ANNE WATSON and ELS DE GEEST

LEARNING COHERENT MATHEMATICS THROUGH SEQUENCES OF MICROTASKS: MAKING A DIFFERENCE FOR SECONDARY LEARNERS

Received: 28 September 2009; Accepted: 16 February 2011

ABSTRACT. We report on the work of 3 schools which set out to make a difference for their previously low-attaining students. Through naturalistic enquiry over 3 years, we built a picture of their practices. Against a national trend, students reported positive attitudes to mathematics and, in 2 out of 3 schools, showed improvement in attainment, so we probe more deeply into the teaching. The lessons broadly conformed to an approach combining inclusive participation with complex and coherent development of mathematical ideas. To understand how the teachers orchestrated this, we developed lesson analysis techniques that focus on how thinking, repertoire and understanding are scaffolded by teachers through sequences of microtasks.

KEY WORDS: improving mathematics learning, lesson analysis, low-attaining students, mathematical coherence, scaffolding, tasks

Introduction and Rationale

Underachievement in mathematics, within countries, of significant numbers of school students is an international problem (Mullis, Martin & Foy, 2008, pp. 35 and 41). Typically, identifiable groups of students, such as those with different language backgrounds and those from lower socioeconomic rankings, underachieve in national and international tests and several countries have initiated policies and practices to tackle this issue and broaden access and achievement beyond those groups that are traditionally successful. Methods of approaching this issue range from macro-changes in policy, curriculum and assessment to institutional change, provision of extra teaching and micro-advice about inclusive teaching in classrooms. In this paper, we report on three schools in the UK context who decided for themselves to improve mathematics learning for low-attaining students. We intend to describe the mathematics teaching by which they hoped to support learning and thus contribute to knowledge about effective mathematics teaching at the classroom level.

The 2007 Trends in International Mathematics and Science Study (Sturman, Ruddock, Burge, Styles, Lin & Vappula, 2008) reports that

International Journal of Science and Mathematics Education (2012) 10: 213–235

English students have improved their mathematics achievement significantly since 1999. However, it also reports a decline in enjoyment, with the sharpest decline for year 9 (grade 8) students of 25 percentage points. Six countries scoring at a higher or similar level of attainment also showed a decline, but none as large. Recently, improvement is not evident in England's national test results in year 9, which show an increased proportion achieving a certain standard until about 2003, but then levelling off. The proportion which does not attain nationally expected levels in year 9 is stuck at around 20%.

When three schools approached us to say that they intended to 'make a difference' for their lowest attaining students, we took the opportunity to observe what they did. Because the decisions about change were the schools', and not ours, we set up an ethnographic study to construct a chronicle of change.

There were good reasons to anticipate success. As Hattie reports in his meta-analysis of meta-analyses of successful change (2009), "everything seems to work" (p. 1) if there is a "constant and deliberate" intention to change (p. 12). Another reason for optimism was in the findings of an earlier project (IAMP, see Watson & De Geest, 2005) which had characterised the regular work of ten teachers who had made a difference for low-attaining students. On many of the usual pedagogic dimensions, their teaching varied between traditional and reform, student-centred and teacher-centred and transmissive and constructivist. What they had in common was something else—the belief that all students could learn mathematics better, could become better learners of mathematics and would feel better about learning mathematics. Would the collective intention of school mathematics departments have the same good effects?

We describe methods of lesson analysis that support a conjecture that improved learning is achieved when teachers scaffold mathematics thinking, repertoire, confidence and understanding through sequences of microtasks, rather than by conforming to generic descriptions of good teaching, or through particular task and problem types.

The Changes in Mathematics Teaching Project

In the Changes in Mathematics Teaching Project, we studied the three mathematics departments over 3 years as they changed the way they taught for one cohort of 11- to 14-year-old students to help previously low-attaining students (PLAS), as well as all other students, to learn more mathematics. Two of the schools, LS and SP, served inner-city areas of social deprivation, one of them highly multicultural, one predominantly white working class. Examination pass rates had been typically at or below half the national average, but the first school scored highly and the second was average on national value-added measures. The third school, FH, served a diverse rural area and had results typically 10% above the national average, but with relatively low value-added scores. Each school had an entry cohort of around 180+ students, organised into seven teaching groups, which were initially mixed-ability groups in all three schools. We knew that national issues such as the politicisation of school mathematics and volatility of mathematics staffing might impact on the continuity of plans and intentions of the departments over the 3 years. Only two of the three departments, LS and SP, maintained a concentrated focus on the cohort we were watching and on PLAS within that cohort, throughout the 3 years of the project. FH shifted their main focus away from this cohort but sustained many of the initial changes they had made to their teaching for these students.

Our methodology was naturalistic enquiry, which aims to develop descriptions of participant perspectives, intentions and interactions in cases that can inform others engaged in similar work (Lincoln and Guba, 1985). Data collection embraces observation and teachers' perspectives on their actions over a period of time and is gathered in collaboration with the teachers. Multiple sources of data are sought to triangulate and interrogate each other and are checked with the teachers to ensure credibility. It is typical of naturalistic enquiry that hypotheses, and therefore methods of enquiry and analysis, emerge and evolve during the research. For this reason, reports of the outcomes of naturalistic enquiry cannot always state theoretical frameworks at the start because the need for frameworks often emerges as enquiry proceeds. Enquiry is complex with multiple strands of enquiry intermeshing as researchers attempt to capture the complexities of real practice. We cannot report all related factors in one paper. The standards of acceptability of naturalistic research are *credibility*, which is achieved through triangulation and member checks; *transferability*, which requires sufficient detail that other practitioners can use the study and *dependability* through provision of archives so others can retrace the research. We meet these standards by (a) coordinating analyses of different data and drawing on literature as issues arise, (b) foregrounding details of practice and (c) referring participating teachers (and other interested people) to a website archive[1].

