Tracking and mixed-ability grouping in secondary school mathematics classrooms: a case study[1]

Hamsa Venkatakrishnan, Dylan Wiliam

School of Education, King's College London

Abstract

This paper reports the findings of a retrospective study of 'tracked' grouping in a mathematics department in a co-educational comprehensive school in Greater London. Tracking consisted here of just two tracks, a 'fast track' for the top 25-30% of a cohort, and 'mixed tracks' for the remainder. The paper outlines the reasons for introducing tracking and explores the effects of this through analysis of interviews with teachers and data on the progress of students from age 14 to age 16. The teachers reported that tracking impacted differently on different students, and this is borne out by the quantitative data. It was not possible to provide for 'setting' across all the mathematics classrooms in the focal cohort, and one mixed-ability class was created. The use of analysis of covariance (ANCOVA) models shows that fast-track students were not significantly advantaged by their placement in these tracks, but the progress of students in the mixed-ability group showed a significant interaction between progress and prior attainment, with placement in the mixed-ability group conferring a significant advantage on lower-attaining students, while the disadvantage to higher attaining students was much smaller.

Introduction

This paper reports the results of a study of the changes in grouping arrangements that were imposed on a mathematics department in a co-educational 11 - 16 comprehensive in Greater London. It begins by briefly describing the school's reasons for introducing this change, which consisted essentially of a shift away from mixed-ability groupings to a 'tracked' system[2]. The main focus of the paper is within the Mathematics Department. There are two primary sources of evidence here. Firstly the views of teachers are traced, from their initial concerns on the possible effects of a change in grouping, their responses as to how the changes ought to be structured and carried through, and their perceptions of the impact of the new 'tracked' arrangement on their classrooms. The second source consists of the examination scores of a cohort of students going through this change. Key Stage 3 (KS3) Mathematics tests scores (taken in May of Year 9, with cohort aged 13/14) provide an indicator of their attainment prior to the change in grouping, and GCSE Mathematics exam scores (taken in May/June of Year 11, with cohort aged 15/16) provide this indicator after regrouping.

With this twin focus on classroom processes and examination performance, this study aims to overcome some of the limitations of earlier waves of research on grouping. These tended on one side to focus either predominantly on equity considerations within setted / streamed environments (Hargreaves, 1967; Lacey, 1970; Ball, 1981; Abraham, 1989), or on the other, to measure in quantitative terms the differences in test performance of students assigned to different sets or 'tracks' (Oakes, 1985; Natriello *et al*, 1989; Hoffer, 1992). This research borrows some of the methods of more recent and emerging research on setting (Boaler, 1997; Boaler, Wiliam and Brown, 2000). These use a combination of descriptive and numerical methods to try and build a more detailed picture of the differences in classroom experiences of students in different sets, and relate these to pupil attainment.

Data from open interviews and questionnaires is used to describe the views of teachers within the Mathematics department. Their attitudes initially to the grouping change, their descriptions of the differences in working cultures between groups in the banded structure, and the differences too, in their own classroom organisation and practices are considered. Their perceptions of the effects of tracking on individual and particular groups or types of students form a part of this section. The primary aim here is to try and identify whether there is evidence, as earlier research suggests (Oakes, 1990b; Boaler, 1997) that tracking produces systematic differences in provision for students. A secondary aim, within a department that raised the issue of differential provision as one of their concerns, is to identify the steps that were taken to address this. These aims are achieved by allowing teachers to describe the effects of tracking on their role within the classroom, and the ways in which this role changes depending on the track they are teaching. Boaler's (1997) qualitative evidence from observations of setted classrooms suggests that the practices adopted by teachers in this environment may relate quite significantly to student performance. The focus here though, is not confined to looking for the common features of setting or tracking in terms of classroom practices, but to look also at the impact of differences in these, on the pattern of results obtained.

Quantitative methods are used to investigate the impact of the change in grouping structure on the examination results of one cohort of students. The Year group that is chosen for analysis studied Mathematics in mixed ability tutor groups throughout Key Stage 3 (Years 7, 8 and 9), and were rearranged into the new tracked set-up for Key Stage 4 (Years 10 and 11). Their Key Stage 3 test results then, provide an indicator of attainment under mixed ability grouping, and GCSE results provide this indicator under tracking. The aim here is to see whether there is evidence of patterns of differential performance between students in different tracks and for students from different backgrounds. The creation of a mixed-ability group allows also for comparison between tracked and mixed ability progress. A diverse range of earlier research (Smith & Tomlinson, 1989; Boaler, 1997; Gaine & George, 1999) suggests the relevance of data on the gender, class and ethnic backgrounds of these students as useful starting points for this analysis. A complex timetable pattern that splits the Year group into three blocks (described in more detail later) allows further, for a quantitative investigation of differences within the overall tracking structure between the blocks. These results are set within the context of previous research on ability grouping in Mathematics.

Within this study, the context within which ability grouping was imposed had a marked effect on the structure that was introduced by the department. In addition to this, individual teachers constructed their own interpretations about operational practices that led to differences in their ways of dealing with aspects of this structure. These practices are likely to have had effects on students, which are more easily explored through an analysis that integrates the context—school and classroom, with examination results.

