



Report

Final Unit 1 Report for
SRI International

Evaluation of Cornerstone Mathematics Pilot in England: Unit 1, linear functions

Linda Sturman
Louise Cooper

January 2012

With minor revisions, March 2012

National Foundation for Educational Research,
The Mere, Upton Park, Slough, Berkshire SL1 2DQ
www.nfer.ac.uk

© National Foundation for Educational Research 2012
Registered Charity No. 313392

How to cite this publication:

Sturman, L., Cooper, L., (2012). *Evaluation of Cornerstone Mathematics Pilot in England: Unit 1, linear functions*. Slough: NFER.

Research Team

| | |
|---------------|---------------------|
| Sarah Maughan | Project Director |
| Linda Sturman | Project Leader |
| Louise Cooper | Researcher |
| Pat Bhullar | Project Coordinator |

Acknowledgements

With thanks to Phil Vahey, Nicole Shechtman and colleagues at SRI International and to Teresa Smart, Phillip Kent and colleagues at London Knowledge Lab.

Contents

| | | |
|----|--|----|
| 1 | Summary | 1 |
| | 1.1 Introduction | 1 |
| | 1.2 Key Findings | 2 |
| 2 | Recommendations | 7 |
| 3. | Introduction and Methods | 10 |
| | 3.1 Background | 10 |
| | 3.2 The Research Questions | 10 |
| | 3.3 The Pilot, Unit 1 | 11 |
| | 3.4 Context and Methods | 11 |
| | 3.5 Context for the Evaluation | 13 |
| | 3.6 The Findings | 14 |
| 4 | Evaluation Findings | 15 |
| | 4.1 Participation and Expectations | 15 |
| | 4.2 Delivery and Effectiveness of the Unit | 17 |
| | 4.3 Effectiveness for Different Groups | 28 |
| | 4.4 Curriculum Fit | 31 |
| | 4.5 Engagement and Motivation | 32 |
| | 4.6 Elements of the Unit | 38 |
| | 4.7 Mathematics in the 'Everyday' World | 43 |
| 5. | Conclusion | 45 |

1 Summary

1.1 Introduction

The NFER, under contract to SRI International (SRI), is evaluating the effectiveness of the 'Cornerstone Maths' pilot in England. The pilot involves trialling two technology-based mathematics units in a small number of key stage 3 (KS3) classrooms in England and, if successful, will be used to inform the potential future roll out of a larger number of Cornerstone units to a larger number of schools. The aim of the pilot is to test how effectively the innovative approach transfers to England from the US context. The approach integrates curriculum, technology and teachers' professional development. The first unit, known as SandCircle Mobile Games, teaches linear functions through a set of related scenarios based on developing games for mobile phones. It uses interactive computer technology to simulate real world mathematics and to support students in making connections between tables, graphs, equations and narrative. It was taught between October and November 2011. A second Cornerstone unit will be piloted in 2012, exploring geometric similarity.

The first unit was begun in 10 schools and completed in nine (one could not complete due to unforeseen circumstances). All schools were recruited by London Knowledge Lab (LKL) in liaison with the Department for Education and SRI. Whilst the small sample for this initial study cannot be considered statistically representative, schools were selected to represent the variety of school contexts in the English education system, including schools with differing demographics, Local Authority maintained schools, the independent sector and academies.

This first report is based on quantitative and qualitative research methods that are standard for this type of small scale exploratory evaluation. These were: a pro forma completed by the main contact in each school; lesson observations; post-observation interviews with teachers; focus group discussions with students; responses to a student questionnaire; and data from students' responses to a test taken before the unit was taught and repeated after the unit had been completed. A summary of key findings from the evaluation follows.

It should be remembered that participants who agree to take part in a pilot are 'early adopters' and likely to be motivated by innovation and change to a greater extent than might necessarily be the case for the wider population. Hence, it is important to bear in mind that, even though the schools in the pilot sample represent a cross-section of the types of schools in England, the pilot teachers may differ in key respects from the general teaching population. A wider roll-out of the Cornerstone approach must take this into account.

1.2 Key Findings

The evaluation addressed seven research questions (RQ1 to RQ7). They are listed below, with a summary of the outcomes for each (or for a pair of related questions in the case of the first two) for Unit 1. The questions are presented in the order in which they are covered in the body of the report. More details of the evidence in each area can be found in section 4.

RQ1. How were the units delivered (implemented) by teachers in each school? Which models of delivery were deemed effective? AND

RQ2. Do teachers and students believe that the materials help in teaching/learning the mathematics topic targeted by the unit?

Overview: Research Questions 1 and 2

The unit was generally manageable to implement and it helped students to learn the difficult concepts covered.

Ahead of the pilot, teachers' expectations of success for the unit were high, although they reported facing a variety of challenges in teaching mathematics to their KS3 students, and had some prior concerns about the technology, the time needed for teaching the Cornerstone lessons and teaching unfamiliar material. Prime reasons for participation related to wanting to improve learning and providing professional development for teachers. Personal contact was important in persuading schools to take part in the pilot, whether or not teachers knew the person making the contact. Contextual factors such as these will have implications for scaling up Cornerstone Maths in England.

Unit 1 was generally delivered in a consistent style, following the common lesson structure in England of starter/recap activity – main activity – plenary, and with the teacher tending to facilitate and/or set the pace. Teachers worked collaboratively within their schools and shared their experience as the unit progressed. They found implementation of the unit manageable and preparation for it no more time-consuming, and sometimes quicker, than their usual teaching of this topic. They felt well-supported by the LKL and SRI teams and found the introductory 'CPD' training session useful, although some wanted less time working through the materials as the students would, and requested more time on planning for implementation in their own context. Individual teachers had different training needs.

Teachers encouraged students to work collaboratively and this aided understanding, although researchers observed that, when working in pairs, a dominant partner tended to control the computer. Some younger students struggled with reading and understanding the tasks. Teachers generally encouraged students to view problems in more than one way.

Both teachers and students believe that the unit helped learning, despite covering a complex topic which students found challenging and in which they, inevitably,

struggled to grasp some concepts at first. In particular, teachers believe that the Cornerstone approach led to deeper learning of these challenging and abstract concepts than would have been the case with their usual approach. Similarly, students believe that SandCircle is helpful, for example, it “*makes them think*” and helps them understand different ways of solving a problem. A few reported liking the element of repetition in the unit as it supports their learning, although some find it unhelpful (see RQ3). Test results confirmed that learning had occurred, with improvements overall as measured by pre-test and post-test results and with more improvement on the ‘M2’ (more complex) components than on the ‘M1’ (foundation) components.

RQ6. What are the perceived impacts on learning using these units in terms of different groups of students (e.g. ability, gender, ethnicity or socioeconomic status - SES)?

Overview: Research Question 6

Perceptions of groups on which the unit impacted most were varied, but test results suggest that the unit is appropriate and effective for the whole range of students, with no group conclusively progressing better than others.

