cover less curriculum material in terms of both breadth and depth (see also Gamoran, 1992). Research also shows that pupils in higher sets are given more homework (Ireson & Hallam, 2001).

As we have shown, given the relationship between set placement and social background, this trend is especially retrogressive.

The suggestion that low attaining groups are allocated poorer 'quality' teachers may seem counter-intuitive, given that these pupils might be seen as in especial need of help – or at least entitled to equality of opportunity in allocation of public resources. However, there are a range of potential explanations for the opposite to be the case. Longstanding sociological perspectives maintain that education systems reproduce social inequality. Bourdieu and Passeron (1977) famously analysed the practices of distinction embedded within the education system, and the symbolic cultural capital meted out to middle class pupils in the form of credentials that recognise and value the same capital in these pupils; reproducing inequality in outcomes. We have ourselves analysed longstanding public discourses on segregation by attainment and social class within schools to conclude that these practices are

produced as ‘natural’, and speak to “long-standing cultural phantasies of identity, aspiration and ‘natural order’” (Francis et al, 2017a, p.9). An array of research has illustrated how middle class parents tend to support setting and/or other forms of segregation by attainment in schools (Oakes & Guiton, 1995; Kohn, 1998) and to assume that their children will be (or ought to be) in the high sets (Ball, 2003; Reay et al, 2011) – and indeed, to campaign for them to be so (Kohn, 1998). Such assumptions are explained by the researchers as expressive of a middle-class ‘habitus’ reflecting competition, aspiration, and entitlement (characterised by the popular media as ‘sharp elbowed parents’).

Hence existing sociological hypotheses which might explain the direction of quality teaching resources towards high sets include a conscious or unconscious reproduction of social inequality via the education system. An alternative, if equally disturbing, hypothesis, is that schools have not been incentivised to prioritise low attaining pupils, and that therefore their needs have been overlooked. Since the early 1990s, the English education system has promoted the award of at least 5 GCSE^{iv} exam awards at grades A*-C (later, 5 GCSEs at grades A*-C to include Maths and English) as the key indicator of expected success for pupils at 16. Crucially, the proportion of pupils achieving this indicator featured in school ‘league tables’ with strong consequences for a schools’ ability to attract pupils, and also for Ofsted^v inspection outcomes. As such, schools were incentivised to focus on getting as many pupils above the grade C borderline as possible (leading Gillborn and Youdell [2000] to allude to practices of ‘educational triage’ operated for pupils at the ‘C/D borderline’). Within this, expending scarce resources on those pupils with little possibility of achieving a C at GCSE may have been seen to be an unproductive investment: indeed this is the suggested explanation by several contributors to ‘The Tail’ (2013), which focuses on England’s relative poor record with low attaining students (though the contributors do not attend to the issue of attainment grouping)^{vi}.

While this hypothesis focuses on the decision-making of managers in the deployment of teachers, it may be that teachers themselves also have a role. Studies such as those of Kelly (2004) and Gamoran (1992) suggest that teachers derive more satisfaction and/or enjoy teaching high sets more (Gamoran found they are more enthusiastic and devote more time to planning lessons for these groups, and conversely that teaching low sets was seen to require greater attention to discipline). As such, it might be that teachers tend to prefer teaching high attainment groups (see Murphy & Hallinger, 1989), or that such teaching is seen as more specialist/higher status. It might then follow that managers seek to reward and retain their highest performing teachers by facilitating them to teach high sets: in other words, that

placement with high set groups is functioning as an informal aspect of a reward and recognition system.

To recap, the research literature shows that teacher quality is of primary significance for pupil outcomes – especially for pupils from disadvantaged backgrounds and for low attainers. And there is evidence that access to teacher quality may be impacted by attainment grouping, with findings that low attainment groups – which contain disproportionate numbers of pupils from disadvantaged backgrounds – are actually subject to poorer quality teaching. If the case, this would lend significant weight to the claim that these pupils face systemic double-disadvantage, and overtly inequitable practices.

What prior research has not explored is efforts to remediate such practices, and whether or not such efforts might be fruitful. In other words, whether practices of segregation by attainment *inevitably* precipitate inequitable trends in teacher deployment quality, or whether these might be mitigated by purposeful equitable assignment of teachers. Moreover, much of the existing evidence is now reasonably dated (we have not identified US research on the topic of teacher quality in relation to tracking published in the last ten years, and there is an even longer gap in the UK context). And further, little among the existing body of work attends directly to issues of social inequality. Hence this article intends to innovate in three key ways. Firstly, it draws on a large-scale data-set that – innovatively within the corpus – includes both qualitative and quantitative data, to examine whether prior findings extend to the present day. Secondly, it analyses findings in relation to the issue of social justice and social in/equality. And finally, it explores the feasibility of more equitable deployment of teachers to attainment sets.

Background to the methodology

The data discussed in this paper draws from a large scale mixed-methods project ‘Best Practice in Grouping Students’, funded by the Education Endowment Foundation. The project seeks to address prior gaps in the literature, by exploring: a) whether practices of setting^{vii} by attainment that remediate factors identified in the literature as negatively affecting those in low attainment groups might improve young people’s progress; and b) what comprises good practice in mixed attainment pedagogy. It includes the following methods:

- Two two-year interventions: ‘Best Practice in Setting’, evaluated by a fully-powered RCT and ‘Best Practice in Mixed Attainment’, constituted as a feasibility study with

RCT design.^{viii} Both evaluations examine the impact or otherwise of practice in grouping students in Year 7 and Year 8 based on research evidence.

- Surveys of pupils and teachers involved in the study
- Qualitative individual interviews with students and teachers, and focus groups with students.

The research is undertaken in 139 secondary schools in England (divided into intervention or control groups), and involves instigating work with and tracking student cohorts from the beginning of Year 7 (11-12 years old) to the end of Year 8 (12-13 years old), focusing on pupils' experiences and outcomes in English and maths. English and maths were selected as the foci because: i) they are two subjects given longstanding priority in the English national curriculum and within school performance indicators; and ii) they represent diversity in content and pedagogy.