The paper begins with a brief setting of the scene. Short descriptions of the school, the Mathematics department, the cohort being studied, and details of the tracking system itself and the way it was timetabled for this year group are given. This is followed by details of the sample used within this investigation.

The school

The school is an 11-16 co-educational comprehensive school in Greater London, in an LEA where approximately half the secondary schools 'opted out' of local authority control to become Grant Maintained Schools in the early 1990s. Its immediate catchment area includes a large council estate that has been described by the Borough as an area of "multiple deprivation". The majority of the school's students are working class but the roll (of approximately 1200) also includes students from the more affluent surrounding areas. The intake of the school is unrepresentative of the national population in several respects as can be seen from the data shown in table 1 which refers to 1997, the mid-point of the period of investigation for the cohort under study in the present project.

	focal cohort	school	nation
male	55%	64%	50%
entitled to free school meals	35%	66%	16%
having special educational needs		48%	18%
from minority ethnic communities	50%	50%	7%

Table 1: characteristics of the school

The mathematics department has used the SMILE[3] scheme—an individualised learning scheme—with all year groups since 1991, supplementing this with some whole class teaching. Between 1991 and 1994, all mathematics classes were taught in mixed-ability groups - in tutor groups in years 7, 8 and 9, and in mixed-tutor groups in years 10 and 11. In September 1995, the department introduced a system of 'tracking' in which the highest-achieving 25-30% of the students in year 8, in terms of mathematics exam scores, were taken out of their existing classes to create two 'fast-track' classes (one in each of the two blocks to which the tutor-groups were allocated). Later in the same year, tracking was extended to year 7, but in the following year, year 7 reverted to mixed-ability grouping, year 8 remained 'tracked' and tracking was extended to years 9 and 10.

In 1996, at the start of their year 10, the focal cohort formed the second 'wave' of students to move to the new 'tracking' system. This structure remained in place until the February half-term of their Year 11, (4.5 of their 5 terms), after which they were regrouped into Higher, Intermediate and Foundation groups to facilitate GCSE revision at the appropriate tier of entry. The arrangement of tracking for this cohort though, had to take into account options choices, and an unusual 'pathways' system that incorporated different routes through years 10 and 11. These interweaved, to varying degrees, an academic (GCSE-based) curriculum with a vocational curriculum, which led either to the more basic Diploma of Vocational Education (DVE) or the intermediate General National Vocational Qualification (GNVQ). This resulted

in the creation of three blocks: Block 1 included all the students who had opted to follow any DVE options, along with some students enrolled on the other routes; Block 2 contained only students following GNVQ and/or GCSE courses; Block 3 consisted of just one teaching group, made up of all the students who had taken Media Studies as an option. Block 3 too, included only GNVQ and GCSE student. Students were taught within these blocks for all their mathematics lessons, and for humanities. In mathematics, blocks 1 and 2 each contained one fast track group, and respectively, four and three mixed-track groups. Block 3 students were taught mathematics as a mixed ability group throughout years 10 and 11.

The net effect of this programme and the blocking structure for the Mathematics department is that the three blocks vary quite considerably in terms of their attainment profiles, measured in terms of mathematics key stage 3 test scores[4] as is shown in Figure 1 (the shaded regions in each box represent 95% confidence intervals for the median of each batch—non-overlapping shaded regions indicate statistically significant differences in medians at the 95% level).

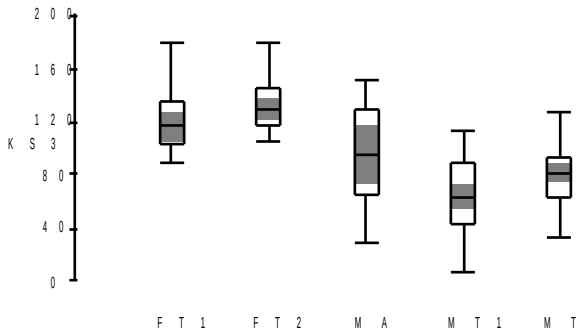

Figure 1: equated key stage 3 test scores for fast-track, mixed-ability and mixed-track groups

Clearly this mode of 'blocking' students makes for a far from straightforward analysis of the impact of tracking, with the variation in composition of students within the blocks. Details of the sample below (Table 2) elaborate on these variations, identifying the ways in which the blocking arrangement interacts with student background factors. The structure of blocking is kept within the scope of this investigation, primarily because omitting it is likely to mask the degree of localisation of any patterns of differences between tracks. In addition though, in a study which explicitly aims to integrate the context of tracking with its results, the blocking arrangement reflects the complexity of the situation that the school has to deal with in terms of an intake which is significantly skewed, but still contains the full ability range. The analysis of students' results therefore, whilst identifying overall the 'winners' and 'losers' under tracking, also looks at contrasts between the performance of students in different blocks, given their different attainment profiles (B1 v. B2 & 3) and grouping arrangements (B1 & 2 tracked v. B3 mixed ability).