Teachers believed that the unit encouraged effective, ‘deep’ learning but they differed in their perceptions of the students for whom the unit was most effective. Some thought that there were no differences between groups, while others felt that there were differences. Various, these teachers deemed the unit more effective for: those with prior experience of the topic; the less able; the more able; boys; girls; and those who like working on computers. It is important to remember that these are perceptions and that some perceived effects might be conflated or might be specific to individuals or to the context, rather than general effects.

The test results confirmed that learning had taken place. However, they showed few differences across groups of students. This might have been because no effects existed. Alternatively, it is possible that effects perceived by teachers were not sufficiently strong to affect test outcomes, or were not visible because of conflating factors. Piloting with a larger sample would allow further investigation of these hypotheses. Based on the pre- and post-test results, the only apparent difference in gains made during the unit was by gender. Girls made a slightly greater gain compared with boys (a small but significant effect). This might be a spurious effect or it might indicate that the Cornerstone materials may help to narrow the gender gap seen for the pilot sample in this topic. It would be useful to investigate this further with a larger sample.

There were no other significant differences in pre-test to post-test gains overall. Students of different ages made comparable gains in their learning, as did students of different ability levels and those in schools with different levels of achievement/progress and with different levels of eligibility for free school meals.

These findings suggest that the approach and materials are appropriate and effective for a range of students.

RQ4. Do teachers believe that the materials help them meet the requirements of the national curriculum and their own mathematics learning goals, and prepare students for their assessments?

Overview: Research Question 4

The unit has good curriculum fit and covers more than most teachers' own schemes of work. Where teachers could make a prediction, they felt that students would do well in their next assessment, based on the students' solid understanding and their willingness to attempt questions.

Teachers feel that the unit fits well with the English national curriculum and, in most cases, covers more than their own scheme of work. They believe that it supports students well in understanding abstract concepts. Only one teacher reported that the unit does not go as far as the school's own scheme of work.

Nine teachers thought that their students would do better or well on their next assessment of their knowledge of linear functions, although some found it difficult to predict how well the unit had prepared students. The nine who felt that their students would do well or better than usual commented on: students' solid understanding of the topic; their expectations of longer-term recall; a need for less revision; and students' increased willingness to attempt questions. Post-test outcomes support the perceived willingness of students to attempt questions but do not allow comment on how well learning may be retained over time.

The pre-test is a key aspect of establishing that learning has taken place but was daunting for students with limited or no prior experience of the topic. A teacher suggested renaming it from 'pre-test' to: 'What do you already know?' to reduce the pressure that students felt when taking it.

RQ5. Do the materials increase student engagement/motivation compared with their general attitude to learning mathematics?

Overview: Research Question 5

Teachers were positive about the unit's ability to engage, although students expressed some reservations. Despite this, students acknowledged that the unit helped them to learn.

The views of teachers were broadly positive, with some reservations expressed by students. Teachers were positive about the software, the materials and the approach, and used a variety of strategies to keep students engaged. They reported a few negative effects on motivation though most felt that students found the unit engaging. Few of the students have significant access to computers for their own use during their regular mathematics lessons and teachers considered that the novelty of using the software was a factor in students' engagement. Further evaluation may help in

establishing the relationship between novelty and learning in this unit in England: whether the perceived novelty supports learning or whether learning through this integrated approach is, as suggested by evaluation in the US¹, independent of the novelty factor.

Over half of students liked their SandCircle lessons and around three fifths of those interviewed think that the lessons improved their ability to learn independently. However, a sizeable number rated the lessons less highly, claiming to enjoy their regular mathematics lessons more, some because they reported finding aspects of the unit repetitive and thus 'boring'. Despite this, as noted earlier, students acknowledged that the unit helped them to learn the key concepts: learning and enjoyment do not always go hand-in-hand².

Most students enjoyed the context of mobile phone games but some felt that it was more appropriate for younger students. In several schools, students were disappointed not to be able to design their own mobile phone games. Motivation may be enhanced by adding this element in a meaningful way.

RQ3. Which parts/aspects of the unit were found to be useful? Which were not so useful?

Overview: Research Question 3

The unit was generally seen as useful, with most parts of it more likely to be rated 'helpful' than 'unhelpful'. Some technical difficulties were encountered, and some changes and additions to resources were suggested.

The unit was felt to be useful generally. Although large proportions of students found some activities less helpful to their learning than others, more students tended to report any given activity 'helpful' than reported it 'unhelpful'. Some students found the language or context of some parts of the unit difficult to understand, and some found the requirements to explain and predict particularly difficult.

Teachers were positive about the software, with most rating it highly. They reported some technical difficulties, relating to the school's ICT system and/or SimCalc. For some, access to infrastructure was an issue. Most used the computer suite or laptops, although some lessons were conducted using an interactive whiteboard

¹ Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J. and Gallagher, L.P. (2010). *Integration of Technology, Curriculum, and Professional Development for Advancing Middle School Mathematics : Three Large-Scale Studies*. American Educational Research Journal, December 2010, Vol. 47, No. 4, pp. 833–878. Originally published online 2 June 2010. <http://aer.sagepub.com/content/47/4/833> Accessed 25 January, 2012.

² See, for example, Mullis, I.V.S., Martin, M.O. and Foy, P. with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., and Galia, J. (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: Boston College, TIMSS & PIRLS International Study Center.

only. Use of laptops simplified lesson set-up. Teachers' own familiarity with the software facilitated their teaching.

Teachers requested some additions to the resources and changes to guidance regarding the length of time needed for the unit. They found the unit too intensive and suggested breaking it down and/or providing resources electronically to make its use more flexible. It was common for them to request resources for differentiated teaching.

RQ7 Do the materials help students see the role of mathematics in the 'everyday' world?

Overview: Research Question 7

Teachers were generally positive about the impact of the unit on students' understanding of mathematics in the real world, although they identified different ways in which the unit achieved this. Students also perceived real-life benefits.

Teachers were positive about the impact of the unit on students' understanding of mathematics in the real world. However, teachers held different interpretations of the concept. Some related 'real-life' mathematics to problems set in realistic contexts, while others interpreted it in terms of students being able to apply learning elsewhere in their mathematics lessons or apply mathematics learning in real life (e.g. understanding a newspaper article relating to mathematics).

Students believe that it is important for the mathematics they do at school to relate to everyday life and 43 per cent think that mathematics in general will be relevant to their own future career, while only 18 per cent think that it will not. The remainder (39 per cent) are not sure. Nevertheless, it is an ongoing issue that students do not always see the links between the mathematics they study and daily life. This is neither an easy nor a quick issue to resolve and it is therefore positive that 70 per cent of the pilot students who completed the questionnaire thought that SandCircle helped them understand how mathematics is used in everyday life, particularly in terms of possible future careers, and that many feel that the unit helped them with problem-solving.

Students in several schools were hopeful of generating their own mobile phone games and the unit would have had more real life relevance for these students had they been able to do so.

2 Recommendations

This section of the report complements the Summary in section 1 and the separate Executive Summary document. It presents recommendations related to the key findings highlighted in those summaries, for each research question in turn. These recommendations relate to implications for the wider roll-out of the Cornerstone Maths pilot in England.

RQ1. How were the units delivered (implemented) by teachers in each school? Which models of delivery were deemed effective? AND

RQ2. Do teachers and students believe that the materials help in teaching/learning the mathematics topic targeted by the unit?