This article draws on two types of data from the broader study, to answer the research questions on 'teacher quality' and set allocation. The first part of this paper analyses the quantitative data from a teacher survey collected at the beginning of the Best Practice in Setting RCT, and the second part analyses qualitative data from interviews and focus groups with students in the same trial.

Quantitative data collected through the Teacher survey

In the academic year prior to the RCT, a teacher survey was designed and piloted with 12 teachers from schools participating in the pilot study for [Project Name] (see Archer et al, 2017, for a full account). Minor amendments were made to improve clarity and validity. The teacher survey was then distributed to teachers in schools participating in the RCT study, at the beginning of the study at the start of the school year. Given the specific interest in setting and assignment of teachers to sets, this paper focuses exclusively on the survey data from teachers' responses within the 'Best Practice in Setting' RCT.

The questionnaire asked teachers about the set groups they were presently teaching; practices of setting within their school; their views of setting and mixed-attainment teaching; and their expectations and practices with young people from different attainment groups^{ix}. It also asked respondents about their academic and teaching qualifications, subject teaching specialism and level of qualification therein, length of time in teaching, position at the school (e.g. classroom teacher, Head of Department), and subject/s which they are presently teaching. All of the above were asked as closed questions, with teachers required to select their responses from a list. These

latter questions built on indicators used in the literature to analyse ‘teacher quality’, with the intention of enabling us to explore any relationship with set allocation. Especially, we wished to answer two particular questions arising within the study:

1. Are teachers who are better qualified, experienced and who have better subject knowledge more likely to be allocated to teach higher performing students in top sets?
2. Is there a difference between the indicative quality of teachers allocated to sets between the intervention and control group, when setting practices were advised to be more equitable in the intervention group?

The control group schools in our study can largely be viewed as indicative of ‘business as usual’ in English secondary schools, in that they were asked to simply continue with their usual practices. However, it is important to note that in order to be recruited for the trial, schools needed to be made aware of the topic and broad requirements of the intervention, prior to randomisation to intervention or control group. As such, these schools had an interest in good practice in setting (hence volunteering as participants in the study), and therefore may indicatively be conscientious schools. Likewise, from the invitation to the trial, control schools would have some idea of what was being asked of intervention schools, and could in theory therefore have chosen to apply some of these practices (such as ensuring fair distribution of teacher expertise across sets) themselves. Nevertheless, the control school respondents are particularly interesting for the purpose of this paper as *broadly* indicative of practice in English secondary schools.

Meanwhile, intervention schools had been asked to implement a range of practices to ensure pupils in all sets had equal access to high quality teaching. The findings from the broader literature reported above led us to include randomisation of teachers to sets as an element of the intervention on ‘Best Practice in Setting’ (Francis et al, 2017a), in order to ensure equitable and even distribution of teacher ‘quality’ to sets within schools. However, it became evident during our study pilot that schools were strongly resistant to randomisation; to the extent that we included an ‘option B’ if schools in the RCT intervention group rejected randomisation. Sure enough, it was notable that almost all schools in the sample chose the option of applying stipulated ‘best practice principles’^x in their own teacher allocation (in spite of the overt study preference for the randomisation method, as the optimal approach for removing bias). Only three of the 64 schools within the ‘Best Practice in Setting’ intervention group subjected their allocation to a process of randomisation conducted by the research team. The questionnaire responses are therefore an opportunity for us to assess to what extent these practices within the intervention group had led to equity in distribution of teaching quality across sets, or otherwise.

Survey Sample

A total of 587 teachers from 81 schools responded to the survey, from a pool of 126 schools. Some of these teachers taught more than one class; therefore their responses were considered for each class that they taught, providing representative responses for 761 classes supplied by 587 teachers. Teacher respondents were recruited by asking for all teachers of English or Maths to the Year 7 pupils in participating departments to respond to the questionnaire. We then (for this analysis) selected the responses of only teachers who taught year 7 and for whom the teaching group could be identified. The total potential pool is unknown. However, 81 schools represent 64% of the 126 schools in our ‘Best Practice in Setting’ sample; 587 respondents providing an average of five teachers per school. In order to answer the research questions of this paper, the teacher survey responses were merged with data provided by students on their set allocation. It was possible to match the responses of 380 teachers to students’ set level. Hence in our final sample for this article there are 380 teachers from 69 schools, 185 from the intervention group and 195 from the control group. Within this sample, there are 262 Maths teachers from 61 schools and 118 English teachers from 35 schools who are matched to the set classes that they teach. The sample characteristics of the teachers who responded to the survey are presented in Table 1. Here and throughout, we present both absolute numbers and percentages of teachers to enable comparison. Note that, throughout, the percentages should be treated with some caution since the numbers in some sub-groups are small.

Table 1. Sample characteristics

Sample Characteristics	Control	Intervention	Total
Gender			
Male	76 (39%)	63 (34%)	139 (37%)
Female	119 (61%)	122 (66%)	241 (63%)
Ethnicity			
Asian	22 (11%)	17 (9%)	39 (10%)
Black	5 (3%)	2 (1%)	7 (2%)
Mixed	2 (1%)	2 (1%)	4 (1%)
White	162 (83%)	162 (88%)	324 (85%)
Other	4 (2%)	2 (1%)	6 (2%)
Age			
Aged 24 or under	29 (15%)	35 (19%)	64 (17%)
Aged 25-39	95 (49%)	99 (54%)	194 (51%)

Aged 40-59	67 (34%)	49 (27%)	116 (31%)
Aged 60 or over	4 (2%)	2 (1%)	6 (2%)
Set allocation			
Top	55 (28%)	50 (27%)	105 (28%)
Middle	103 (53%)	96 (52%)	199 (52%)
Bottom	37 (19%)	39 (21%)	76 (20%)
Whole sample	195	185	380
	(53%)	(47%)	(100%)