The focal year-group for this study consisted of a total of 240 students. Some of these students arrived at the school during years 10 and 11, and since KS3 results were unavailable for the majority of them (and where they were available, could not be taken as an indicator of attainment prior to a *change* in grouping), they were removed from the data. GCSE scores were not available for some students (the vast majority of whom were those students who had left the school at Easter), and these were also removed, leaving a total of 180 students for whom complete data were available. A standardised mathematics KS3 test score was calculated for each student by equating the four tiers of the KS3 mathematics test. The mean score for those whom GCSE scores were available was 89.7 and for the others, it was 62.6, showing that the students who did not take GCSE cannot be considered representative of those who did. Since the vast majority of the excluded students were in the mixed tracks, any significant differences found in achievement, either between tracks or between blocks is likely to understate, rather than overstate the real situation.

Block	Sex		Ethnicity*							Total	FSM
	F	M	AC	AFR	AOG	ASI	CHI	MR	WHI		
1	30	47	14	5	5	14	1	2	36	77	45%
2	44	40	15	1	7	15	0	2	44	84	20%
3	11	8	1	1	6	3	0	1	7	19	37%
	85	95	30	7	18	32	1	5	87	180	32%

*AC: African Caribbean; AFR: African; AOG: Any other group; ASI: Asian; CHI: Chinese; MR: Mixed race [one 'white' parent]; WHI: White

Table 2: characteristics of focal year group for who matched data are available

The background to tracking

The decision to impose some form of banding (the term used by the school for ability grouping in a subject-specific sense) in Mathematics followed two years of deepening concerns at Senior Management level about the increasing proportions of pupils with special educational needs (SEN) entering the school. Problems at the curricular level were compounded by cuts in funding for SEN provision. At the same time, the school was receiving rising numbers of letters from parents, largely, but not exclusively, middle-class parents, expressing concerns about the levels of disruption their children were facing in some classrooms in a range of subject areas. Some of these children were the high attaining pupils that research on education markets (Whitty *et al*, 1998) has shown to have become increasingly valuable to schools within the league table culture (Ball *et al*, 1994).

Interviews with members of the Senior Management team and documentation from that period reveal the conflicting demands of how best to balance the needs of a skewed intake with the needs of higher attaining pupils. The decision to impose some form of banding in Mathematics, was, they admitted, fuelled primarily by the letters from a small minority of parents:

> Deputy Head: " It was where we were getting the most unhappiness from the parental point of view."

Perceptions from parents and individuals within the SMT that setting was the most 'natural' and 'effective' context in which to teach Maths also played a part in this decision:

> Senior Teacher: "There was a perception from parents, and from research and the wider view—Ofsted, for example—that Maths was usually streamed."

Tracking began in 1995 with years 7 and 8 since this was where the student cohorts were most skewed, and also where a significant number of parental concerns were being voiced.

The mathematics department's response

Following the instruction, quite late in the Summer Term of 1995, that some form of banding needed to be introduced, several general concerns about banding /setting were raised within the department. There was widespread concern at the negative labelling that could be induced by banding, and particularly at the impact that this could have within the classroom in terms of the motivation and behaviour of students. Issues surrounding the possible effects on students at the 'cut-off' points in a banded environment were also raised. Particular emphasis was placed on the effects of this coinciding with students on the C/D borderline, with the possibility of creating unnecessary anxiety for them at the 'bottom' of a 'top' set, and demotivating them if they missed out on this and were placed in a lower group. Empirical research into students' responses to setting (Boaler, 1997; Boaler *et al*, 2000), suggest that these issues are 'real', and not just negative perceptions of setting from teachers. Concerns were also expressed that the skew in the ability profile of the Lower School, and in years 7 and 8 in particular, made any form of banding difficult, since the number of higher-ability students was relatively small. In making the 'top' sets large in order to alleviate strains on lower sets, the ability range within them was likely to be wide—the attainment at the start of year 8 in a 'top' set was likely to be anywhere between National Curriculum Levels 4 and 7 (a range that would include over half the population). This, in turn, the department argued, was likely to preclude the prospect of large scale whole-class teaching (if this was what parents and the school were expecting), and also to produce lower sets consisting entirely of students working at or below the national average. The organisational problems of setting such as ensuring that each class had sufficient numbers of SMILE cards and other resources to deal with the reduced ability range in some year groups was also stated.

The system of 'tracking' proposed, with just two tracks—a 'fast' and a 'mixed' track—together with the decision to continue using SMILE across *all* classes, alleviated some of these concerns. It involved a minimisation of the 'amount' of banding that was to occur, and was likely to be the most akin to a fully mixed-ability context. It facilitated too, the continued use of SMILE, the introduction of which had coincided with substantial improvements in the school's GCSE results in Mathematics. The top 25% or so of students in each year half who were 'creamed off' into the fast track also contained the majority of students with parents who had voiced concerns about their learning. Tracking therefore provided a 'visible' response to their criticisms.