Recommendations: Research Questions 1 and 2

Recommendations specific to the SandCircle unit:

1. In the introductory section of the student workbook, make explicit reference to the level of challenge of the tasks (perhaps in the section ending “*Don’t worry, it will all be clear by the time you are finished!*”, p1).
2. Add sample responses to the guidance materials to help teachers support students in giving explanations.
3. Review the language of the units, to make them accessible to younger students who are capable of accessing the mathematics but might struggle with reading and understanding the tasks.
4. Add materials for differentiated teaching and add starter/recap activities.
5. Amend guidance regarding time needed for the unit and provide materials in a form that facilitates flexible usage in different contexts (e.g. electronic copies of the student workbook).
6. Explore pre- and post-test data further, in particular investigating why there was no improvement for item 2a.

Recommendations for a wider pilot in England:

7. Retain existing support mechanisms and continue to encourage a collaborative approach to teaching and learning.
8. Be aware of the challenges teachers say they face in teaching mathematics generally; consult with them about lessons learned in accommodating these challenges during the pilot and, if applicable, add relevant guidance in the Cornerstone Maths context.
9. Add guidance on managing open-ended learning, to support teachers and learners as they begin engaging with contexts that do not require ‘a right answer’.
10. Provide guidance on effective management of discussion and computer use, for situations where students work individually or in pairs.
11. Consider alternative ways of organising training to meet the range of needs (e.g. parallel workshops to accommodate teachers’ different preferences and needs).
12. In scaling up the pilot, use feedback about the factors that encouraged participation to maximise participation in later trials. Consider what additional support and encouragement less experienced or non-specialist teachers of mathematics might need. Take account of practical constraints such as the availability of computer hardware.

RQ6. What are the perceived impacts on learning using these units in terms of different groups of students (e.g. ability, gender, ethnicity or SES)?

Recommendations: Research Question 6

Recommendations specific to the SandCircle unit:

1. Investigate group performance further, with larger sub-samples: to investigate whether the small gender effect indicated here is real and whether the existence of comparable gains is confirmed, on a larger scale, for all other variables investigated.
2. Based on current data, draft teacher guidance outlining that the unit has similar impact on different groups of students. Expand and/or make it more specific as the pilot continues.

RQ4. Do teachers believe that the materials help them meet the requirements of the national curriculum and their own mathematics learning goals, and prepare students for their assessments?

Recommendations: Research Question 4

Recommendations specific to the SandCircle unit:

1. Consider exploring further, in light of the teachers' view that SandCircle led to deep and long-term learning, the extent to which learning is sustained over time in the English context.

Recommendations for a wider pilot in England:

2. Given that some students may have little or no prior experience of a Cornerstone topic, consider following the suggestion of one of the teachers to rename the pre-test to something more supportive for students, such as: 'What do you already know?'.

RQ5. Do the materials increase student engagement/motivation compared with their general attitude to learning mathematics?

Recommendations: Research Question 5

Recommendations specific to the SandCircle unit:

1. Consider an extension activity allowing students to design their own mobile games.
2. Review the materials in light of students' perceptions of repetitiveness (see RQ3 below). Consider making it explicit in the student materials how elements that they may perceive as 'repetitive' can aid their learning.
3. Consider guidance for teachers on how to maintain the attention of students who disengage, so that they can benefit fully from the learning potential of the SandCircle unit.
4. Conduct further evaluation of the unit to explore the relationship between novelty and learning in England: whether the perceived novelty supports learning or whether learning is, as suggested by evaluation in the US, independent of the novelty factor.
5. Investigate further the finding that a number of students did not rate their SandCircle lessons as highly as their usual mathematics lessons, despite learning from the unit. Establish whether this is a general effect or specific to this group of students. If the former, provide guidance to minimise the effect.

RQ3. Which parts/aspects of the unit were found to be useful? Which were not so useful?

Recommendations: Research Question 3

Recommendations specific to the SandCircle unit:

1. Retain the current activities, since they were generally rated useful, but review the materials to make them appear less repetitive (e.g. by combining some activities or adding activities that do not use the software).
2. Where perceptions of repetitiveness may be unavoidable (e.g. where the context is related but the mathematics is different), make it explicit in teacher guidance how repetition will help learning and emphasise this in the CPD training session for teachers. This corresponds to a recommendation made for the student materials (see RQ5).
3. Amend the layout of the student booklet so that the 'predict' section is separate from the 'explain' section. This will support students in developing their prediction skills whilst reducing their fear of recording a 'wrong' prediction next to their correct explanation.
4. Add clarification to sections where the context might not be familiar to all students.
5. Emphasise the benefits for learning of allowing students to use the software themselves during this unit. However, also acknowledge potential difficulties in accessing ICT resources and identify the activities that might best be done using an interactive whiteboard if absolutely necessary.

Recommendations for a wider pilot in England:

6. Add guidance and trouble-shooting information on integrating the software with systems commonly in use in schools in England.
7. Act upon relevant suggestions for resolving minor issues with the SimCalc software. Review teacher guidance on managing the software and related technical issues, to allay any prior fears. Consider branding this as 'an introduction to using SimCalc'.

RQ7 Do the materials help students see the role of mathematics in the 'everyday' world?

Recommendations: Research Question 7

Recommendations specific to the SandCircle unit:

1. In guidance notes and the student workbook, make examples of 'real maths' and the relevance of the unit to students' current and/or future lives explicit, enabling a shared understanding among teachers and students, as it applies to the unit.

3. Introduction and Methods

3.1 Background

Based on research in the US, SRI International (SRI) has shown that a combination of professional development activities for teachers, along with SRI's digital mathematics materials, can improve students' learning in mathematics³. In response to a request from the Li Ka Shing Foundation, SRI, in liaison with London Knowledge Lab (LKL), is conducting a pilot to investigate how effectively its proven mathematics teaching materials transfer to the English context. The effectiveness of the pilot is being independently evaluated by the National Foundation for Educational Research (NFER), under contract to SRI.

The main purpose of the evaluation is to investigate the effectiveness and impact of two technology-based mathematics units in key stage 3 (KS3) classrooms in schools in England. The materials are known as Cornerstone Maths in England. The study is part of a wider project and will inform the potential future roll out of a larger number of Cornerstone units to schools. Evidence of successful implementation will be demonstrated through improved learning, as measured by the extent to which a number of agreed positive impact measures are achieved in the pilot schools (based on both hard data and the views of teachers and students).

3.2 The Research Questions

The following research questions are addressed in the evaluation study:

1. How were the units delivered (implemented) by teachers in each school? Which models of delivery were deemed effective?
2. Do teachers and students believe that the materials help in teaching/learning the mathematics topic targeted by the unit?
3. Which parts/aspects of the unit were found to be useful? Which were not so useful?
4. Do teachers believe that the materials help them meet the requirements of the national curriculum and their own mathematics learning goals, and prepare students for their assessments?
5. Do the materials increase student engagement/motivation compared with their general attitude to learning mathematics?
6. What are the perceived impacts on learning using these units in terms of different groups of students (e.g. ability, gender, ethnicity or socioeconomic status - SES)?

³ Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J. and Gallagher, L.P. (2010). *Integration of Technology, Curriculum, and Professional Development for Advancing Middle School Mathematics : Three Large-Scale Studies*. American Educational Research Journal, December 2010, Vol. 47, No. 4, pp. 833–878. Originally published online 2 June 2010. <http://aer.sagepub.com/content/47/4/833> Accessed 25 January, 2012.

7. Do the materials help students see the role of mathematics in the 'everyday' world?

3.3 The Pilot, Unit 1

The Cornerstone materials were originally developed for 7th and 8th grade students (ages 12-14) in the US. The first Cornerstone Maths unit was piloted in England in October and November 2011. The earliest pre-tests took place in the week beginning 3 October, and the last post-tests in the week ending 25 November. Schools were free to schedule their testing and teaching as they chose within that period. A second unit will be piloted in 2012, exploring geometric similarity. The first unit, known as SandCircle Mobile Games, teaches linear functions through a set of related scenarios based on developing games for mobile phones. As part of the pilot, it was taught in 10 schools, all of which were recruited by London Knowledge Lab (LKL) in liaison with the Department for Education and SRI. A limited number of schools was chosen for this initial exploratory study for Unit 1, and the results are not intended to be statistically 'representative' of the broader school population in England. However, sample schools were selected to represent the variety of contexts in the English education system, including comprehensive schools, Academies and an independent school. In each school, one or two classes participated, drawn from Years 7 to 9 (ages 11-14) and including students of differing abilities. This allows some conclusions to be drawn about the use of Unit 1 in the range of school contexts found in England. Due to unforeseen circumstances, one school was unable to complete its participation in the first unit and so is not included fully in this stage of the evaluation.

The focus of the pilot is to investigate how effectively the materials transfer from one context (the US) to another (England). The findings will be used to inform the potential roll-out of a wider number of units to a wider number of schools. As a result, piloting of Unit 1 (and hence the evaluation) focused only on the sample of schools and classes trialling the materials; it did not include any element of comparison with a 'control' sample of schools or classes taught the same topic by other methods, since that was not directly relevant to the aims of the pilot. The scope and size of the sample was agreed by the pilot team in liaison with the pilot steering group.

3.4 Context and Methods

This report focuses on outcomes following the teaching and assessment of the first pilot unit. It reports outcomes from a series of research exercises, using standard research methods for this type of small scale exploratory evaluation. These included quantitative methods to provide broad information across the student sample as a whole, and qualitative methods to provide depth of detail in key areas. In summary the research methods were:

- a pro forma completed by the main contact in each of the 10 pilot schools prior to the evaluation visit;
- observation of a lesson with one or two participating classes in each of the nine schools that completed the unit;

- post-observation interviews with the teacher(s) of the observed class(es) in nine schools;
- students' responses to a questionnaire completed at or towards the end of the unit (eight schools completed and returned);
- a focus group discussion with a sample of students from each observed class (nine schools);
- pre-test and post-test data, from students' responses to a test taken before the unit was taught, and repeated after the unit had been completed (nine schools).

The pro forma

The brief pro forma was completed by the researcher in discussion with each school's main contact, generally prior to the evaluation visit to their school but in a few cases on the day of the visit. It was designed to establish a baseline 'context for teaching' indicating the reasons why staff in these schools chose to participate in the pilot, background information about mathematics teaching in the schools, expectations for the pilot and any challenges anticipated. It was designed so that the second phase of the evaluation could potentially gather teachers' views regarding parallel questions at the end of the pilot, comparing earlier expectations with post-pilot views. All 10 schools provided a completed pro forma (including the school which could not complete the unit) and all were included in analysis.

Lesson observations

Researchers from NFER and LKL observed 16 lessons across nine schools. During each evaluation observation, the researcher took notes using a semi-structured schedule. The observations covered three main themes:

- manageability (how the unit transferred to schools in England, and how teachers controlled the lesson content and pace);
- engagement/behaviour (attitude, engagement and understanding, behaviour and interaction);
- teaching style/approach (interaction/intervention, use and application of mathematics).

Post-observation interviews

The 16 observed teachers in nine schools participated in an interview, covering the participation of 17 classes (one teacher talked about his own and a colleague's experience). The aim of the interviews was to collect relevant information about: the teachers and students observed; preparation for teaching the unit; the implementation of the unit; the technology; and the impact on students in terms of their engagement and learning.

Student questionnaire

Students in the observed classes were asked to complete a questionnaire. This asked for their views about mathematics generally and about the SandCircle unit in particular. Most teachers administered the questionnaire to their classes after the day

of the observation, closer to the end of the unit. Questionnaires were completed by students in eight schools.

Focus groups

A sample of students from each observed class also took part in a focus group interview. Teachers selected 6-8 students for each focus group and were advised to provide a mix of students (e.g. of different ability levels or different language backgrounds). Students from nine schools participated in the focus groups.

Test data

Students completed two identical tests: a pre-test, before the unit was taught, and a post-test, taken at the end of the unit. Nine schools and 17 classes took part in the testing, and 429 students took both the pre- and post-test. The test incorporated questions on two curriculum areas: M1 and M2. On the whole, M2 questions are of a greater difficulty level than M1 questions. The difference between M1 and M2 questions is summarised in Table 1 below.

Table 1: M1 and M2 components⁴

| M1 Component <i>Foundational concepts typically covered in the grade-level standards, curricula, and assessments</i> | M2 Component <i>Building on the foundations of M1, essentials of concepts of mathematics of change and variation found in algebra, calculus, and the sciences</i> |
|---|--|
| <ul style="list-style-type: none"> ➤ Categorizing functions as linear/nonlinear and proportional/non-proportional ➤ Within one representation of one linear function (formula, table, graph, narrative), finding an input or output value ➤ Translating one linear function from one representation to another | <ul style="list-style-type: none"> ➤ Interpreting two or more functions that represent change over time, including linear functions or segments of piecewise linear functions ➤ Finding the average rate over a single multirate piecewise linear function |

3.5 Context for the Evaluation

The observed teachers are generally experienced mathematics specialists. Most (eight) are in their fifth to ninth year of teaching. Four have more than ten years' experience and three are in their fourth year or less. Fourteen have a mathematics or mathematics-related degree, while two are not mathematics specialists but have done or are planning to do a conversion course.

Of the 17 classes, eight were year 7 (Y7, age 11-12) and eight were Y8 (age 12-13). One teacher taught the unit to a Y9 class (age 13-14). The target grades for the unit

⁴ Source: Shechtman, N., Haertel, G., Roschelle, J., Knudsen, K., & Singleton, C. (2010). *Design and development of the student and teacher mathematical assessments* (SimCalc Technical Report 05). Menlo Park, CA: SRI International. (p5).

are US 7th and 8th grade (England Y8 and Y9 equivalents) so almost half of the pilot teachers in England had chosen to pilot the unit with students slightly younger than those for whom it was developed. Most of the Y7 classes were higher ability classes, but one was the lowest 'set' in the year group. The Y8 classes were generally of middle ability. It is important to bear in mind, however, that the criteria for setting students according to their ability levels may vary across schools and according to the school's intake. There was a range of attainments and ethnicities within the observed classes; some students in the classes had Special Educational Needs (SEN) or spoke English as an Additional Language (EAL). See Appendix A in the separate Appendices document for more information about the context of the observations and Appendix B for sample characteristics.