Qualitative data collected from interviews and focus groups with students

Qualitative data was collected from pupils at ten different schools in the main study, chosen to represent diversity of location across England, and diversity of school type and practice (see Appendix 1 for the details of schools). Across the 10 schools, a total of 118 students (56 boys and 62 girls) were interviewed in 31 focus groups of between 2 to 6 students. Teachers were asked to use students' English and maths set level to create groups of students with similar experiences of setting. In terms of attainment level, we interviewed 38 top set students (19 girls, 19 boys), 41 middle set students (21 girls, 20 boys), and 39 low set students (17 boys, 22 girls). Socio-economic status and ethnic categorisations were assigned on the basis of student self-report. The higher status occupation between two parents was used to classify students into the following categories: higher SES (n = 19), middle SES (n = 24), low SES (n = 63), and unknown (n = 1). Students belonged to the following ethnic groups: White British (n = 59), Irish (n = 1), Gypsy and Traveller (n = 1), White Other (n = 9), Black African/Caribbean heritage (n = 14), South Asian heritage (n = 19), any other Asian background (n = 1), Arab (n = 4), and any other mixed ethnic backgrounds (n = 10).

Qualitative data Analysis

The data were coded thematically 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.

To ensure the consistency in coding and to enhance the trustworthiness of the process, a coder comparison coding exercise was undertaken mid-way through the coding of the bulk of the data (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 enabled identification of potential problem areas, such as the lower agreement scores between different coders on individual nodes across the document. The coders achieved a high agreement score for the student transcript analysis, reaching as high as 95 percent on most nodes, with disagreement in coding no more than 15 percent on any node. Any identified variations in coding were collectively discussed and resolved.

Findings

The first research question explored the relationship between indicators of ‘teacher quality’, and the set level taught in the sample of schools within the control group, indicative of ‘business as usual’ in English secondary schools. Numbers of sets differ markedly between schools. Therefore, for ease of analysis, we have distinguished between the lowest set in each school, and then included all remaining sets within each school as ‘higher’ sets.

Our analysis against a range of measures of ‘teacher quality’ show inconclusive results, with few clear trends between set allocation and the specified measures (see Appendix 2). There is some suggestion from our descriptive statistics that teachers with senior or middle leadership roles were more likely to teach top sets, but that teachers with higher qualifications i.e. a doctorate, masters or postgraduate diploma or with more than 5 years’ experience appeared equally likely to be allocated to bottom sets than to higher sets in schools. However, variations are slight, and numbers are small, so should be treated with caution.

A clearer indicative trend was found in relation to teachers’ subject qualifications and set placement. As noted above, we recognise the limitations of using qualification as a proxy for quality: nevertheless, this indicator also acts as an indicator of schools’ judgments about teacher quality. In addition, for subject-specific qualifications, we focus on how teachers with a relatively low level of qualification in the subject are deployed (Baumert et al. [2010], for example, indicate subject knowledge to be a foundational first step in developing strong pedagogical content knowledge). We compare teachers at three levels: a basic Level 2 lower secondary school qualification in the subject (GCSE), a Level 3 upper secondary school qualification in the subject (A-level or equivalent), and a higher education qualification in the subject.

Table 2: Highest qualification in subject taught (English or mathematics)

Teacher qualification level	Top and middle sets	Bottom set
Undergraduate or higher degree	99 (65%)	17 (46%)

A level or equivalent	45 (29%)	15 (41%)
GCSE or equivalent	9 (6%)	5 (14%)
Total	153 (100%)	37 (100%)

Although numbers were small, this pattern lends support to prior findings of resources being directed towards higher attainment groups at the expense of the bottom group. Where for higher sets, two-thirds (65%) of teachers had an undergraduate degree or higher in the subject being taught, less than half (46%) of bottom group teachers did. This seems especially striking given maths and English are such key subjects. Bottom sets also had the highest proportion of teachers with only a GCSE qualification in the subject they were teaching. Numbers were particularly small in this latter case, but they may corroborate Sims & Allen's (2017) recent evidence that disadvantaged pupils were significantly more likely to be taught by an unqualified teacher.

Hence, our findings from the descriptive frequencies of set allocation in the control group schools representing 'business as usual' were somewhat mixed and inconclusive, but the finding on allocation in relation to subject qualification corroborates the patterns previously established in the literature concerning deployment of 'high quality' teachers disproportionately to high attainment groups, and indicates that at least one element of such patterns remains in the present. And given that previously established trends for pupils from low socio-economic groups (and certain minority ethnic groups) to be over-represented in low groups were also borne out by our study (see Taylor et al, forthcoming), this illustrates a socially unjust trend in English schooling.

Our second research question sought to explore whether any unjust deployment patterns could be addressed via an intervention designed to secure equity. Our analysis explored the relationship between set allocation and teacher quality when comparing the intervention group, who should have employed equitable practices of setting, and the control group, who were considered to be adopting 'business as usual' practice. As with the data from the control group schools, patterns of allocation in the intervention group against indicative measures of teacher quality were likewise mainly inconclusive. However, given the stronger (unjust) pattern according to qualification in subject taught in allocation to sets within the control group, it was notable that the pattern was less evident in the intervention group, as shown in Table 3:

Table 3: Comparison of allocation of teachers to bottom sets, by highest qualification in subject taught (English or mathematics in intervention and control groups)

Teacher qualification level	Intervention	Control
Undergraduate or higher degree	24 (62%)	17 (46%)
A level or equivalent	10 (26%)	15 (41%)

GCSE or equivalent	5 (13%)	5 (14%)
Total	39 (100%)	37 (100%)

This suggests that, against this measure at least, the intervention group does illustrate more egalitarian practice with greater direction of teacher subject expertise towards bottom sets.

In summary, the quantitative analysis exploring the relationship between teacher quality and set allocation found descriptive evidence of a correlation between teacher subject qualification and deployment to sets, which lends some support to previously-established trends of inequitous practice in teacher assignment to sets. In the descriptive analysis between the intervention and control group, we also found potentially encouraging evidence against one measure, that these trends can be mitigated by an intervention designed to encourage more equitable deployment of teachers. The extent of this reduction did not show differences large enough to constitute statistical significance between the intervention and control groups; albeit this is not surprising given small numbers and a large number of variables. We explore the implications of these findings further in the Discussion section. However, what this indicative quantitative data is unable to address is the quality of pedagogy that pupils experience in the classroom. Our qualitative data offers some insights here.