In the interests of trying to maintain equitable provision across the ability range, and minimise what were perceived as the negative effects of setting, a decision was taken to make the 'fast' tracks large, and thereby allow greater individual attention and alleviate behaviour concerns in the 'mixed' tracks. Also the 'mixed' track groups in block 1, (which contained all the DVE students and therefore had a more negatively skewed attainment profile), were generally smaller than the mixed track groups in block 2. Additionally, it was agreed that students' progress in the two tracks would be monitored regularly and kept fluid in terms of movement between them.

The department therefore, with their choice of 'tracking', tried to produce a compromise structure that fitted the constraints they faced. Tracking, at this point (1995) was still relatively rare, and where it existed, tended to consist of 'enrichment' or 'express' type pathways for the most able. This was not the function of the fast track in this context—the consideration was more to give higher-attaining students the opportunity to interact more constructively than was possible in mixed ability groups where their numbers were limited. This aim, while not restricted to fast-track groups, finds considerable support in both theoretical and empirical studies (Cobb & Bauersfeld, 1995). The decision to continue using SMILE with both tracks, though partly forced by financial considerations, also remained one that was at odds with the wishes of some quarters within the Senior Management Team and the parental lobby.

Tracking in operation

Three teachers in the department each taught two groups within the cohort being considered. Two of them, Karan Davda and Janet Dyner, taught a 'fast-track' group in one block and a 'mixed-track' group in the other block. The third teacher, Enya Lama, taught two mixed-track groups (one in each of the tracked blocks). Their views of how tracking impacted on the classroom, collected through open questionnaires and interviews are presented below.

These three teachers were chosen for a number of reasons. Firstly, they were all at the school during the period in which the introduction of banding was discussed and were therefore aware of the policy context involved. Secondly the fact that they each taught two groups allowed them to make comparisons between groups in terms of the progress, attitudes, working cultures and atmosphere within these classrooms in ways that relate to the tracking and blocking structures.

Although the introduction of tracking was imposed on the department, how it was to be operated was largely left to the department to determine. This point is important to note because all three teachers interviewed spoke about the decisions they made in the classroom regarding either their operational practices or the progress of their groups in terms of the curriculum and/or individual students, and not with reference to the concerns of SMT (or indeed parents). This practice of allowing teachers to interpret the implementation of policy as they see best appears to be part of the rationale of operation within the school, and has been viewed in previous work as both positive and negative in its effects. Broadbent *et al* (1992) suggests that leaving the policy making process to the senior managers, with reference in their study of how institutions interact with the market, actually preserves the sense of educational priority for the majority of staff:

> "Therefore *reorientation,* where systems do become changed, but in such a way that the values in the lifeworld are not compromised." (p.66)

Robertson (1996a), and Angus (1994) amongst others though, argue that the loss of control over policy which is implied in the imposition of tracking in this context, actually de-professionalises teachers. Whilst this question was not addressed directly by the teachers here, the impression left from discussion was that there was a degree of sympathy with both these viewpoints.

Both Karan and Janet stated that behavioural problems were more common in their mixed-track (MT) group than in the fast-track (FT) group, a feature that has been documented extensively within previous research on setted environments (Ball, 1981; Abraham, 1995). In terms of developing a working culture in the MT classrooms, the behavioural problems were compounded by the poor literacy levels of several individuals in these groups (an issue located almost exclusively within the mixed tracks), and generally poorer self-management skills. Whilst attitudes towards working varied between individuals within particular tracks, tracking overall did appear to affect students' perceptions of themselves as learners of Mathematics - Janet stated that FT students " perceived themselves as 'doing well'", whilst MT students "perceived themselves as 'low'".

The impact of setting on students' self-esteem has been widely researched. Sukhnandan & Lee's (1998) review summarises this evidence thus:

> " - research suggests that streaming and setting, compared with mixed-ability teaching, have a detrimental effect on the attitudes and self-esteem of average and low ability students. Research suggests that poor attitudes and low self-esteem can lead to a decrease in achievement which can create a vicious circle from which it is difficult for low ability students to escape." (p42)

These affective issues had implications for curricular provision for students in different tracks.

As part of the department's policy to try to avoid the negative stereotyping of lower attaining students, a 'standard' target, in terms of the number of SMILE tasks to be covered in a term, was set for the year group. Both teachers commented on the fact that whilst almost all students in the FT were capable of meeting or exceeding this target, several students in the MT groups were unable to achieve it. It was therefore easier to set individual targets for HT students, because this didn't entail contradicting department policy. As Janet put it: "this made it harder to be consistent with the M [mixed-track] groups."

The use of SMILE was generally viewed as positive, but particularly so in relation to this issue, since the individualised programme allowed more able and/or more motivated students in the MT groups to progress at a faster rate than the norm for these classes. This is likely also to have at least reduced, if not avoided, the teacher-led acceleration of pace that was a feature of the top sets observed in recent studies (Boaler, 1997; Boaler *et al*, 2000), since here too, students were working at their own rates.