Since we did not know what teachers were going to do, we could not be precise at the start about the focus and direction of our research. We used three broad research foci to guide the collection of data:

1. How departments and teachers worked in relation to the cohort
2. The attitudes, experience and learning of PLAS as evidenced by cohort test results and through interviews
3. What teaching was like as evidenced from observations and video analysis

We collected a range of data, visiting each school for between 9 and 12 days each year and also organising off-site meetings with the heads of department. Analysis was ongoing throughout 3 years. Hypotheses and findings based on data collected in the first year informed collection methods and direction of enquiry in the second year and so on. In the first 2 years, we collected teacher interviews, lesson observations and videos, audio recordings from department meetings, schemes of work, lesson ideas, interviews with a sample of PLAS, test scripts, students' work and background data about past achievements and school statistical predictions. In the third year, we revisited schools to observe any further changes to practice and to interview heads of department again and collected national test scores.

The schools had taught earlier cohorts in attainment groups, and this study focuses on those who would have been in 'low' groups under this system. All schools started with mixed-ability groups; two schools, LS and FH, changed to attainment groups for year 8 and all schools used attainment grouping for year 9.

For this paper, we shall describe the cohort national test outcomes first in order to show that reporting this study is worthwhile. We shall then report on the teaching[2]. Relevant literature and methods will be introduced as the issues arise, in the spirit of naturalistic enquiry. Finally, we shall conjecture connections between the two.

Achievement in the Three Schools

Our commitment to the departments was that we would adopt a non-intervention stance, not imposing any extra testing. Since schools and individual students were judged by national test results in year 9, we used these to compare the achievements of the cohort to the previous year—a contextually valid measure that is meaningful for teachers. These tests included non-routine questions that require some interpretation and application, as well as procedural questions. It is only partly possible to 'teach to the test' by using a procedural and question-spotting approach (Ofsted, 2005). Test results for individuals are given as levels from 1 to 8, level 5 being considered the 'pass' score. Results are given as 'pass' percentages of the cohort. See Table 1.

In their context, all three schools had been noticeably successful for the whole cohort, either by improving performance above the national and school background or, for FH, by maintaining performance in a context of unusually falling levels in other tested subjects. Nationally, test results in all three subjects had levelled out after a rise at the start of the decade, pass percentage scores for 2007/2008 being science 73/71, English 74/74 and maths 76/77.

Our first interest was in PLAS. For this, we looked at progress of students who entered the school with levels 2 and 3 in their pre-entry national tests, for which 'level 4' is considered a pass. Results for PLAS were not distributed normally, being from the lower end of the distribution for the whole cohort and hence skewed, and could not be treated as interval data since they consisted of a few integers. Another problem is that the tests are not standardised. We used the Mann–Whitney test based on ranking individual test scores of PLAS for 2007 and 2008 to look for significant differences between the 2 years (Table 2).

In SP and LS, who maintained the focus on this cohort, PLAS results for our cohort in 2008 were significantly higher than for similar students in 2007, but with a small effect size. In FH, there was a significant negative change for PLAS. The achievement of PLAS had improved in two schools but fallen in FH.

However, despite FH failing to make a positive difference for PLAS and effect sizes being small, overall all three schools had achieved a

TABLE 1

National test results in core subjects for our cohort 2008 and 2007

'Passes' in year 9 national tests as %	*SP*		*FH*		*LS*	
Year	*2007*	*2008*	*2007*	*2008*	*2007*	*2008*
Maths	47	61	79	80	53	62
English	59	56	76	69	48	43
Science	46	45	77	69	53	54

TABLE 2

Comparisons of PLAS' achievements in 2007 and 2008

School	*Mean rank 2007*	*Mean rank 2008*	*Mann–Whitney U*	*Effect size*	*p*
SP	55.50 (n = 62)	64.89 (n = 57)	1,488	−0.148	0.048
LS	38.92 (n = 45)	47.59 (n = 40)	716.5	−0.189	0.041
FH	34.05 (n = 29)	24.95 (n = 29)	288.5	−0.292	0.013

broadly positive effect for the whole cohort, when considered in relation the two other core subjects. It is worthwhile looking at their practice and, in particular, to look for differences between SP and LS on the one hand and FH on the other.

Affective Factors

Some major studies relevant for mathematics relate dissatisfaction of low attainers to the nature of teaching groups (Boaler, Wiliam & Brown, 2000; Wiliam, Brown & Boaler, 1999) while others draw on perception of self, such as confidence or self-efficacy (Ireson, Hallam & Plewis, 2001; Pietch, Walker & Chapman, 2003) or classroom environment (Dorman, Adams & Ferguson, 2002). Dorman et al. found that environments that offered more support, cohesion, involvement, cooperation and equity led to stronger feelings of capability and satisfaction. It was therefore critical to collect information from students to check that achievement was not at the expense of well-being.

At the start of the project, we selected 43 students (15 LS + 13 FH + 15 SP). The selection was based on:

- Students for whom we had data who had previously attained levels below 4
- Students for whom we had no entry data but who had below average scores on a test of general cognitive ability (given by schools to all students). These constituted about 10% of the target students

We used alphabetical lists of target students from each group, selecting every *n*th student where *n* would give us roughly equal numbers per teaching group. Thus, each teacher of year 7 was similarly represented in the sample. Informed consent to be interviewed regularly was obtained from students and parents and most teachers.

Of our sample, only 25 students of out of 43 were present for all three interviews (interview 1: n = 41, interview 2: n = 33, interview 3: n = 33). We planned for up to 30% attrition. Ethical constraints prevented following up absentees, mostly due to students leaving or erratic attendance. Interviews were conducted by the same researcher every time to develop a familiar relationship. Students were interviewed in pairs or threes to avoid them feeling intimidated. The interviews consisted of structured question sequences. Two of the interview questions related to students' feelings and perceived capability in mathematics, and we will discuss these in detail. We did not find any existing instruments which could be used with this sample without significant support to handle language and length. All our questions were displayed on A4 card in large print, while at the same time being read aloud by the researcher. Pictures were used to illustrate concepts which might not be clear. We did this because many were relatively new to using English and/or were weak readers, and hence, a traditional questionnaire would be inaccessible.