Pupil perceptions of teacher and pedagogic quality

We did not ask students directly about teacher quality. Asking specific questions on this topic risked precipitating personalised discussions about individual teachers within focus groups, and we were also mindful of not wanting to suggest (or illuminate) inequity for pupils. The closest question was one that asked if they thought different sets receive ‘different types of teaching’. Perhaps not surprisingly given the vague nature of the question, a large majority of focus groups – whether representing high, middle or low sets - said that the type of teaching does not differ according to set. Likewise, very few groups spontaneously volunteered speculation about teacher quality depending on set allocation. However, a few did, and their responses warrant attention for the themes they raise.

Lower-middle set students from School X explicitly discussed distinct teacher quality for different sets, following Sauba’s (Set 3 maths, mixed attainment English, African, low SES) claim that “Some teachers say that Set 1 teachers are, like, the smarter teachers, that they have better teachers; and Set 5, they give to those teachers that don’t really understand.” When a boy in the group initially contested her claim, Sauba provided an illustration from their maths class, whom the group agree is “not even a maths teacher”. Sauba elaborates,

I don't think he's experienced, he would just make us watch 'My Maths'^{xi} or [...] yeah, he would make us watch clips every day, that's it. And people say in Set 5, they just listen to music and they don't do much. Whereas people in Set 1, if you're late you get in trouble. The teachers actually want you to learn, like, they teach you and everything. I feel like Set 1 teachers are more experienced.

Later when Sauba concludes that "I don't think they even bother with Set 5", others in the group concur, albeit Joaquin (Set 3 maths, mixed attainment English, Other White, low SES) protests "But that should not happen because we all need good teachers to learn."

Hence Joaquin expresses the inequity of such perceived discrimination. Sauba also went on to explain how the unchallenging tasks set by this particular teacher made her feel: "He would give us addition and subtraction, which made me feel, kind of like, really stupid. Like, this is [equivalent to] Year 1, like, Reception work. Like, I don't think that Set 4 should be that simple. And that's what made me really angry."

Such views were not exclusive to School X. In a middle set group at School S, Mukthar (Set 3 maths, Set 6 English, Bangladeshi, low SES) explains, "If you're in set 6 you will not be taught that well because you'll be taught, like, simple stuff. But when you're Set 1 you'll be taught like higher stuff so it will really help you." And Asima (Sets 3 maths & English, African, low SES) also observes, "Because, like, I'm not trying to be rude but like when you're in the higher set you get, like, better teachers than the lower set or the middle set, you just get teachers in the middle."

However, it is important to remember these views were articulated by a small minority of students. One middle set pupil at School R, a project intervention school, suggested the contrary view that lower sets got better teachers; noting that in his year group the low maths set is taught by the Head of Maths:

Like, I think they're at, like, the lower set, because if she's the leader of maths, she probably knows a bit more than the other teachers. And that is going to help them get better, if you know what I mean. (Leroy: Set 2 maths, Set 3 English, White and Black African mixed heritage, low SES)

And a top set pupil at School U alluded to the school's timetable rotation that allowed his maths and English teachers to teach both bottom and top sets. These latter examples show that it is possible to 'do differently' and that some schools are managing to find more socially just ways of distributing teachers.

Teacher expectations and related pedagogy

However, while few pupils spoke of teacher quality differing according to set level, many more pupils volunteered opinions concerning differing pedagogic approaches accordingly to set level: for example, concerning various levels of ‘strictness’, and/or the level of expectation and challenge applied to different classes. There were varying views on the benefits of levels of strictness (i.e. whether or not students welcomed it), but there was a strong tendency, irrespective of set level, for students to believe that higher sets had stricter teachers. The following quotes are indicative:

When you go to higher sets, the teachers are strict. (Ahmed, Sets 2 maths & English, Bangladeshi, low SES, School T)

In the higher sets, you get stricter teachers, but in the lower sets you get more, like, laidback teachers that just explain it more. (Andrea, Sets 1 maths & English, White British, low SES, School U).

When you are in a higher set teachers become strict and they just give comments out, even if you do like a small thing. In the lower set, they try more to help you instead of just giving comments out or start[ing] shouting at you. (Mikaal, Sets 5 maths & English, Pakistani, middle SES, School T)

Yeah, I think as you get lower down in sets the teachers get nicer. Then the higher up, they get more strict. (Lyle, Sets 4 maths & English, White British, low SES, School U)

This view was not an exclusive one: for example Cary (Sets 5 maths & English, White British, low SES), a low set student at School R, argued that bottom set teachers are stricter, because “they’re trying to make us like improve on our learning and get up in that middle set or top.” However, this was a notably isolated case in the data, while the view that high set teachers expect or enforce tighter discipline was frequently articulated across different schools and attainment groups. And while not all pupils welcomed this rigorous discipline, it was uniformly seen as an *investment* in pupils. Significantly, when Ahmed (Sets 2 maths & English, Bangladeshi heritage, low SES, School T) opines that in comparison to high set teachers, lower set teachers are ‘not that strict’, Nafisa (Sets 1 maths & English, Pakistani heritage, high SES) agrees, “They don’t really care.” Hence strictness represented engagement and high expectations – and perhaps, by extension, care for pupils.

As some of the above quotes indicate, this perception of more rigorous discipline in the high sets was intimately bound up with perceptions of higher expectations applied to high sets – and, conversely, lower expectations of pupils in low sets. Middle set pupil Mandy (Sets 3 maths &

English, White British, low SES) at School U captures the view expressed by many others when she characterises top sets as “pushed to the limit” while lower sets “get it a little bit easy”. In a middle set focus group at School W, Edith (Sets 3 maths & English, White British, middle SES) affirms that, “the lower down set you are, the easier the work. The higher you are, the harder the work will be”, and her classmate Connor (Sets 3 maths & English, White British, low SES) agrees, “I remember my old teacher when I was in set four used to be a bit more laid back so it wouldn't be... I mean we'd still be working loads but it wouldn't be as challenging at times and then now I've gone into set three it is different.”