Practical considerations guided the choice of pilot classes. They were selected based on three predominant themes: the sets or combinations of sets with which teachers wanted to trial the materials; timetabling issues, and teacher/class match. Access to ICT hardware was also an issue.

Teachers typically reported having 3 to 4 hours a week with each selected class (12 of those who answered). Two had only 2 hours, 30 minutes per week. At the time of the observations, all were at or past the half way point of teaching the planned Unit 1 (SandCircle) lessons. The way in which the lessons had been planned and/or organised varied across the schools.

Few of the pilot students have significant access to computers for their own use during their regular mathematics lessons. The least frequent access was once each half-term or less often (five teachers), while only four said that their students could access computers in their mathematics lessons more than once a month. Only two reported access in most lessons and for one, this was for assessment not teaching. Teachers more commonly use an interactive whiteboard (IWB), which students sometimes use directly. Hence, the degree of technology involved in the SandCircle unit was new to most of the pilot students.

3.6 The Findings

Section 4.1 below describes the context of the pilot lessons, while sections 4.2 to 4.7 describe the outcomes from the research exercises. These outcomes informed the Summary and Recommendations presented earlier in Sections 1 and 2, and the separate Executive Summary document.

4 Evaluation Findings

4.1 Participation and Expectations

The main contact in each school identified the reasons why that school chose to participate, challenges they face in teaching mathematics to their key stage 3 (KS3) students and their aspirations and expectations for the unit. These are outlined in sections 4.1.1 to 4.1.5 below and provide the context for the evaluation findings that follow in sections 4.2 to 4.7.

4.1.1 Reasons for participating

Schools invited to participate had been identified as part of a targeted sample meeting several key criteria (see sections 1.1 and 3.3). The main contacts most commonly cited personal contact as the reason why staff in their school elected to participate in the pilot, regardless of whether the teachers concerned knew the person making the contact. Their positive responses arose following invitations from London Knowledge Lab (LKL), a numeracy consultant or other consultant connected with their school, or from an unidentified source. One contact mentioned that respect for the work of a member of the LKL team was a deciding factor. The motivation in two schools was to improve the use of Information and Communication Technology (ICT) in mathematics lessons and another contact saw it as an opportunity for useful Continuing Professional Development (CPD) for the teachers involved. For two respondents, the decision related to the CPD training session for Unit 1 (linear functions): one participated because 'cover' was paid for attendance at the CPD session, and another signed up for the pilot as a result of being impressed by the presentation. Other reasons related to a desire to learn about research in the mathematics world, being interested in 'real maths', and because of a general approach of responding positively to research requests.

4.1.2 Challenges in teaching mathematics to the pilot students

The pro forma asked contacts to specify the challenges they face in teaching mathematics to their KS3 students (those aged 11-14). Respondents identified a range of challenges. Three related to limited motivation among their students or to the need to keep able students engaged, while two referred to language or literacy issues affecting learning in mathematics, and another two to enabling students to see mathematics as functional and relevant to life. Other challenges were specific to individual schools and included factors such as: identifying particular topics that students find difficult; needing to stimulate discussion during lessons and develop independent learners; challenges relating to overly repetitive lessons; and encouraging problem solving, process skills and creativity.

4.1.3 Aspirations for the pilot

Main contacts in the 10 schools identified a range of potential gains they hoped to achieve by participating. Five focused on students' understanding, albeit from different perspectives. One wanted to improve students' understanding and application of mathematics, while another wanted to raise attainment and move students' understanding ahead of their chronological age. Helping students to make connections in mathematics was mentioned, as was motivating learners and exploring methods that might suit particular groups of learners (such as 'kinaesthetic' learners). One contact felt that it was important to spend time on Unit 1 as it underpins other mathematical concepts and another wanted students to become more independent learners, able to relate their learning to their daily lives. For three contacts, their aspirations for the pilot related to developing mathematics teaching within their school: exploring different ways of delivering mathematics, working with new organisations and/or technology to enhance mathematics teaching and developing teachers' skills. Staff in one school reported that they might extend their pilot activity by adding a 'control class' in their school in future.

4.1.4 Anticipated challenges

The contacts anticipated some challenges in implementing the pilot Cornerstone units. This question was asked in advance for only seven of the 10 schools: opportunity for discussion ahead of the visit was limited in the remaining three schools. Some teachers raised more than one challenge. Teachers in six schools perceived challenges centred on the technology. These were: issues of software compatibility and initial problems with setting up the technology; decisions over whether to use PCs or laptops; and the logistics of managing multiple laptops in the classroom or getting access to the school's computer suites. Other concerns related to potential time issues in preparing the Cornerstone lessons (two schools) and uncertainty about how to teach the new material (one school).

4.1.5 Expectations

In conclusion, researchers asked the main contact in each school to give a rating to each of three statements using a scale of 1 to 10.⁵ The main contacts had generally high expectations, answering as follows.

a) How positive are you about the likelihood of the Cornerstone approach improving your pupils' understanding of linear functions (compared with your usual approach to teaching this topic)? Ratings were broadly positive (one rating of 5 and seven ratings above 5), with an average rating of 7.4 from the eight teachers who responded. Four of the ratings were at or above 8. Reasons for ratings across the range related to the materials (positive comments that they encourage students to make connections and

⁵ where 1 is not positive and 10 is highly positive

explain their answers) or to external limiting factors (such as students not being ready or being unfamiliar with the computer suite).

b) How positive are you about the likelihood of the Cornerstone approach improving your pupils' understanding of geometric similarity (compared with your usual approach to teaching this topic)? Again, ratings were generally positive but showed a narrower range than for the previous question (perhaps because teachers had already seen Unit 1 (linear functions) but not yet seen Unit 2 on geometric similarity). Their ratings ranged from 6 to 9 with an average of 7.7.

c) How positive are you about the likelihood of the Cornerstone approach improving your pupils' engagement with mathematics in general? This was the most positive outcome, with an average rating of 7.8 albeit with a wider range. Most responses were within the range 5 to 9, although one teacher gave a provisional rating of 1 (alongside an overall rating of 6 and thus discounted in calculating the average) based on the fact that one student had already disengaged.

4.2 Delivery and Effectiveness of the Unit

This section addresses the first two research questions:

RQ1. How were the units delivered (implemented) by teachers in each school?
Which models of delivery were deemed effective?

RQ2. Do teachers and students believe that the materials help in teaching/learning the mathematics topic targeted by the unit?

4.2.1 Planning, preparation and support

CPD training session

All teachers involved in the pilot were invited to attend an introductory CPD training session for Unit 1. They were generally positive about the session, although two felt that it had not been useful. One of these felt that there was too much time spent on the technology, which was sufficiently straightforward that teachers would understand it easily. The other had never used IT in lessons before so found it difficult. This latter teacher worked through the unit again afterwards with a young relative and found that this made it clearer.