Students' Feelings About Mathematics

In the first interview, we asked students how they had felt about mathematics in year 6, in their previous schools, as well as their current experience. In the two later interviews, we asked them how they felt about mathematics in their current year. Feelings were 'measured' using five smileys (see Figure 1).

In all years, most feelings were slightly smiley or neutral. There were some extreme feelings at both ends at the start of the study, but these extremes reduced by the end to settle around 'slight smile' or neutral (see Figure 2).

A slight dip from smiley 4 to smiley 3 in winter of year 7 appeared to be from one school, FH. This school had the most staff difficulties and abandoned the special focus on 'our' cohort, choosing instead to review their practices for the following cohort. Students' feelings at FH improved in general in year 8, however.

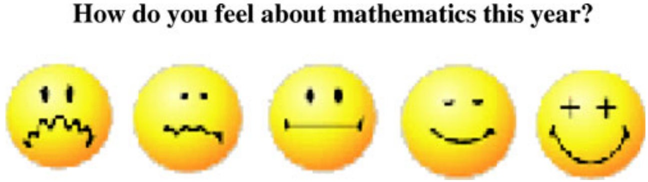

Figure 1. The 'smiley' question

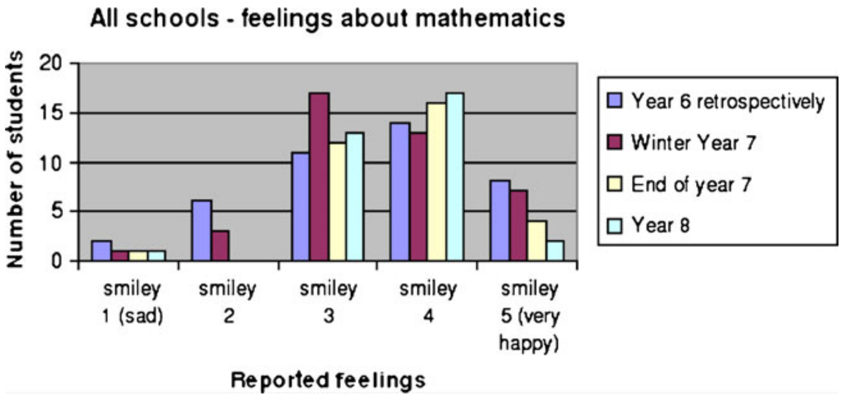

Figure 2. All schools—feelings about mathematics

Student Confidence

In the IAMP study, teachers had agreed that students need to feel better about learning mathematics (Watson & De Geest, 2005, date), so for this study, achievement should be seen against a backdrop of students' confidence about mathematics. We could not use standard questionnaires because many would be daunted by the length, density and language. We asked the students to indicate which description fitted their view best:

- I am very good/good/ordinary/not so good at mathematics

Students did not in general have a low opinion of their own mathematics (Figure 3) and 6 at the 'not so good' end developed a more positive impression, but also 8 shifted from 'good' to ordinary in schools that changed to attainment groupings. This change did not occur in the school that stayed with mixed ability in year 8.

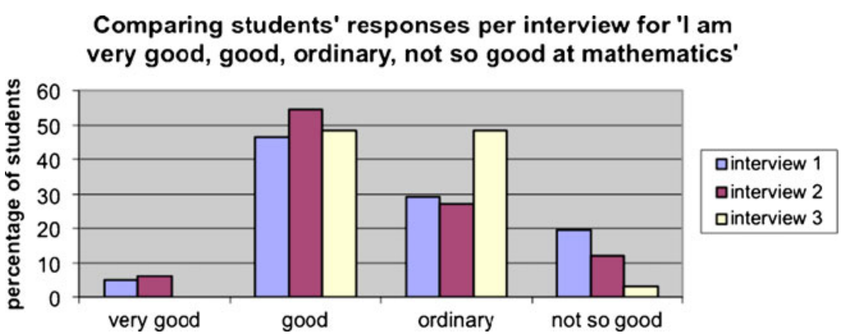

Figure 3. Comparing students' responses per interview for 'I am very good/good/not so good at mathematics'

Summary of Affective Outcomes

Several studies show that negative comments about mathematics are easy to elicit from students, so there was no strong reason to suppose that our research process would hide these (Nardi & Steward, 2003; Brown, Brown & Bibby, 2007). In the study by Boaler et al. (2000) of 943 students moving from year 8 to year 9 in six schools, four schools changed from mixed-ability to 'ability' grouping at this transition, and this was associated with an increase in dissatisfaction. In our three schools, PLAS' good attitudes to mathematics were maintained through the 3 years of secondary school, and self-confidence was largely maintained with some small negative shifts, but we did not find evidence of dissatisfaction among our small sample.

What the Teaching Was Like

We now focus on qualities of the teaching. We recorded and observed with field notes 40 lesson videos from 21 teachers of PLAS who had agreed to be in the study. For the majority who taught PLAS in all 3 years, we have two videos, one towards the end of year 7 and one towards the end of year 8 and some further observation data from the beginning of year 8. For year 8, as well as videoing teachers who were currently teaching PLAS, we videoed those who had been teaching several PLAS in year 7 but were now teaching groups which contained a few PLAS. We also included three videos of teachers in FH and LS who had taught PLAS in year 7 but no longer did in year 8. We did this because they reported that their experiences in year 7, and the development work they had done on their teaching, had influenced their work with all students and we wanted to understand their sustained teaching, rather than their initial innovations.