These differing expectations were evident in the very different pedagogical approaches frequently reported by pupils (see also Mazenod et al, 2018, for teachers' explanations of these practices). High sets were seen to have more independent learning, and faster paced, more demanding work. Brad (Sets 1 maths & English, White British, low SES, School W) observes, “I think as well that with the higher sets, you're more, like, trusted to do stuff, like go on the learn pads and it's more different ways of learning suited to you”. And Luke (Set 5 maths, Set 3 English, White British, high SES, School P) concurs, “if you're in the top sets they let you like get on with it.” Whereas low sets were seen to be slower paced and scaffolded: Anah (Set 1 maths, mixed attainment English, Arab, low SES, School V) cites a teacher that tends to teach “the lower sets because he's quite slow paced and he likes to do it step by step”; and Aaliyah (Sets 1 maths & English, African, low SES, School V) reports her friend's experience that their lower set teacher is “really slower paced”, in contrast to a high set teacher who is “very strict and wants you to do as best as you can in your lessons”.

It was notable that some low set students appeared highly aware of the different expectations and connected pedagogic practices that were being applied to them (and that this insightfulness belied the low expectations!). For example, in a bottom set focus group at School Q, Ella (Sets 3 maths & English, White British, low SES) explains of different set groups,

We all do, like, the same lessons but some lessons are harder than some. So like, when we were doing poetry [...] we were doing, like, ‘what does the phrase mean?’ Whereas, if you were in a higher set it would be, like, ‘what does the phrase mean and why have they used it in that concept?’

Likewise, a bottom set group at School U:

Sabrina (Sets 4 maths & English, Other White, low SES): It's like different sets do different things. So [...], let's say set six, for example, they would be able to do—instead of doing really hard stuff like learning what implicit information, some of them instead of learning

that, they would probably write, like, a story board or draw a story board. They'll probably, like, draw it. Put the basic connectors and all that that they know.

Interviewer: Is it in English?

Sabrina: Yeah. Then in maths, let's say again set six. They would do really easy stuff, like, let's say, rounding numbers to the nearest ten, for example. That's really easy.

These reports from pupils illustrate the difficult balance between appropriate 'support' for low groups, and the low expectations and consequently impoverished pedagogy that Ireson & Hallam (2001) and others have identified as associated with the teaching of low sets. We have written elsewhere (Mazenod et al, 2018) about a tendency to infantilise low-set students that we found on the part of teachers, and the dangers of 'over support' and low expectations. This tendency was also expressed by students, either in perception of the approaches directed at low sets (i.e. by students in higher sets), or via the reported experiences of low-set pupils themselves.

I think that mainly in the lower set sometimes the teacher can kind of feel like a mum when you're a baby, because a lot of the time when you're, like, a toddler or a lot younger the mum would start, like, talking in a very, like, way that, they'd be like, "Oh, are you okay?" and everything. But like, in the higher set they will be quite strict. (Jackie: Sets 2 maths & English, White British, low SES, School Q)

Bobby (Set 5 maths, Set 3 English, White British, middle SES, School P) observes, "In lower set I think they talk slow and they go through the questions like a second time and they help you more." He goes on to explain that he doesn't like maths because "sometimes like they're helping you even if you want to do it by yourself", and;

Because sometimes they just treat you like you're little. They have to like speak slowly and then they start like speaking slowly and using hand gestures and stuff. I know that some people like need that but like sometimes it gets annoying.

Hence these pedagogical approaches were experienced as patronising and demeaning by lower set students. The trope of infantilisation (being treated like 'babies') manifested frequently in the data from these pupils:

Yes, I think that, like, the people in the lower sets, they get taught like easier methods. And like the teacher might talk, like, babyish to them to make sure they understand. (Charlene: Set 3 maths, Set 2 English, White British, low SES, School S)

Yeah. One of my friends came out of a class, one of the people in the lower set, quite upset because she felt like the teacher was kind of babying the class. So I feel that if you are in the

lower set than the teachers won't put as much pressure on you. And then when you're in the middle they just treat you normally. (Jackie: Sets 2 maths & English, White British, low SES, School Q)

A boy from the bottom set at School U who later withdrew himself from the study likewise complained that teachers made them feel stupid. His classmate Helen (Sets 5 maths & English, White British, middle SES) elaborated, "so in English you need to do one paragraph, and we do one paragraph and then the teacher just [...] like 'oh carry on now' where we've already done. I'm just like 'but we've already done it the stuff we are carrying on'.

These quotes illustrate both frustration and bitterness generated among pupils in response to slow-paced and insufficiently demanding pedagogy which they experience as patronising and inadequate. Some students appeared aware that through this 'spoon-fed' approach they were losing out on important life skills. For example, a bottom set focus group at School Y complained at the lack of independent learning they experienced, and would like to work more independently:

Yaday: Yeah because it's a good skill for you because you have to work...

Umaima: You don't have to depend on someone.

And Graham (bottom set maths and English, White British, low SES, School W) critically observes of his teacher, "Miss B – she does it more fun, but we need to do the learning as well because when it comes to actually doing a test, we don't get it." Likewise, Charlene (Set 3 maths, Set 2 English, White British, low SES, School S) shows an awareness of the extending benefits of high expectations which she feels low groups miss out on:

They [teachers] talk easier, however in the higher sets they talk with long words. Words that like children don't really understand but they get to know; but then in the lower sets they don't speak like that, they speak the easier words.

Hence these articulations emphasised the ways in which this low expectation pedagogy was serving to reinforce inequalities by setting a cap on learning and limiting the development of skills that students in higher groups benefitted from, and which those in low groups arguably especially need.