The variation in attainment profile between blocks 1 and 2 makes the investigation of differences in the classroom contexts between these relevant. Enya Lama taught one MT group in block 1 and one in block 2. In order to avoid making generalisations about the differences between the blocks for MT groups from the views of just one teacher, the descriptions given by Karan and Janet of their MT groups within their comparisons of the FT and MT groups are used again here.

Enya's description of the differences between the two groups does in some respects follow the patterns that might be expected from the attainment profiles of the 'mixed' tracks. However, due to the measures taken by the school and the department to compensate for the lower attainment of B1, it does in some ways, depart from expectations too. As mentioned already, MT classes in B1 were considerably smaller than their B2 counterparts—the mean MT class size for the two blocks were 22 and 26 respectively. In addition, since there were more students with statements of special educational need, and students with English as an additional language (EAL) in B1, the MT groups in B1 tended to have more support staff in their lessons. In Enya's B1 group, which due to a large number of EAL students benefited from particularly high levels of support, these factors made a significant impact on the classroom ethos:

> "Well in block 1, the group was much smaller and this made it easier to teach. They were almost always on task and well-behaved—almost silent sometimes."

Janet confirms that the smaller classes were particularly useful, allowing her to give greater individual attention to students in her B1 (mixed-track) class.

The development of a positive work culture appears to have been harder within the MT groups in block 2 due mainly to the larger class sizes and the low (and in some cases non-existent) level of in-class support. This resulted in less teacher attention for individuals, and somewhat higher levels of disruption. Enya felt that these factors had more negative effects on the 'borderline' students in the MT in block 2 than was the case in block 1, and commented about students who were moved down in block 1: "… I don't think they suffered because of this" whilst in block 2, one student who had been moved down was described as having " 'lost it' in this group." Both teachers also felt they had problems in motivating the higher-attaining students within the mixed-tracks in block 2 who had just failed to be placed in the fast-track.

In its operation of tracking, therefore, the department's efforts to compensate for the attainment skew in block 1 that was created by placing all the DVE students together, did, to some extent, alleviate problems in the classroom. However this in turn, resulted in large classes, and greater classroom management problems in the mixed-track classes in block 2. Overall though, the differences between the blocks in terms of the atmosphere and progress within classrooms (to the extent that this can be gleaned from these qualitative data) appear to be smaller than the differences between tracks.

However, it is also important to note that there were substantial differences between teachers in the way that they taught the two tracks. Although detailed differences in teaching style are beyond the scope of this paper, there were differences between the teachers in their day-to-day operation that appeared to be directly related to the tracking system**.**

Boaler (1997) and Boaler et al (2000) noted that teachers often develop notions about what a 'top-set' group ought to be able to do:

> "The identification of students as 'top set' seemed to set off a whole variety of heightened expectations for the teachers about their learning capabilities. It was almost as if the teachers believed that they were dealing with a completely different sort of student, one that did not experience problems, one which understood the meaning of examples flashed up on the board for a few seconds and one which could rush through questions in a few moments and derive real meaning from them as they did so". (Boaler, 1997 p129)

Karan, who taught the fast track group in block 1, expressed this expectation, both in terms of ability and in terms of attitudes and motivation:

> "With the top set—what I always felt right from the beginning was it wasn't really a top group. There were big differences in the group. It was really like teaching two groups. One half worked really well, well motivated, they could start a task and carry it through on their own and discuss it, and come back when they needed help. The other half were like a mixed (MT) group—they were just there to make up the numbers." (KD)

Janet, who taught the fast-track in block 2, did not express problems in terms of a disparity between expectations and reality in the attitudes and attainment of her group although this could be due in part to the fact that her group had both a higher level, and smaller spread of attainment making the group more like a traditional 'top' set.

Both teachers described the fact that the initial act of placing students in a tracked environment had different effects on different students. For example, Janet said:

> "Some thrive, some feel threatened when low on the higher [fast-track] group …… Some thrive when high in the M [mixed-track] group"

Both were clear too that for many of these students, the attitudes seen in a tracked environment represent a *change* in attitude for these individuals that would have been unlikely to occur if they had continued in mixed ability groupings in which the majority of these students had obviously experienced a degree of 'success'. Janet described "tensions at the cut-off points", tensions that were directly related to the structure of tracking.

Karan had been in charge of collating students into tracks, and she confirmed that whilst the vast majority of decisions were made on the basis of test achievement, she had followed up the cases where the level achieved on the key stage 3 tests differed from the level awarded by the teacher taking on board where possible representations made by teachers on behalf of particular students. These may have incorporated teachers' views on the 'educability' of students, a notion discussed later in this section, but there was no evidence of the kind of biases found elsewhere (such as the over-representation of working-class students in the lower sets).

Another aspect of the 'borderline' between tracks is the issue of movement between them. The placement of students was reviewed every term, and in some cases every half term, and during years 10 and 11, a total of 21 of the 161 students in the tracked blocks changed classes. Most of these movements were movements into or out of the fast tracks, attributable to particularly good, or poor, progress in mathematics, measured either in terms of class work or performance in end of year examinations, although some changes were within mixed tracks (usually as strategies to deal with inappropriate behaviour). In this area, the practice of the two teachers of the fast-track groups differed quite significantly (although both operated within the department's agreed structure). Karan, whose FT group, as already described had a lower ability profile, viewed being in the FT group as a benefit for more able students:

> "A lot of these students, if they did well, did well because of the environment they were in, where they could actually work. They didn't have the behaviour problems going on at the same time, not to the same degree anyway, and they were learning from one another".