Some teachers made specific comments about the activities that involved 'role playing' students: two regarded it as useful to experience the unit from a student's perspective, while two did not. Each of these responses was made by one experienced teacher and one less experienced teacher. In total, four wanted less time spent on the students' perspective during the training.

Several teachers said they would have welcomed more time to plan and discuss how to implement the unit in their own school, i.e. rather than focusing on the content, they would prefer to focus on how to *teach* the content in their context. Several also

felt that the guidance given on the time needed for each lesson needs to be more accurate.

Other ideas for inclusion in the training session included how to address misconceptions and deal with the challenges anticipated at that point, how to adapt their teaching style, and how to use the materials flexibly and creatively. Two teachers requested additional resources, such as examples of students' work or teachers' accounts of it, further graphs for homework, and 'lesson starter' activities. This was mirrored by similar comments from other teachers during the Unit 1 teaching period.

Overall, therefore, the CPD session was deemed a success and, where changes were requested, these tended to be for specific reasons which varied between individuals. This suggests that an element of tailoring future CPD sessions might be useful, if feasible.

Preparation time

Compared with their usual lessons on this topic, most teachers reported spending a similar amount of time, or less time, preparing. There were no apparent patterns in preparation time related to teaching experience or specialism. Those who found that preparation took less time commented that this was because the resources were in place. Some who found it took a similar amount of time added their own resources to the pack or prepared materials for differentiation (e.g. extension activities for more able students). Elements that took time were familiarisation with the unit and setting up the technology, but these could be traded off against time gains in the other areas identified.

Planning and collaboration

All interviewed teachers had a colleague in the school also teaching the unit and joint planning and discussion were common. Only in one school did teachers not discuss the unit at all and this was for practical reasons of limited access to each other. Teachers varied the timing of discussion: some talked before each lesson, some after. Some planned everything together after the CPD session, others did it as they went along, learning from each other's successes and experience. Some taught in the same way and some differently. Approaches varied because of student needs, attitudes and maturity, the teacher's preferred teaching style and practical constraints such as having access to the computer room. Examples of differences included that one class in a school might 'play' with the software more than the other, or one might do more extension activities than the other.

Adaptations

The CPD session had encouraged teachers to adapt the unit to their own context. Only two teachers made no changes to the teaching material for Unit 1. Most commonly, teachers added to or changed the structure of the lesson. For example, they devised their own lesson plans built around SandCircle activities, or produced the activity in a different format. Additions included further graphs for homework,

starter activities, contextual information, devising worksheets to check understanding of the previous lesson, and mini-plenaries throughout a lesson. One teacher did some activities without access to a computer. Teachers also used support mechanisms, such as scaffolding activities, adapting the language used in activities, and reducing the demand for independent working.

Support

Teachers were generally content that they had received sufficient support, both in school and from the Cornerstone team. Support in school included support from senior staff and IT staff, being given priority when requesting access to IT facilities, being able to observe each other teaching, and being given time to attend the training. Support from the Cornerstone team that was appreciated included both face-to-face and email support from LKL, opportunities to discuss early concerns, the CPD session, and the fact that SRI was responsive to queries. The website and online forum received several positive mentions. Teachers valued being able to learn from the experience of other pilot teachers and mathematics experts. One commented that it would be useful to meet with other pilot teachers after the unit: the post-unit presentation session will presumably have met this need. Another teacher commented that the license key for SimCalc had not been provided (the teacher is aware that there were technical difficulties in arranging this).

4.2.2 Delivery

Lesson structure and style

The delivery of the unit tended to follow the general UK model. Teachers usually opened with a starter activity that they had prepared themselves (i.e. not a standard Cornerstone activity), coupled with a recap of what the students had previously learned. The teachers finished the lessons with a plenary, often directly referring to the SandCircle work.

The researchers observed different approaches to delivering the lesson content. Some teachers used the bulk of the middle part of the lesson to move around the classroom, helping individuals or pairs, and occasionally stopping the whole class to explain a common misunderstanding, or to clarify what students should be doing. Other teachers took a more structured approach and set the class specific targets (e.g., finish tasks a, b and c). At regular intervals they stopped the class and took in answers, checking the students' understanding. This teaching style required the teacher to periodically reinforce the pace of the lesson.

Collaborative working and discussion

All teachers encouraged collaborative work, between pairs of students or other peer collaboration or in small groups. Most teachers asked students to use their peers to help their learning: e.g. *"If you're stuck, talk to your partner as well"*.

In all cases, in line with common classroom practice, teachers led any class discussion, although students were able to ask questions if they wished. Examples of

good practice included giving students thinking time to respond, and giving students time to discuss in pairs before responding. Some questions were clearly aimed at individuals, whilst others were intended for the whole class to respond collectively. Teachers re-phrased questions or answers where necessary, and used carefully directed questions to draw out students' understanding, as is shown in the following example:

T: Do you know the speed for Geneva?

S: 10 cm per second.

T: That's one way. What's another way?

S: In the steepness.

T: In the table?

S: The gap between the positions.

T: What does the gap between positions mean?

S: In the first second, it travelled 10 cm. In the second second, it travelled 20 cm. The gap is always 10 metres."

Many teachers adopted a discursive style. They encouraged students to think of multiple ways of responding to a problem. For example, they might ask, *"How else could I work this out?"* They asked students to predict solutions, and then test their prediction on SimCalc. This was a particularly effective method because it engaged students with problem-solving. Teachers invited students to respond to each other: for example, asking student 'b' to comment on whether he/she thought that student 'a' was right - *"do you think that's going to work? How do you know what's right?"* Teachers actively encouraged students to explain their strategy for reaching a particular answer, or to *"expand on that"*.

In terms of language (word selection) and context (game design), there seemed to be few problems. There were issues with reading/understanding the tasks required in one school, although this could be related to the age of the students (Y7). However, in another school with a high number of students with English as an Additional Language (EAL), reading did not cause any issues (Y8). In general, the language of the workbook did not seem to present a problem, although some students in the focus group also mentioned difficulties, for example: use of *"big words"* and *"some questions are not well explained"*, *"the wording is too hard/too formal"*. SRI could consider a review of the language in terms of the amount of reading required and accessibility issues, particularly where the unit might be used with students younger than those for whom it was originally developed. Some students may also require more explanation about certain topics: for example, one Y8 observed class did not understand how bank balances worked.

A few teachers felt pressured to keep to the lesson plan, and this had an impact on the length of time they allowed discussions to continue. It meant that sometimes teachers rounded off discussions in order to keep the pace going, or rushed students to finish off activities: *"Last minute. I want everyone to answer 1A, B and C. I don't want you to think too much about it. Just put in the book what you think is correct."*

One teacher took a more mechanical approach, checking that students' answers were 'right': *"Who got that right? Good"*. There were fewer discussions and the focus of the work seemed to be getting the (right) answers. This teacher encouraged students to think about the problem in a systematic way, showing how the examples always fitted into the same format, $y = mx + c$. Another teacher made sure that each discussion ended with a definite, *"right answer"*, which students were asked to write down in their books.