Discussion

Our data analysis provides a contemporary update that lends some limited corroboration of trends found in prior quantitative studies, where more highly qualified and expert teachers tend to be designated to higher attainment groups. While the data against the majority of measures was inconclusive, there was some descriptive statistical evidence that supported

this hypothesis, especially on the measure of qualification in subject taught. We recognise that qualifications are a relatively weak proxy for teacher quality (Shulman, 1986): indeed the existing measures of ‘teacher quality’ remain lacking in conceptual and empirical validity, and more work is needed to provide robust analytical tools in this regard. Nevertheless, the indicator of qualification in subject taught acts as an indicator of schools’ judgments about teacher quality, as well as about teachers’ level of expertise in their curriculum subject area. Any trends in this direction raise issues for social justice, given that – as with so many other studies – our study findings are that pupils from socially-disadvantaged groups are under-represented in top sets, and over-represented in low attainment groups (Taylor et al, 2018). Hence the very pupils that would most benefit from high quality teaching (Ainscow et al, 2012), and whom most need compensatory input from the school system, may be less likely to receive high quality subject expertise and pedagogy.

We have shown how trends for inequality in teaching practice deployed with different attainment groups was also strongly reflected in young people’s perceptions of pedagogy according to set level. There was an overwhelming view among pupils that expectations were higher for top set groups. Distillation of the young people’s responses reveals a view that pedagogy directed at higher set groups is characterised by:

- rigorous expectations of discipline,
- ‘pushing’ pupils to do their best,
- respect conveyed by the provision of independent learning opportunities.

By contrast, pedagogy for low sets was widely perceived to be:

- more tolerant and relaxed,
- ‘spoon-fed’, with less opportunities for independent study and skill development
- slow-paced and less demanding

(the latter as documented by Hallam & Ireson, 2005; Boaler et al., 2000; Oakes 1985).

Coupled with our findings on teacher subject qualifications and deployment, pupils’ articulated experiences suggest strongly inequitable practices, and classroom observational research is required to investigate these pupil perceptions further.

A further significant finding, also revealed from our qualitative work with pupils, is the frustration experienced by low set pupils in relation to this. It is worth noting that throughout our study we have found that, due to the high valuing of respect by pupils, pupils are especially sensitive, and resistant to, being patronised. Segregation by setting already hierarchises pupils and renders low set pupils vulnerable in regard to currencies of respect.

The infantilisation felt by students in low sets (and/or projected on to them by pupils in higher sets and by teachers) has been a strong point of complaint throughout our data (Francis et al, 2017b; see also Mazenod et al, 2018).

Our findings indicating some inequity in deployment of teachers and the likely impact of this inequality in access to ‘quality’ teaching on pupil outcomes, coupled with the fact that socially-disadvantaged pupils are over-represented in low set groups (Taylor et al, forthcoming), highlights socially unjust practice. Pupils also feel that set group impacts the quality of pedagogy and teacher expectation applied to set groups. Taken separately or together, these findings may be regarded as a potential contributory factor in the persistent extent of the socio-economic gap for educational attainment in the English education system. Given what we know about the importance of respect for teenagers (e.g. Archer et al, 2010), and our findings on their experiences of pedagogy in low sets, it is also no surprise to find that setting impacts on young people’s liking for school, with a detrimental effect for those in low groups (Archer et al, 2018).

In our introduction we presented some of the theoretical and pragmatic potential explanations for iniquitous patterns. In analysing the findings, we find ourselves less concerned with identifying why these practices occur, than the importance of halting them. To that end, it is potentially encouraging that in regard to the measure wherein we found indication of inequity in teacher deployment within the control group, the Best Practice in Setting intervention appears to have shown indication of with greater direction of teacher subject expertise towards bottom sets. If schools are determined to continue application of setting, adoption of these principles may be important. Nevertheless, given schools’ resistance to randomisation of teacher allocation, our findings from the intervention highlight both the need for greater fidelity in intervention application to facilitate equity – and the strength of the risk that this greater fidelity will not occur. The pupil interview data also highlights the general dangers of setting for pupil and teacher self/perceptions and self-fulfilling prophecy (see also Francis et al, 2017b; Mazenod et al 2018), which are inherent to setting practice. These findings therefore demonstrate the urgency of action – for teachers and school leaders, in professional reflection on practice in schools; and for researchers, in refining measures of analysis in this areas, and continuing exploration of optimal mixed attainment practice as an alternative to segregation by attainment.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago Public Schools. *Journal of Labor Economics*, 25, 95–135.

- Ainscow, M., Dyson, A., Goldrick, S., & West, M. (2012). *Developing equitable education systems*. London: Routledge.
- Allen, R., Mian, E., & Sims, S. (2016). Social inequalities in access to teachers. Social Market Foundation Commission on Inequality in Education: Briefing 2.
<http://www.smf.co.uk/wp-content/uploads/2016/04/Social-Market-Foundation-Social-inequalities-in-access-to-teachers-Embargoed-0001-280416.pdf> Accessed 7.1.18
- Archer, L. , Hollingworth, S. & Mendick, H. (2010). *Urban Youth and Schooling*. Maidenhead: Open University Press.
- Archer, L., Francis, B., Miller, S.J., Taylor, B., Tereshchenko, A., Connolly, P., Pepper, D. & Travers, M-C. (2018). The symbolic violence of setting: A Bourdieusian analysis of mixed methods data on secondary students' views about setting. *British Educational Research Journal*. 44:1, 119–140 <https://onlinelibrary.wiley.com/doi/abs/10.1002/berj.3321>
- Argys, L. M., Rees, D. I., & Brewer, D. J. (1996). Detracking America's Schools: Equity at Zero Cost? *Journal of Policy Analysis and Management*. 15(4), 623-645.
- Ball, S.J. (2003). *Class strategies and the education market: the middle class and social advantage*. London: RoutledgeFalmer.
- Bazeley, P., & Jackson, K. (2013). *Qualitative Data Analysis with Nvivo*. London: Sage.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5:1, 7-74.
- Boaler, J., Wiliam, D., & Brown, M. (2000). Students' experiences of ability grouping-disaffection, polarisation and the construction of failure. *British Educational Research Journal*, 26(5), 631-648.
- Bourdieu, P., & Passeron, J-C. (1977). *Reproduction in education, society and culture*. London: Sage Publications.
- Boyd, D., Lankford, H., L., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management*, 27(4), 793-818.
- Clotfelter, C., Ladd, H., & Vigdor, J. (2006). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of Human Resources*, XLI(4), 778-820.