This perception that the FT group provided an environment that was more conducive to learning impacted on her policy on moving students out. Through Year 10, in line with department policy, students who she felt were making insufficient progress on SMILE were moved out of the FT, with, in most cases, other

students replacing them. In addition though, she makes it quite clear that she did shuffle, down and up, on a quite regular basis, students who in terms of behaviour or motivation, were not pulling their weight in the classroom:

> "What I found was, I had a group of children who didn't want to work—like [name of student] who was just there to be in the top group. She had the ability, but she didn't want to work, but once I moved her out of the group, she wanted to be back because she couldn't work in the other group."

Whilst the attainment and progress criteria for moving students between the two tracks were followed, other criteria such as behaviour and motivation clearly figured in decisions to move students between the fast track and the mixed-track in block 1. This arrangement meant that Karan also took into her group on a sporadic basis, students that were being particularly disruptive. Two students in particular, one from the FT and one from the MT, spent a significant proportion of their time in the 'other' track, although several others were moved temporarily. The use of an individualised learning scheme allowed for this to happen without unduly disturbing these students' access to the curriculum.

Research on teachers' perceptions of students' 'educability' (Haller & Davis, 1980; Haller, 1985; Dreeben & Gamoran, 1986; De Avila *et al*, 1987) matches the sentiments expressed here. Karan views some students as being able to profit from being in the fast track, and others as not being able to do so, with these judgements being formed on the basis of student behaviour and motivation, as well as attainment. Secada, (1992) in summarising the evidence of studies on teachers' views on the 'educability' of their students states:

> "This more generalised notion of educability seemed based on a combination of factors that went beyond notions of student ability to include student classroom behaviour, student ability to learn, and the like." (p.644)

As in the investigation above, no patterns were found that related these temporary movements with any of the student background variables.

Janet, teaching the fast track in block 2, was clearly aware of some of the problems commonly associated with setting, and in particular that within-class differentiation was necessary even with 'tracked' groups:

> "introducing new material required a multi-level approach for each group—if not done well the level 10s [….] and the level 3s [….] were not sufficiently assisted."

- echoing the concerns of the Chief Inspector of Schools (Ofsted, 1998, p.38) that "grouping alone is not the solution to providing effectively for students at different levels of attainment".

She was also concerned about the potential demotivating effects of moving students 'down' to the MT. This affected her decisions relating to moving students out of the FT, resulting in her 'hanging on' throughout Year 10, to students whose progress indicated transfer to one of the mixed track groups, making her group very large. This policy also made it difficult to move high-performing students in the MT into the FT since there was no room. Janet, whilst defending her actions, acknowledged that this worked against some students in the MT:

> "a few in the M [MT] group needed to be seen individually and would benefit less from discussion with fellow classmates."

Karan too, takes up this issue of interaction stressing its positive effects on students who were prepared to engage with their work. This recognition of the role of interaction in learning, across the school and within the department, raises the central tension of the tracking arrangement. Mixed ability grouping, due to the skew in numbers decreases the opportunities for higher attaining students to interact constructively whilst tracking does the same for lower ability students.

Janet maintained too—having seen the exam results of the students she moved down—her opinion that the move overall had negative effects, particularly on the two girls:

> "[Names the two students] seemed to do less well after leaving the H [FT] group."

The implementation of tracking clearly benefited those students placed in the fast track though the provision of a generally 'well-behaved' atmosphere that allowed for the establishment of a stronger work

ethic, and faster rates of progress. The negative features of pressurised conditions of learning, and emphasis on speed associated with top sets in the research of Boaler (1997) and Boaler *et al* (2000) seem to have been avoided in the fast-tracks here through the continued use of individualised learning programmes. However, progress of students in the mixed-track groups was hampered by higher levels of disruption, lower levels of literacy and poorer self-image in terms of mathematical proficiency although it should be noted that all the teachers stressed that the structure of tracking impacted on different students in different ways, and that this variation was particularly marked at the borderline between the two tracks. Some students became more motivated simply by virtue of being placed in either the FT or the MT, while others found this demotivating.

The impact of tracking on achievement

The relationship between key stage 3 tests score and GCSE score[5] for the 180 students for whom both data are available is shown in figure 3, with separate regression lines shown for mixed-ability (MA), fast track (FT) and mixed-track (MT) students. Analysis of covariance (ANCOVA) shows that there is a significant interaction ($F = 12.19$, $p = 0.006$) between key stage 3 test score and assignment to either the mixed-ability block (block 3) or the blocks containing fast tracks (blocks 1 and 2). The model predicts that on average those scoring around level 4 at key stage 3 would achieve grade G at GCSE if placed in the tracked blocks and grade E if placed in the mixed-ability group. Those scoring around level 5 at key stage 3 would achieve an E if placed in the tracked blocks and grade D if placed in the mixed-ability group. Only for the highest attaining 12% is placement in the tracked blocks (and therefore, by definition, the fast-tracks in these blocks) advantageous—in other words, placement in the fast track is beneficial for only the upper half of each fast-track. For other students, in this school, tracking made no difference or was deleterious.