What Good Teaching Is Like

We did not have preconceptions of 'more' or 'less' desirable ways of teaching. In this sense, our research is of the first type described by Krainer (2005), that is, 'refusing norms', in that we did not define good teaching and then set out to see if we could observe it (his second type) but sought for new insights. Nor was it his third type, in which teachers and researchers collaborate to develop new norms. In the IAMP project, there had been wide differences in observable pedagogy, but between teachers' basic shared values and observable pedagogic differences, we had found a layer of similarities (Watson & De Geest, 2005). Many of these similarities were in accord with the generic findings of Day, Sammons and Kington (2008) who undertook a national study of 136 effective UK primary and secondary teachers, based on achievement of higher test scores than expected. The features they found which typified the most effective teaching included:

1. Tasks were challenging and students had some control over what they did.
2. Relations built self-esteem, trust and respect.
3. Teachers stimulated students intellectually, scaffolding their thinking.
4. Students use a range of resources, media and working methods.
5. All students were valued as part of the community.
6. There was dialogue about learning the subject.
7. Expectations were high, individualised and consistent.

Effective teaching, therefore, even when measured in terms of national test scores, is more complex than instrumentally ‘teaching to the test’. From studies with the highest effect sizes on learning, Hattie (2009) found that “it is what teachers get the students to do [that is] the strongest component … rather than what the teacher, specifically, does” (p. 35) and questions whether there are subject-specific ‘better’ methods of teaching (p. 12). These studies suggest that a focus on tasks and problem types might reveal reasons behind the relative effectiveness of our three schools.

Effective Mathematics Lessons

Effective mathematics teaching is culturally specific (Askew, Hodgen, Hossain & Bretscher, 2010) because it depends on how knowledge, schooling and the teacher–student relationship is perceived within the wider political, social and historical context. For that reason, we cannot draw on ideas about ‘best’ teaching by looking at high achieving countries and comparing overt features of lessons. Furthermore, the effectiveness of a method is relative to how the method outcomes relate to the test material. Merely practising standard algorithms is not ‘best’ in relation to helping students become better learners, but achievement is measured by national tests of concepts and procedures, so focusing solely on mathematical thinking without building up repertoire might not be ‘best’ either. Typical mathematics lessons in high achieving countries vary widely, from being largely based on silent individual algorithmic exercises to being mainly about whole-class problem solving (Hiebert, Gallimore, Garnier, Givvin, Hollingsworth, Jacobs et al., 2003). It made sense for us, therefore, in the spirit of naturalistic enquiry, to look for a method of analysis that related to the aims and context of the teachers in our schools rather than apply frameworks for general teaching or to seek for international standards.

Methods of mathematics lesson analysis used in the METE project by Andrews and colleagues (e.g. Andrews, 2007) include identifying the sequence of mathematical purposes. Their method develops descriptions of the epistemological foci in lessons. Our teachers wanted to develop students' thinking: the *effortful* selection of givens, knowledge and methods of combining these (Sternberg, 1986), so we needed insight into where effort might be exerted. The TIMSS seven-nation video study (Hiebert et al., 2003) reported the frequency of use of answer-only problems; concurrent problems; time taken or problems in individual or class work; connectedness of problems; whether problems were for practice, new content or review and whether the teacher made summary statements or goal statements. Their classification of problem types itemised: procedural complexity (the number of decisions and operations and subproblems), reasoning, contextualization, number of solutions, students' choice of methods, critical reflection and connections between concepts. The focus on task and problem types showed differences between typical lessons in higher-achieving countries, so clearly problem and task choice is not a key factor in high achievement.

The Mathematics Quality Analysis Group (MQAG; Hiebert et al., 2003, pp. 189–202) probed the mathematical content of lessons and found that what lessons in high achieving countries had in common was the presence of deductive reasoning, strong mathematical rationale, use of generalisation, maintaining complexity and coherent progression of mathematical ideas. The METE and MQAG studies suggest that what matters is not the general mode of teaching nor the nature of tasks and problems but the way mathematical content is made available for learners.

When we set out on this project, our expectations were of a repetition of IAMP findings, namely a variety of observable practice, with similarity of underlying values and mathematical pedagogical principles. We did not foresee that we would have to look at the fine grain of classroom tasks to describe what happened. The methods of analysis we devised for fine grain analysis are new and are an outcome of the study.

Development and Use of Analysis Methods for Lessons

We used a cycle of three methods in our analysis of lessons. The first was grounded analysis of the first six videos to get a sense of the range of practice before we continued observing and videoing. It was this process that showed us that, unlike the IAMP study, lessons were surprisingly similar at the level of describing the range of teacher strategies, methods of participation and epistemological foci.

The second method was to use an observation schedule on all lessons in the sample, using what we had learnt about teachers' aims from their interviews. The third method focused on details we realised were missing from the second, using the MQAG ideas as a starting point. We also used interview data from teachers throughout to contextualise our observations and analysis, so all statements about teachers' purposes in what follows have been triangulated.

Possibly because of ongoing discussion about teaching in the three schools, all teachers except one adhered to broadly common practice. This included *whole-class teaching*, *eliciting and using students' ideas*, *providing opportunities for students to collaborate*, *problem solving and reasoning tasks*, *procedural and conceptual work* and *clarity about lesson aims*. The extent to which individual teachers used any of these varied widely but the conditions for stronger feelings of capability reported by Dorman et al. (2002) were broadly present. All teachers agreeing to be in the research stated similar social and affective aims for students such as resilience, learning from each other, making efforts, organising themselves, listening to each other and caring about their work. Most said that achievement depended on *mathematical thinking*—and suggested various ways to generate this including: thinking about unfamiliar ideas; having a repertoire of factual knowledge; using mathematical expressions, representations and language; reasoning to follow through ideas; hypothesising; posing questions and learning from mistakes. In department discussions, it seemed that these methods were supposed to be embedded in all teaching, and not dependent on particular types of task. Indeed, the lessons we analysed for the first cycle all had a range of task features: quick closed questions, fully directed tasks, repetitive exercises, extended tasks, realistic contexts, open questions, discussion of key ideas, student choice, exercises and extended problems. Lessons focused on procedural and exploratory, contextual and abstract, practical and symbolic tasks, done individually or in groups. In other words, the teachers used a very wide variety of task types, not just those identified by Day et al. (2008).