Coe, R., Aloiso, C., Higgins, S., & Elliot Major, L. (2014). *What Makes Great Teaching?: Review of the underpinning research*. Retrieved from the Sutton Trust website:
<https://www.suttontrust.com/research-paper/great-teaching/>

Connolly, P., Taylor, B., Francis, B. et al (forthcoming) The Misallocation of Students to Academic Sets in Maths and English: A Study of Secondary Schools in England, under review at *British Educational Research Journal*.

Dunne, M., Humphreys, J., Sebba J., Dyson, A., Gallanmaugh, F., & Muijs, D. (2007). *Effective teaching and learning for pupils in low attaining groups*. London: DCSF.

Finley, M.K. (1984). Teachers and tracking in a comprehensive high school. *Sociology of Education*, 57, 233-243.

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., Mazonod, A., Pepper, D. ... Travers, M-C. (2017b). 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.

Gamoran, A. (1992). Synthesis of research: Is ability grouping equitable? *Educational Leadership*, 50(2), 11-17.

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

Good, T., & Marshall, S. (1984). Do students learn more in heterogeneous or homogeneous groups? In P. Peterson & L. Cherry-Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes*. New York: Academic Press.

Hallam S., & Ireson J. (2005). Secondary School Teachers' pedagogic practices when teaching mixed and structured ability classes. *Research Papers in Education*, 20(1), 3-24.

Hallinan, M. (1984). Summary and implications. In P. Peterson & L. Cherry-Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes*. New York: Academic Press.

Peterson, P., L. Wilkinson, & Maureen T. Hallinan (Eds.). (1984). *The Social Context of Instruction: Group Organization and Group Processes*. San Diego, CA: Academic Press.

- Hanushek, E., & Rivkin, S. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2), 267-271.
- Harris, D., & Sass, T. (2006). *Value-added models and the measurement of teacher quality*. Unpublished manuscript.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.610.1061&rep=rep1&type=pdf>
[accessed 1.1.18]
- Hattie, J. (2013). *Visible Learning*. London: Routledge.
- Higgins, S., Katsipatakis, M., Coleman, R., Henderson, P., Major, L. E., Coe, R., & Mason, D. 2015. *The Sutton Trust - Education Endowment Foundation Teaching and Learning Toolkit*. London: Education Endowment Foundation.
- Husbands, C. (2012, September 6). *Teaching talent: raising quality in education*. Speech at Social Market Foundation, London.
- Illich, I. (1971). *De-schooling society*. London: Calder & Boyars Ltd.
- Ireson J., & Hallam S. (2001). *Ability Grouping in Education*. London: Paul Chapman Publishing.
- 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.
- Kane, T., Rockoff, J., & Staiger, D. (2007). What does certification tell us about teacher effectiveness? *Economics of Education Review*, 27, 615-631.
- Kelly, S. (2004). Are teachers tracked? On what basis and with what consequences? *Social Psychology of Education*, 7(1), 55-72.
- Kulik, J. A. & Kulik C.-L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly* 36(2), 73-77.
- Kutnick, P., Sebba, J., Blatchford, P., Galton, M., Thorpe, J., with MacIntyre, H., Berdondini, L. (2005). *The effects of pupil grouping: Literature review*. London: DfES.

Kohn, A. (1998). Only for *My Kid*: How Privileged Parents Undermine School Reform. *Phi Delta Kappan*, 79(8), 569-577.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24, 37–62.

Anna Mazenod, Becky Francis, Louise Archer, Jeremy Hodgen, Becky Taylor, Antonina Tereshchenko & David Pepper (2018): Nurturing learning or encouraging dependency? Teacher constructions of students in lower attainment groups in English secondary schools, *Cambridge Journal of Education*, DOI: 10.1080/0305764X.2018.1441372

Marshall, P. (Ed.) (2013). *The Tail: How England's schools fail one child in five - and what can be done*. London: Profile Books.

Murphy, J. & Hallinger, P. (1989). Equity as access to learning: curricular and instructional treatment differences. *Journal of Curriculum Studies*, 21:2, 129-149.

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

Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal*, 32, 3-33.

Old, A. (2017, November 11). Why all the research on teacher qualifications is worthless. [Web log post]. Retrieved from <https://teachingbattleground.wordpress.com/2017/11/11/why-all-the-research-on-teacher-qualifications-is-worthless/> Accessed 1.1.18

ONS. (2010). ONS Standard Occupational Classification (SOC) Hierarchy. Retrieved from http://www.neighbourhood.statistics.gov.uk/HTMLDocs/dev3/ONS_SOC_hierarchy_view.html Accessed 21.3.17

Papay, J., & Kraft, M. (2014). Productivity returns to experience in the teacher labour market: methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*, 130, 105-119.

Reay, D., Crozier, G. and James, D. (2011). *White Middle Class Identities and Urban Schooling*. London: Palgrave Macmillan.

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1), 43-74.

Sanders, W. L., & Rivers, J. C. (1996). *Research project report: Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Sims, S. (2017, September 2017). Could we get the best teachers into the most deprived schools? [Web log post]. Retrieved from <https://ioelondonblog.wordpress.com/2017/09/19/could-we-get-the-best-teachers-into-the-most-deprived-schools/> Accessed 1.1.18

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.

Suk Wai Wong, M., & Watkins, D. (2001). Self-esteem and ability grouping: A Hong Kong investigation of the Big Fish Little Pond Effect. *Educational Psychology*, 21(1), 79-87.

Sukhnandan, L., & Lee, B. (1999). *Streaming, setting and grouping by ability*. Slough: NFER.