Whilst overall this finding in favour of mixed ability grouping is consistent with those of Boaler (1997) and Linchevski (1995) this result should be interpreted with considerable caution, being based on just a single mixed-ability class, in which it is impossible to disentangle other factors such as teacher effect. At the very least though, such a highly significant difference across a wide range of student attainment does indicate that mixed-ability grouping in mathematics does merit further attention.

The differences in performance between students assigned to the mixed-tracks and fast tracks were also significant ($F = 8.61$, $p = 0.004$), with placement in the fast track proving advantageous - a finding supported by earlier research (e.g. Kerchkoff, 1986; Hoffer, 1992; Linchevski, 1995).

The impact of moving between tracks also needs to be considered here, and reveals a mixed picture. Upward movement resulted universally in achievement at or above the model for the relevant fast track while downward movement produced much more mixed results, with many students achieving lower GCSE grades than would be predicted on the basis of their key stage 3 test scores (whether based on the track they left, or the one to which they moved).

The ANCOVA models discussed above showed no significant interactions with sex, but there was a strong ($F = 5.72$, $p < 0.001$) interaction between key stage 3 score and ethnicity (ethnicity data was available on 200 of the 240 students in the cohort). Taking key stage 3 scores into account, Asian students scored 11.3 marks higher on the GCSE than whites, (standardised effect size, $d = 0.44$), with Afro-Caribbean students scoring 3.2 fewer marks than white students on the GCSE, again taking key stage 3 attainment into account ($d = 0.12$).

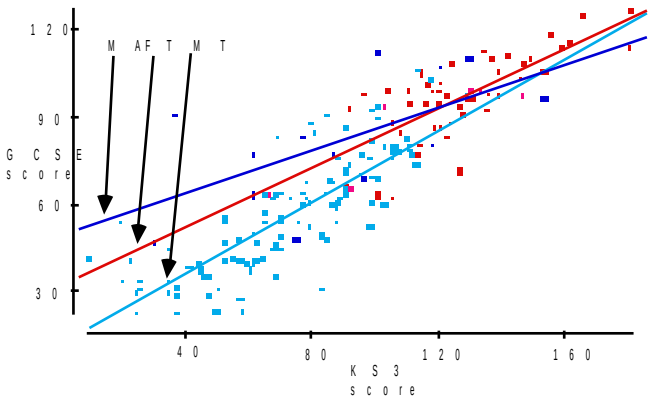

Figure 3: relationship between key stage 3 score and GCSE score for different groups

As noted earlier, and as illustrated in figure 1, the composition of the blocks differed, with more high attaining students assigned to block 2. However, there was no significant interaction between block assignment and KS3 test score, so that the 'value-added' in each block was comparable. However, while the regression lines of the mixed track and the fast track in block 2 are almost collinear (indicating that whether a student was assigned to the fast-track or the mixed-track in that block made little difference), students assigned to the fast track within block 1 were at a significant advantage. In block 1 students in the fast-track whose attainment on the key stage 3 tests was in the overlap area (i.e. those achieving scores between the lowest score achieved by a student assigned to the fast-track and the highest score achieved by a student in the mixed track) performed approximately one grade better at GCSE than those with comparable scores in the mixed-track ($F = 10.98$. $p = 0.001$). There was also an interaction between ethnicity and key stage 3 test score in block 1 (though not in block 2), and the interaction of key stage 3 test score with socio-economic status (as measured by entitlement to free school meals) was bordering on significance in block 1 ($p = 0.090$) but not in block 2 ($p = 0.744$).

Discussion

The fact that mixed ability grouping appears to strongly benefit the majority of students in this cohort, is worth considering from several standpoints. Firstly, the political climate continues to press schools to move away from heterogeneous grouping:

> "Unless a school can demonstrate that it is getting better than expected results through a different approach, we do make the presumption that setting should be the norm in secondary schools." (White Paper, Excellence in Schools, 1997)

The findings of research on the relative merits of heterogeneous and homogeneous grouping are mixed (Slavin, 1990), but a large number of studies over an extended period have found a pattern which is in accord with the one seen in this context—that setting (or tracking) produces slightly higher results for higher attaining students, whilst lowering the results of average and below average attainers (Lacey, 1974; Postlethwaite & Denton, 1978; Askew & Wiliam, 1995; Boaler, 1997). Within this study, it is also notable that the significant gains in achievement seen in the mixed ability group are made *in spite* of the fact that both tracked and mixed ability classes used SMILE. The studies by Boaler (1997) and Boaler *et al's* (2000) suggest that the poorer performance of students in setted environments was due to the imposition of lower level content and reduced pace in lower sets, and an emphasis on speed in the top set. This marked difference of content and pace between the tracked and mixed ability environments is not a significant feature within this school since both use individualised work schemes (although it has not been entirely eradicated). The gain in this context then, would appear to be related more fundamentally to the grouping structure than was the case in those studies.