As well as questioning, teachers checked understanding when they talked with students on their tours of the classroom. Mini whiteboards⁶ were also used to check understanding amongst the whole class, although this was normally a quick visual check to see that the majority understood. One teacher went to the length of checking every single board in the class and making sure that everyone had the right answer, explaining, when necessary, why students were wrong.

Teachers gave one-to-one help on their tours of the classroom. Many students struggled with having to 'explain' their reasoning. Despite this being a key part of the national curriculum, not all students find it a natural skill, and some teachers had identified that this was one of the reasons they wanted to participate in the pilot. Teachers tried to make students feel confident about what they wrote, encouraging them to *"have a go"*, and, *"imagine that you are talking to someone and explaining the answer"*, and to *"write more"*. Some teachers reminded students that they had already covered that topic, encouraging them to work a bit harder to gather their thoughts. For example, when the student asked how to do a particular problem, the teacher responded, *"You know, you've done it before"*. SRI could consider adding a few sample responses for these types of questions to the teachers' manual to give more guidance on what is expected.

Differentiation

One of the issues teachers experienced in delivering SandCircle was managing students who work at different speeds. Overall, it was difficult for teachers to differentiate learning for students with the resources provided, a fact that was recognised during the CPD training session, when teachers were encouraged to adapt the unit as necessary. The strategies teachers used involved either trying to keep the whole class at the same pace and the same stage in the workbook, or allowing high ability students to race ahead in the workbook therefore not following the work set for the lesson, or leaving the high ability students with no further activities to stretch their learning once they had finished the set tasks.

Paired work helped lower ability students because of the support of the partner. They also sought support from other peers. In some cases, it appeared that the lower ability students needed more support from a teacher (as would be the case in any

⁶ Students use mini whiteboards by writing their answer on their individual board and then raising the board so that the teacher can see their answer.

class). However, in general, the lack of differentiation seemed to be more of a problem for the high ability students. SRI could use teachers' ideas as a starting point for developing extension activities for all students, particularly the high ability students who race through the standard activities.

Using the technology

Most of the observed classes took place in a mathematics classroom, although four of the 16 lessons took place in a computer suite. In one lesson, the students worked on paper only and watched the simulations on an interactive whiteboard (IWB).

In the classrooms where laptops were in use (10 out of 16 classes), lessons normally began with students getting these out and setting them up. Occasionally laptops/computers did not work properly, causing minor disruption. The process of logging on and accessing the appropriate files was relatively straightforward. Students worked on their own laptop or computer in seven classes, and in eight classes students worked in pairs. Students in one class did not use SimCalc during their lesson; instead they watched the SimCalc simulations on the IWB. There were no cases reported where the level of computer skill required was beyond the students: all students were familiar with the basic requirements.

An advantage of sharing technology was that more discussion was generated and students helped each other fill in gaps in their understanding. However, researchers observed that when working in pairs a dominant partner tended to control the laptop/computer: one student in the pair moved the mouse and ran the simulations. This is not necessarily a problem but poses a potential risk to learning if the weaker partner does not understand how the software is working. This is a potential issue that can be overcome through using technology that allows dual control or through effective classroom management (e.g. the teacher could encourage each partner to take turns to control the mouse, and employ strategies to ensure that both partners understand how the software works). SRI might consider adding guidance about this to the training material.

In some schools, teachers had prohibited internet access. Even when this was the case, a number of students in each class played around with other computer applications from time to time, instead of focusing on the Cornerstone work. Teachers had a range of strategies for managing laptop/computer use. Often, when they wanted the full attention of the class, teachers asked the students to close or partially shut the laptop lid, or to turn towards the front (and away from their computers).

4.2.3 Effectiveness

Perceptions of the effectiveness of the unit were generally positive amongst both teachers and students, and these were broadly confirmed by the test data.

Observations and teachers' perceptions

All teachers who commented were positive about their students' learning, with almost all saying that it was good and/or better than when using their usual approach. In particular, depth of learning was emphasised as a benefit by several. The following summaries from the researchers' observation notes are all taken from different lessons:

- students seemed confident with some difficult mathematical ideas (Y8);
- students showed good understanding of some tricky ideas (Y7);
- some students showed clear understanding of key ideas and completed all set questions (Y8);
- in the end, most students seemed to have progressed their understanding of the key idea (Y8);
- the majority of students understood the lesson material (Y8);
- key mathematical idea appeared to be understood by all students (Y7);
- levels of understanding were very good (Y8).

However, as might be expected for a unit on such a complex topic, this learning was not without its challenges. One teacher (over halfway through the unit) reported that some concepts had been grasped but that others were still proving difficult and five of the observers also noted such difficulties. Additionally, observers noted that some students had difficulty specifically with:

- new technical language (e.g. velocity);
- aspects of the task (for example, how to develop an equation using the values in the table/graph);
- units (for example, "minutes per metre");
- equations (for example misunderstanding 'x' as 'times').

In pair work, and generally, there were a number of students who were observed copying answers rather than working them out themselves. Some students thought that they could use shortcuts to get to their answers, for example, by finding equations on SimCalc, only finding out much later that this did not produce the correct answer.

The task set by the mobile games company in the context of the unit lends itself to referring explicitly in the students' workbook to the level of challenge they will face during the task. That might be useful in order to minimise any risk of disengagement where students find the unit challenging.

Despite the challenges, half of the teachers commented that the software is good for helping students make connections and one felt that the Y8 pilot students had gained better understanding than a Y10 GCSE class taught conventionally. The role of engagement in supporting learning was mentioned by four: some in terms of motivating students to engage with the concepts and others in terms of making the concepts real (further information on engagement is given in section 4.5). The ability

to 'play' and manipulate the software was also seen as useful in supporting learning and one teacher commented that the technology allowed students who would not normally 'shine' to do so.

The final question in the interview asked teachers to summarise their views, on a scale of 1 to 10,⁷ about how good the unit was as a way of teaching linear functions. Eleven teachers gave ratings at or above 8. Five gave lower ratings (a 5, three 7s and a 7.5), either because they were reserving judgement until they had finished teaching the unit, or because they had views about improvements needed. These improvements related to the time allowed for the unit and to the need for differentiation, and they are discussed further in section 4.6 below.

Students' perceptions

Students tended to agree with the teachers that the unit was helping their learning (all were part way through the unit). They believe that SandCircle has helped them develop a range of skills, in particular the 'top five' skills of (in order of percentage size of those who ticked, 'A lot' or 'Some'):

- reading and interpreting graphs;
- drawing graphs;
- reading and interpreting tables;
- completing tables; and
- writing equations.

They think that SandCircle made a smaller, but still positive contribution to their skills in 'making predictions', and 'explaining' (see Table 2).