Sutton Trust (2011). *Improving the impact of teachers on pupil achievement in the UK – interim findings*. London: Author. Retrieved from <http://www.suttontrust.com/wp-content/uploads/2011/09/2teachers-impact-report-final.pdf>

Becky Taylor, Becky Francis, Louise Archer, Jeremy Hodgen, David Pepper, Antonina Tereshchenko & Mary-Claire Travers (2017) Factors deterring schools from mixed attainment teaching practice, *Pedagogy, Culture & Society*, 25:3, 327-345, DOI: 10.1080/14681366.2016.1256908

Waldfogel, J., & Washbrook, E. (2010). *Low income and early cognitive development in the U.K.* London: The Sutton Trust.

William, 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.

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, 100, 61-78.

Appendices

Appendix 1. Details on schools in which pupil focus groups were conducted

School	Trial group	Location	Ofsted rating	Free School Meals	Number of set levels
Q	Intervention	City, South coast	Good	16%	4
R	Intervention	Market town, North west England	Good	12%	4
T	Intervention	City, Yorkshire	Requires improvement	24%	5
Y	Intervention	Outer London	Outstanding	22%	3 attainment bands, 3-4 sets within the bands
X	Intervention	Outer London	Good	17%	5 – mathematics, Mixed attainment – English
P	Control	Market town, Midlands	Outstanding	5%	6 – mathematics 3 – English
S	Control	Outer London	Good	21%	6
U	Control	Rural, South East England	Requires improvement	28%	5
V	Control	Outer London	Good	21%	7 – mathematics Mixed attainment – English
W	Control	Semi-rural, Midlands	Good	7%	6

Appendix 2. Exploring teacher quality across sets in the control group

Measure of Teacher Quality	Top or Middle Set	Bottom Set
Position in school		
Head or Deputy Head of Department	23 (18%)	5 (18%)
Teacher	80 (63%)	19 (68%)
Newly-Qualified or Unqualified Teacher	23 (18%)	4 (14%)
<i>Total</i>	<i>126 (100%)</i>	<i>28 (100%)</i>
Highest qualification in subject taught (English or mathematics)		
Undergraduate or higher degree	99 (65%)	17 (46%)
A Level or equivalent	45 (29%)	15 (41%)
GCSE or equivalent	9 (6%)	5 (14%)
<i>Total</i>	<i>153 (100%)</i>	<i>37 (100%)</i>

Highest Qualification (any subject)

Postgraduate diploma or higher degree	44 (29%)	8 (24%)
Undergraduate degree or equivalent	107 (71%)	26 (76%)
<i>Total</i>	<i>151 (100%)</i>	<i>34 (100%)</i>

Length of time in teaching

6 years or more	81 (51%)	22 (59%)
5 years or less	77 (49%)	15 (41%)
<i>Total</i>	<i>158 (100%)</i>	<i>37 (100%)</i>

Note: Missing responses excluded from these analyses. Members of the senior leadership team, other middle managers and trainee teachers excluded from position in school analysis. "Other" qualifications excluded from highest qualification (any subject) analysis.

ⁱ Segregation by attainment includes many different practices, many of which may overlap or be practiced in tandem. Between school segregation occurs where different types of school are provided for pupils deemed to have different ‘abilities’. Within-school segregation includes practices such as: tracking (or ‘streaming’, as it is referred to in the UK), wherein students are placed in the same ‘ability’ group for most or all lessons; setting (common in the UK), wherein pupils are placed in attainment groups for particular subjects; and ‘ability tables’, where pupils are separated within class according to attainment (commonly practiced in primary schools in the UK).

ⁱⁱ While (often dated) meta-analyses tend to find no significant overall effect of attainment grouping (see Education Endowment Foundation, 2016), some studies have found a small advantage from attainment grouping in outcomes for students in higher attainment groups (see for example: Argys, Rees & Brewer, 1996; Kulik & Kulik, 1992; Steenbergen-Hu et al, 2017). Meta-analyses suggest that grouping by attainment has no significant impact overall, conferring a slight advantage to high attainers simultaneous to a more significant disadvantage for the (small group) of low attainers (e.g. Slavin, 1990; EEF, 2016). Nevertheless, it is also worth pointing out that the research remains complicated by a range of factors, and occasionally contradictory.

ⁱⁱⁱ We do not ascribe to a view of ‘ability’ as fixed, hence our adoption of inverted commas.

^{iv} ‘GCSE’ stands for General Certificate of Secondary Education; the exams taken by all pupils in England at age 16.

^v Ofsted is the schools inspectorate organisation for England.

^{vi} The GCSE grading system is now undergoing change, with letter grades being replaced by number grades. Nevertheless, the notion of a ‘pass’ grade remains, meaning that they issue has not been removed. Focus on the progress of all pupils may, however, be levered by the recent application of ‘Progress 8’ measures, which reward schools for the relative progress made by each pupil, rather than (only) raw attainment outcomes.

^{vii} See endnote i). Setting is sometimes applied only to certain subjects; sometimes to all/almost all; and is sometimes applied in conjunction with streaming.

^{viii} NFER are commissioned by the Education Endowment Foundation to perform the post-testing and to evaluate our key intervention outcomes.

^{ix} The large majority of teachers in both control and intervention school groups qualified via the PGCE route, albeit there is a difference here between the intervention and control groups, with the control group reflecting slightly more diverse spread in qualification routes followed. The patterns for length of time in teaching are very closely matched between the two groups, with the greatest numbers having taught between 3-5 and 6-10 years.

^x The ‘Best practice principles’ for teacher allocation were: 1) Newly-qualified teachers and non-subject specialist teachers should not be generally allocated to the lowest set level. 2) Teachers should be allocated to the sets with a view to ensuring an equitable distribution both across year groups and year-on-year; and 3) Teachers should not teach the same set in the Year 8 in the 2015/16 academic year as in year 7 in the 2014/15 academic year.

^{xi} ‘My Maths’ is an interactive maths teaching website, for which schools pay a subscription.