The differences in performance between the tracked blocks allow for an examination of how different student compositions, and to an extent, different teacher practices impact on eventual attainment. The significant 'gap' in performance between the fast and mixed tracks in block 1 is particularly interesting here. This 'gap' is caused to a large extent by the sizeable cluster of fast track students at the lower end of the attainment range for this group who achieved particularly well in relation to their group. Several factors could have contributed to this. Firstly, the student composition in the fast-track in block 1, which contained more students working at lower levels than in block 2, may have led to a greater emphasis on 'lead-in' or 'consolidation' type work on core topics in whole-class work, to ensure the achievement of a grade C at GCSE level. The more balanced attainment profile in the fast track in block 2 could have switched this emphasis to differentiated work across all levels, with extension work being provided for the most able. Evidence of this focus on differentiation is available in the qualitative data for the fast track in block 2. A similar argument could hold for the mixed tracks in both blocks where a similar pattern of performance was seen.

Differences in the policies of the two fast track teachers on moving students could also have contributed to the cluster of positive residuals seen in the overlap region in block 1. The evidence here suggests that the use of movement between tracks regularly, on a fluid and temporary basis for students who were not well motivated, has succeeded in raising levels of achievement in this range. These gains are not limited to students involved in these movements. The fluidity of movement seen here, both on an official basis across both blocks and more unofficially in block 1 match the recommendations that are often made about how to improve the effectiveness of setting (see for example Hallam & Toutounji, 1996).

The significance of 'track' and 'ethnicity' (and the near significance of FSM) in block 1 raises issues about the ethics of the structure of blocking used within the school, since it appears that having a more negatively skewed population in terms of prior attainment is associated with poorer performance for higher ability students in both tracks, and for some ethnic groups ('African Caribbean' and 'white' students in particular here). One theory for these findings may be related to research which points out the positive effects of a more 'comprehensive' mix on eventual attainment (Willms, 1986; Gewirtz, 1997). The ANCOVA analyses for blocks 2 and 3 support this, with the results for block 2 in particular showing that none of the student background factors made significant difference to them.

Conclusion

What is clear from this study is that the effects of ability grouping are highly complex, and currently theorisation is inadequate to account for them. What goes on within classes, whether tracked or setted or not, is as important as how those classes are constructed. While this study has replicated some findings of earlier studies, others are apparently contingent, rather than necessary, features of homogenous ability grouping. In particular, it appears that concerted attention to the possible negative consequences of homogenous ability grouping may ameliorate its impact, allowing some of the positive features of mixed ability learning to be retained (see Sukhnandan & Lee, 1998, p56, 58 for a summary of these). The government's uncritical support of homogenous ability grouping, irrespective of whether the students are taught individually or as a whole-class, and of the scope and flexibility of the grouping structure, is certainly not grounded in evidence. To the extent that the evidence points either way, in fact, it is towards the greater use of mixed-ability grouping, in which case the advice to schools should be that:

"Unless a school can demonstrate that it is getting better than expected results through a different approach, we do make the presumption that mixed-ability grouping should be the norm in secondary schools."

Notes

References

Address for correspondence

Dylan Wiliam, School of Education, Franklin-Wilkins Building, Waterloo Road, London SE1 8WA; Fax: 020 7848 3182; Email: dylan.wiliam@kcl.ac.uk

[1] Paper presented at British Educational Research Association 26th annual conference, September 2000, University of Cardiff, Cardiff, Wales.

[2] In this paper, ‘blocking’ is used to refer to a process in which the year-group is divided into two or more groups for ease of timetabling, and where each block has a representative range of ability. ‘Banding’ is usually used to refer to a process of dividing an age-cohort of students into two or more ‘bands’ based on perceptions of general (as opposed to subject-specific) ability, but within this context teachers used the term to describe the 'broad' division of the cohort into 'sets' on the basis of perceptions of Mathematics-specific ability. Where the number of bands is equal to the number of teaching groups, then this is termed ‘streaming’. Where classes are grouped into particular classes based on ability in a particular subject, this is called ‘setting’.

[3] Originally the Secondary Mathematics Individualised Learning Experiment, now called the Secondary Mathematics Individualised Learning Experience, or just Secondary Mathematics Individualised LEarning.

[4] Since the students were entered for different ‘tiers’ of the key stage 3 tests, the different tiers were equated using standard linear test equating within the level cut-offs specified by the Qualifications and Curriculum Authority. The equated scores have been scaled to the scores on the easiest tier. The scores on key stage 3 tests (and indeed GCSE) are not particularly reliable (estimates of the reliability of the key stage 3 tests range from 0.80 to 0.85), but this has little impact on the current study because the focus here is on groups of students (rather than individuals).

[5] Again, linear equating was used to place scores from different ‘tiers’ of the GCSE on the same score scale, with scores scaled onto those on the lowest (Foundation) tier.