For our second analytical cycle, which we applied to 38 lessons, we focused systematically on ideas from MQAG and teachers' stated aims. Two lessons were omitted: One did not contain any mathematics but was concerned with organising work folders and equipment; the other consisted of individual low level arithmetic tasks with inaudible one-to-one help.

Lesson observation notes were taken in the following format:

Lesson segment	Task set	How the task was set	Specific things said about maths by teacher	Examples used	Pupils' contributions about maths	Questions and prompts used for whole class

In addition to the schedule, there were summary questions such as:

Task types used
Questions and prompts used for whole class/small groups/individuals
What mathematical ideas were emphasised in whole class/small groups/individuals?
What was said about good mathematics or mathematical thinking?
Habits and patterns of interactions
How were right/wrong answers dealt with?
How were learners helped to deal with complexity?

From these notes, we collated a range of features of the observed lessons to see if we could characterise the teaching. We shall report on the ranges of mathematical ideas, how learners were helped to deal with complex tasks and question types. These were aspects in which we found most consistency across all lessons and purposes and which have been found credible and transferable among readers of the website from whom we have had, and continue to have, feedback.

Mathematical Ideas

A range of mathematical ideas and methods was emphasised in lessons. By 'mathematical ideas', we mean characteristics of the subject that extend across topics. Those we observed being emphasised included the purpose of definitions versus descriptions, algebra as a way to express generalities, equivalence, the meaning of '=', when to be accurate and when to estimate, the importance of inverses, use of alternative representations and considering non-integers and negative numbers. Thirty-two lessons included explicit comments on at least one such idea. These emphases indicated that most teachers were aware of the nature of mathematics and incorporated ways of working and mathematical habits of mind in their lessons. Fourteen teachers went further in the lessons we saw and made explicit statements about how to work on mathematics such as:

- Looking for difference and sameness
- Creating examples to explore ideas

- Thinking about the efficiency and power of solutions
- Talking about relationships and expressing them algebraically
- Contrasting everyday observations with mathematical ones
- Including reasons and justifications in answers

Most of these are specifically mathematical, and in most lessons, explicit or implicit praise was given for displaying these kinds of mathematical behaviour. In this way, these teachers were generally conforming to the qualities of 'more effective' teachers described by Day et al. (2008) but with a specific explicit mathematical focus. In three lessons, no statements about the nature of mathematics, or what is mathematically important, or associated praise, were heard at all: Two of these lessons we earlier described as 'omitted' and the third was with a teacher whose second video did contain such utterances.

Complexity of Tasks

To focus on the qualities identified by MQAG, we analysed the complexity of tasks and examples used. We made a distinction between tasks which focused on one feature and tasks which involved two or more interconnected variables and perhaps their relationship. Fourteen teachers offered some of the latter kind of complexity during every lesson that we saw. These teachers did not simplify the questions for PLAS in the observed lessons. In about half the lessons the examples used were not simple enough to be resolved using ad hoc methods. We did see two teachers during years 8 and 9 breaking questions into smaller steps for some students thus reducing multi-stage tasks to calculation questions. They also changed tasks so that PLAS could use empirical approaches rather than reasoning. Most teachers did, however, maintain complexity and the need for reasoning.

Teachers used a variety of ways to scaffold students' work to new levels of complexity. We classified these into those which develop thinking, repertoire, confidence and structural understanding. They scaffolded:

- New ways of *thinking* by thinking about mathematics out loud, posing questions whose answers cannot be guessed but can be reasoned from current understanding, demonstrating critical reflection on a 'wrong' answer and asking students to imagine the effects of methods
- Extensions of *repertoire* by comparing the scope of different methods, and evaluating the consequences of choice; demonstrating how learnt methods can be linked to solve a multi-stage problem; emphasising multiple representations; asking for more than one

example, or for unusual examples and thinking about classes rather than single objects
- *Confidence* by having harder ideas and examples available; discussing difficulties as sources of knowledge; using formats, grids, layouts, writing frames, to organise work and asking students to make up ‘hard’ examples
- *Structural understanding* by focusing on relations, systematic variation, increasing parameters from one to two to three and encouraging appropriate generalisation

Our second method of analysis had therefore yielded a rich description of scaffolding strategies for use when working with PLAS and also showed clear differences between most teachers and those who simplified tasks for PLAS rather than provide scaffolding. We began to understand that it was this kind of complex scaffolding, rather than task types, that might make the difference for PLAS and others.

Question Types and Microtasks

We then looked at question types, but in most whole-class interactive episodes, teachers asked short closed questions, waited a very short time for replies and dealt very briefly with student responses, whether right or wrong, conceptual or procedural. Boaler & Brodie (2004) also found that, in a study contrasting question types between ‘reform’ and ‘traditional’ teachers, such questions were frequent in both classifications. Our teachers told us in interviews of the value of open questions, longer wait times for answers[3] and the value of elaborating answers. This apparent contradiction provided a reason for us to focus on the public availability of mathematical ideas rather than question types—the dominance of short closed questions had not prevented improvement of learning. We therefore asked what is the learners’ experience of mathematics during these questioning sequences? This led to the realisation that *all* lessons could be analysed as sequences of teacher-influenced microtasks through which students’ engagement with a mathematical idea might undergo change. Sometimes this was structured by teacher intervention with individuals or groups, but more often the structuring of understanding was public and common for all students in the class in interactive episodes alternating with individual, pair or group work.

All three schools had started out thinking that to foster mathematical thinking they needed to provide extended exploratory projects in which students could follow their own lines of enquiry, in line with suggestions in many national curricula, including ours. About half the teachers supplemented these with procedural episodes to develop necessary skills through exposition and exercise. The common use of sequences of conceptually focused microtasks, replacing both these approaches, arose during year 7. It was associated with collaborative thinking among teachers about how students learn particular ideas rather than selecting resources which might, in some generally vague way, lead to learning. We noticed this change of focus in interviews with teachers and also our observations of teachers' meetings. LS started making such shifts during year 7; in SP and FH, some teachers used much more structuring than others from the start, and others changed during year 7 or year 8. The focus that all schools started with—developing ways of working—had led several teachers to worry about 'coverage' and 'basics', so they sought for ways to be more explicit about curriculum topics without losing the emphasis on developing mathematical thinking, inclusion, collaboration and mathematical challenge. They started using fewer tasks in which students could make choices that 'go anywhere' and 'work at their own pace' and more structured work in which students experienced mathematical ideas beyond what arose naturally for them. There seemed to be no conflict, in year 8, between teacher–student interaction shaping what is learnt and attending to the curriculum (Wilson, Cooney & Stinson, 2005).

There were still some lessons in which students worked on extended tasks, and these too could be viewed as sequences of microtasks prompted by teachers. Microtasks were strung together, orchestrated by teacher for the whole group or for individuals, each one influencing what happened next, to accumulate into one extended experience. With the exceptions already mentioned, we saw no differences for most teachers in how this was done in mixed-ability and setted groups. Nor did we find differences between the ability of new and experienced teachers to construct such sequences of microtasks, whereas Leinhardt (1988) had found that expert teachers could weave 'segments and routines' together in more coherent and flexible ways than new teachers. The team approach (Beswick, Watson & De Geest, 2010) may have contributed to our different experience because planning discussions often included talk about the need to scaffold thinking, repertoire, confidence and conceptual understanding. We conjecture that teachers had internalised PLAS' need for frequent, regular and consistent support in these four aspects of mathematical work and that this guided their in-the-moment classroom decisions.

Coherence: How Content Is Made Available in the Lesson

Our third analytical method incorporated many of the aspects that emerged from this analysis but in chronological order so that we could see how teachers constructed mathematical coherence, another feature of the MQAG findings, through strings of microtasks. We developed a tool, an analytical typography, to identify the mathematical actions that are afforded by the lesson and their sequencing (Figure 4). The tool coordinates theories of learning that depend on variation of examples (Marton & Tsui, 2004), mathematical affordances (Greeno, 1994), structural complexity (Biggs & Collis, 1982), mathematical structure (Mason, Graham & Johnston-Wilder, 2005) and exemplification (Watson & Mason, 2005) and names the opportunities to make conceptual sense of experiences during a lesson. The theoretical background of the tool is

Teacher makes or elicits informational/factual statements Information giving Tell/know/ask facts, definitions, techniques 'Research' facts, definitions, techniques *Purposes: remember, prepare* **Learners are expected to** Imitate/follow instructions Use procedure Tell answers and methods *Purposes: fluency, accuracy*	**Teacher asks learners to** Think about ... Use prior knowledge Find answer without known procedure Visualise Seek/break pattern Compare, classify, connect Describe Hypothesise Informal reasoning Give examples with one/many features Application decisions and adaptation *Purposes: personal orientation towards concepts, methods, properties, relationships*	**Integrate and connect mathematical ideas** Associate ideas Generalise Redescribe Summarise ideas Abstraction Formalisation New definition *Purposes: synthesis, connection*
Teacher directs perception/attention Tell/show objects/examples with single/multiple features Tell/show multiple objects Identify properties Classify/compare Identify variation Summarise actions *Purposes: public orientation towards concepts, methods, properties, relationships*	**Discuss implications** Evaluate method(s) Varying variables Special/extreme cases Adapt procedures Identify relationships Induction/ Prediction Explain/ Justification Deductive reasoning *Purposes: analysis, focus on outcomes and relationships, what is interesting about this?*	**Affirm/ act on what we know …** Explore properties Adapt/ transform Apply later Evaluate development Prove *Purposes: rigour, use, objectification*

Figure 4. Analytical frame for mathematics lessons. The columns are not significant; the tool is one list

described in more detail in Watson (2007). The actions are organised in groups which relate to a set of learning purposes, for example: remembering, preparing for use; being accurate; having attention drawn to concepts; acting on concepts; synthesising; objectifying and many others. The list is first used to identify and record the sequence of microtasks that makes up the lesson as a trajectory of lines criss-crossing the list; the sequence is then written out in chronological order to show how the lesson proceeds.

The lessons in Table 3 were taught by different teachers on different topics. Our method strips away content-specific matter to show the structures of pedagogic interaction which are intended to promote conceptual understanding through mathematical reasoning. In this way, we show how the teacher sequences microtasks and elicits and orchestrates students' individual and collective engagement with concepts. The sequences expose the nature of mathematical engagement in the lessons, so that we are able to notice general, but still mathematical, features of how teachers construct such engagement. For this paper, we have included some indications of content so that readers can access these descriptions more easily, but for our work we omitted these. Each item can be seen as a microtask for the learner. For example, in lesson 1, the first microtask offers an organiser for the lesson for the learner to adopt; the second microtask is to bring pertinent definitions and facts to mind ready for what follows; the third is to interpret what the teacher is offering and listen to the interpretations of others, then to spot some properties, then to make up their own examples, then to think about a more complex example and so on. This method of analysis allowed us to see what is being made available for the learner to experience, exactly how the teacher expects learners to engage with mathematical objects and ideas.

This method showed how teachers generated effortful engagement by scaffolding students' thinking, repertoire, confidence and structural understanding. By the end of year 8, all except two teachers were weaving whole-class orientations towards particular aspects of a mathematical idea through drawing attention to single or multiple examples and features of them, summarising and expounding on the implications of work done, relating it to other work, engaging students in public exemplification, explanation and hypothesis, as well as developing fluency and recall. They created and used environments in which students could continually match their constructions of meaning to other examples, methods and definitions coming from the teacher, other students or empirical experience. Their teaching was similar for all groups they taught. The other two teachers remained stuck in the top left-hand aspects

TABLE 3

Examples of lessons showing the sequences of microtasks

Lesson 1	*Lesson 2*	*Lesson 3*	*Lesson 4*
T says what this lesson is about; how it relates to last and next lesson	T gives object with multiple features; asks students to think about how they relate (symmetries of polygons)	T introduces the idea (of 'equivalent equations', e.g. $3x + 2 = 17$; $2x - 2 = 8$)	T tells students what to think about (choose likely outcomes of spinners)
Recap of definitions, facts (% change)	T indicates a classification via variation (number of axes of symmetry; order of rotational symmetry)	T introduces one example; students make examples with certain characteristics	T gives procedural information: how to play game and an example; students follow procedures (to play)
T introduces new idea and asks for interpretations (expressing percentages over 100)	Students use prior knowledge to identify properties of given diagram (rectangle)	T summarises so far, identifies variables (combinations of operations) in their examples, and compares selected examples	T reminds them: 'think about …' (more or less likely outcomes)
T offers example; students identify its properties	T summarises discussion; gives out grid with a two-way classification	Students do examples (solve equations) made by others and compare methods	T associates ideas (compares likelihood with different spinners)
T gives examples with multiple features; students identify properties of them (repeated application of % change)	Students create objects with multiple features; classify them on the grid; make combinations of properties (polygons with order of rotation n; and m axes symmetry)	T leads public deduction of how methods relate (comparing universal methods and case-specific methods)	Students compare and classify results (for different spinners); informal deduction
Students have to produce examples with several features (particular effects of % change)	Students induce generalisation and new definitions (regular and impossible polygons)	T summarises ideas, and shows application to more variables	T leads public exploration of variables and prediction processes
Three concurrent tasks for individuals and small groups	Students transform existing ideas in light of this new evidence ('regular' and 'symmetrical' are not the same)	Students work in groups to express the meaning (of equations) in own words	T compares; asks for justification
describe properties in simple cases	Predictions and conjectures are written on the public board		Students vary cases to justify; (suggest other designs)
describe properties in complex cases	Informal inductive reasoning		Students summarise in public statements
create own cases	T indicates association of ideas and possible generalisations		
T varies variables deliberately (to increase and then decrease)	Students have to compare examples, (relations and inverses among % changes)		
	explain and justify their comparisons and conjectures		
Students classify examples and identify relationships (those that end up more, less, the same)			
T asks about concepts and properties (relations and inverses among % changes)			

of the list in Figure 4. The third teacher who in some videos was not using the whole list developed to do so by the end of year 8.

We have already said that most teachers avoided simplifying mathematics. This final way of analysing lessons shows how learners' actions and responses played a central role in generating examples and ideas. Reasoning of various kinds, including justification and deduction, is evident (e.g. in lessons 3 and 4). The engagement of learners in the coherent development of mathematical ideas can also be seen in each example. Nearly all lessons contained discussion episodes in which teachers used learner responses to develop shared interactions, language and meanings, while teachers directed learners towards salient features of examples and methods. We have not shown examples of lessons in which all that happened was a zig-zag sequence between early categories—recall, instructions and telling answers—but there were five of those.

Differences between lessons analysed at this fine-grained level were usually due to the mathematical purpose. For example, lessons for becoming more fluent with a procedure necessarily focused on procedural work, although most teachers would then have some kind of review of the implications, relating methods to other mathematical ideas, or comparing different methods. Lessons whose purpose was to compare representations of data would necessarily involve much interpretation and discussion of implications. Lessons whose purpose was to continue some ongoing extended exploration would necessarily involve pattern seeking, expressions of relations, hypothesizing and testing in new situations. Nearly all lessons, for all students, whatever the overt shape, included teaching that focused on coherent development of ideas, through experience, reasoning and critical reflection. It was only at this final level of analysis that differences between mathematics specialist and non-specialist teachers showed up. In all schools, a few PLAS who were taught by non-specialist teachers were not offered opportunities to reason deductively, discussions of implications or connections within mathematics. In other words, use of a wide variety of forms of engagement to create a coherent mathematical experience did not guarantee the presence of a strong, explicit, mathematical rationale. This tool helped us to see that nearly all teachers had developed skills in enticing students into mathematically authentic ways of thinking and engaging with mathematical concepts—we conjecture that these skills contributed not only to improved results but also to the relatively positive attitudes of students to their own learning. It also helped us to see how non-specialist teachers, whose lessons looked like everyone else's in generic features of effective teaching, and in terms of mathematical ideas, complexity and coherence, were less able to communicate how the ideas they were working on related to the subject in its fullest sense.

This finding is separate from the comments about two teachers above. The two teachers whose practice was most consistently different from the general descriptions immediately above were sharing the teaching of about half the PLAS in school FH in year 8, the school in which PLAS' achievement fell. In the study reported by Boaler et al. (2000), changes in teaching approach were commonly observed when schools changed their grouping methods, and less effective teachers were allocated to lower attaining groups. Two of our schools managed to avoid that kind of change, and in FH, it was only two teachers, who shared the teaching of one group, who taught PLAS differently[4]. In year 9, these teachers were replaced.

Final Conjectures

We can only conjecture, due to the nature of naturalistic enquiry and the complexities of the teaching–learning relationship, that the common features we have identified 'caused' stronger learning of PLAS or for other students. However, this way of looking at lessons—focusing on how conceptual development is structured through sequences and collections of microtasks which scaffold thinking, repertoire, confidence and structural understanding—illuminates the constructions of complexity and coherence in mathematics lessons that were associated with improved performance on tests accompanied by the maintenance of positive attitudes. Further, we conjecture that this level of analysis is more informative for mathematics teachers than generic descriptions of good teaching, which may be culturally bound, or question types, task types and problem types which do not focus on the effortful thinking students are expected to do.

Acknowledgements

This study was funded by Esmee Fairbairn Foundation, grant number ED 05-1638. The views expressed are those of the researchers and not of the foundation. We are grateful to the schools and teachers who allowed us to research their changing practice and to the anonymous reviewers who helped us communicate this account of an aspect of a complex research project.

Notes

[1] Archives, which have also been used for member checks, are available at www.cmtp.co.uk.

[2] Department practices have been reported elsewhere (Beswick, Watson & De Geest, 2010) and can also be accessed at the website.

[3] Such guidance is widely available in English schools and its use is encouraged by the inspection regime.

[4] At the time of writing, we know the FH end-of-school national exam results: All students who were still attending school achieved a mathematics grade; 72% achieved grades in the highest range (20% above the national average), and this was a 10% increase over the previous year.

References

Andrews, P. (2007). Mathematics teacher typologies or nationally located patterns of behaviour? *International Journal of Educational Research, 46*, 306–318.

Askew, M., Hodgen, J., Hossain, S. & Bretscher, N. (2010). *Values and variables: Mathematics education in high-performing countries*. London: Nuffield Foundation.

Beswick, K., Watson, A. & De Geest, E. (2010). Comparing theoretical perspectives for describing mathematics departments: Complexity and activity. *Educational studies in mathematics, 75*(2), 153–170.

Biggs, J. & Collis, K. (1982). *Evaluating the quality of learning: the SOLO taxonomy*. New York: Academic Press.

Boaler, J. & Brodie, K. (2004). The importance, nature and impact of teacher questions. In D. E. McDougall & J. A. Ross (Eds.), *Proceedings of the 26th Annual meeting of the North American International Group for the Psychology of Mathematics Education* (pp. 773–781). Toronto, Japan: University of Toronto.

Boaler, J., Wiliam, D. & Brown, M. (2000). Students' experiences of ability grouping—disaffection, polarisation and the construction of failure. *British Educational Research Journal, 26*, 631–648.

Brown, M., Brown, P. & Bibby, T. (2007). "I would rather die": Attitudes of 16-year-olds towards their future participation in mathematics. *Paper presented at annual conference of British Educational Research Association*. Available from http://www.lancs.ac.uk/fass/events/mathematicalrelationships/comments/documents/Margaret%20BERA%202007.doc. Accessed 25 Nov 2009.

Day, C., Sammons, P. & Kington, A. (2008). *Effective classroom practice: A mixed method study of influences and outcomes: Full research report, ESRC End of Award Report, RES-000-23-1564*. Swindon, UK: ESRC.

Dorman, J., Adams, J. & Ferguson, J. (2002). Psychosocial environment and student self-handicapping in secondary school mathematics classes: A cross-national study. *Educational Psychology, 22*, 499–511.

Greeno, J. (1994). Gibson's affordances. *Psychological Review, 101*, 336–342.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon, UK: Routledge.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video* *study (NCES 2003–013)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Ireson, J., Hallam, S. & Plewis, I. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts. *British Journal of Educational Psychology, 71*, 315–326.

Krainer, K. (2005). What is 'good' mathematics teaching, and how can research inform practice and policy? *Journal of Mathematics Teacher Education, 8*, 75–81.

Leinhardt, G. (1988). Expertise in instructional lessons: An example from fractions. In D. Grouws, T. Cooney & D. Jones (Eds.), *Perspectives on research on effective mathematics teaching* (pp. 47–66). Reston, VA: National Council of Teachers of Mathematics.

Lincoln, Y. & Guba, E. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.

Marton, F. & Tsui, A. (2004). *Classroom discourse and the space of learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mason, J., Graham, A. & Johnston-Wilder, S. (2005). *Developing thinking in algebra*. London: Paul Chapman.

Mullis, I., Martin, M. & Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Nardi, E. & Steward, S. (2003). Is mathematics T.I.R.E.D.? A profile of quiet disaffection in the secondary mathematics classroom. *British Educational Research Journal, 29*, 345–367.

Ofsted (2005). *Mathematics in secondary schools 2003/4*. Available from www.ofsted.gov.uk/publications/annualreport0304/subject_reports/secondary/maths.htm. Accessed 25 Nov 2009.

Pietch, J., Walker, R. & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology, 95*, 589–603.

Sternberg, R. (1986). Toward a unified theory of human reasoning. *Intelligence, 10*, 281–314.

Sturman, L., Ruddock, G., Burge, B., Styles, B., Lin, Y. & Vappula, H. (2008). *England's achievement in TIMSS 2007: National report for England*. Slough, UK: NFER.

Watson, A. (2007). The nature of participation afforded by tasks, questions and prompts in mathematics classrooms. *Research in Mathematics Education 9* (111–126).

Watson, A. & De Geest, E. (2005). Principled Teaching for Deep Progress: Improving Mathematical Learning Beyond Methods and Materials. *Educational Studies in Mathematics, 58*, 209–234.

Watson, A. & Mason, J. (2005). *Mathematics as a Constructive Activity: Learners Generating Examples*, Mahwah, NJ: Lawrence Erlbaum Associates.

Wiliam, D., Brown, M. & Boaler, J. (1999). 'We've still got to learn': Low attainers' experiences of setting. *Equals, 5*, 15–18.

Wilson, P., Cooney, T. & Stinson, D. (2005). What constitutes good mathematics teaching and how it develops: Nine high school teachers' perspectives. *Journal of Mathematics Teacher Education, 8*, 83–111.

Anne Watson

Department of Education, University of Oxford, 15 Norham Gardens,
Oxford OX2 9JZ, UK
E-mail: anne.watson@education.ox.ac.uk

Els De Geest

The Open University, Buckinghamshire,UK