Table 2: How much do you think SandCircle has helped you develop the following skills?

| Skills | Percentage ⁸ of student response | | | | |
|---------------------------------|---|------|----------|------------|----------|
| | A lot | Some | A little | Not at all | Not sure |
| reading and interpreting graphs | 37.4 | 39.7 | 18.1 | 3.8 | 1.0 |
| drawing graphs | 31.7 | 38.1 | 19.0 | 7.9 | 3.0 |
| reading and interpreting tables | 29.3 | 39.4 | 22.6 | 5.6 | 2.8 |
| completing tables | 32.1 | 36.2 | 20.0 | 8.5 | 3.1 |
| writing equations | 44.0 | 22.4 | 17.6 | 13.2 | 2.5 |
| making predictions | 24.4 | 32.9 | 27.2 | 12.3 | 3.1 |
| explaining | 25.2 | 30.3 | 25.4 | 12.2 | 6.6 |

⁷ where 1 is poor and 10 excellent

⁸ Percentages are calculated from the total number of students who answered each item (and excluding those who omitted the item).

Even so, despite the positive views expressed overall by students and teachers, a significant number of students in the focus groups do not think that they are learning as much in SandCircle as they do in their regular mathematics lessons.

The students with positive views, around three fifths of those interviewed, thought that SandCircle improves their ability to learn independently. They are able to *“use their brains, not calculators”*, and SandCircle *“makes us think”*. They liked the fact that they were given different ways of learning the same thing, for example, by looking at similar graphs in different contexts, and they liked how the sequence of table, graph and equation gave them different ways of working out the same thing (*“we saw how the graph fitted with the equation”*). The questionnaire results showed that 68 per cent of students agreed or strongly agreed that SandCircle helps them understand different ways of solving a problem, and a similar percentage (63 per cent) considered that SandCircle broadens the range of problem-solving techniques available to them. A few students liked the repetition because *“it gets the ideas stuck in your head”*. Students also pointed out that SandCircle is great for visual learners, and many spoke with enthusiasm about the fact that there is less writing and reading required. A summary comment from one student was that *“you learn without realising it”*. A couple of students thought that they learned more through SandCircle than they would in their regular mathematics lessons: *“we learn more than we do in normal maths”* (Y8 student). A Y7 student said, *“SandCircle is pushing you further in your maths”*. When students reflect, many believe that SandCircle is improving their learning.

In contrast, the students holding negative views, around two fifths of those interviewed, had the feeling that *“you leave the lesson not knowing anything [new]”*. The major problem for students in this group was the perceived repetition in SandCircle: they wished there was more variety and thought that they spent too long on the same topic. For these students, the sequence of table, graph and equation made SandCircle boring. A large proportion felt that what they were learning was too easy – *“all the games are based on the same principle”*, and *“we keep going over the same stuff”*. Some of the Y7 students thought that *“older pupils would get bored”*, and it was clear from the focus groups that older students were indeed more critical of the unit. In a similar vein, some students thought that *“SandCircle will be good for bottom sets because they will understand more”*. A small number of students (from a Y8 mixed ability group, and a Y7 middle and top set group) thought that SandCircle was too hard and that they needed more support from the teacher. The students in the focus group were of mixed ability and these negative views did not appear to be linked to any particular ability level.

Test outcomes

Despite these mixed views, the test data showed that students had generally progressed in their understanding. Analysis was carried out at overall sample level and by school and class. Analysis was also conducted for each individual test item. An outline of the content of the test items is given in Appendix C.

- Sample level

- The pre-test mean for the whole sample was 10.72; 6.35 for component M1 and 4.37 for component M2.
- The pre-test reliability alpha for the whole sample was 0.84, indicating acceptable test reliability; reliability was 0.68 for component M1 and 0.75 for component M2.
- The post-test mean for the whole sample was 18.52; 8.76 for component M1 and 9.76 for component M2. These figures showed an overall learning gain on all measures.
- The post-test reliability alpha for the whole sample was 0.90; 0.81 for component M1 and 0.89 for component M2, again indicating appropriate reliability.
- School level
 - On average, students in eight of the nine schools scored significantly higher in the post-test compared with the pre-test. All but one of these results was significant to $p < .0001$ (i.e. the probability that these results occurred by chance is less than one in ten thousand).
 - The remaining school appeared to show a small increase in mean score, but this was not statistically significant (see the class level results below for more information).
 - For one of the eight schools showing a significant improvement at the school level, the improvement was only significant overall and not when the results were broken down by the individual components of M1 and M2 (this is probably because of the sample size).
- Class level
 - On average, all 17 classes scored significantly higher in the post-test. Note that, despite this, when the results were collapsed at the school level, one school did not show a significant difference between pre- and post- test scores. In this school overall, it would appear that the two classes which took part in the testing were less similar to each other than classes tested at other schools, so that the difference was only visible at class level not at overall school level.
 - The significant improvement overall at class level was not significant for some classes when the results were broken down by M1 and M2.
 - The M1 results reveal that all classes show an increase in mean in the post-test, but the results for three classes are not significant. Among the 14 classes that showed a significant improvement, nine of the differences were highly significant ($p < .0001$).
 - The M2 results were even more positive: all 17 classes showed significant improvements, and 14 out of 17 were highly significant ($p < .0001$).
- Item level

Descriptions of the content of each of the test items discussed here can be found in Appendix C.

- For M1 items, 17 out of 18 items showed a greater percentage of students achieving correct scores in the post-test. There was one item (item 2a), where slightly more students achieved correct scores in the pre-test. This is a multiple choice item requiring translation of a linear function from narrative to

symbolic form (students are asked to pick the correct equation representing the number of months it will take for a company to have 900 workers). This might indicate a problem with students interpreting and understanding equations or there may have been some confusion with the unit 'm' (which stands for 'months' in the test, but was also used to describe 'metres' in the unit).

- The top five M1 items, showing the highest increase between the pre- and post-test, were the following (in order of size of difference): 11a; 8b; 10b; 5; 13 (see Table 3).
- For M2 items, all 18 items showed a greater percentage of students achieving correct scores in the post-test. The top five items, showing the highest increase between the pre- and post-test were the following items (in order of size of difference): 16c; 15c; 15d; 15b; 17a (see Table 4).
- Omission rates
 - These were higher for the pre-test than for the post-test, as would be expected, indicating perhaps that students taking the post-test were either more able to attempt questions, or more confident to try than they had been during the pre-test.
 - Nineteen students (four per cent) omitted the first question in the pre-test but only four students omitted this in the post-test (less than one per cent). Similarly, 295 students omitted the last question on the pre-test (69 per cent) compared with 140 students in the post-test (33 per cent).
 - Other items in the pre-test with particularly high omission rates of more than 100 students, other than the questions at the end of the test (items 14a to 17b), were the following:
 - item 10b – 135 students omitted;
 - item 11c – 138 students omitted; and
 - item 11d – 234 students omitted.
 - In the post-test, the number of items with particularly high omission rates was lower. Apart from the last two questions (items 16d to 17b), the only item which more than 100 students omitted was item 11d (122 students omitted). This item is at the end of a 4-part cumulative question (i.e. each item in the series builds from the one before). There are other multi-part items in the test, but the others contain items that are free-standing (i.e. the answer to one is not dependent on the answer to the preceding ones). This may be one explanation for the high omission rate. Another possible reason may be that it is the second and final time during the test that students are required to write an equation: students who find writing equations challenging may have attempted to write an answer for the earlier question, but may have given up on the second attempt.
 - In the pre-test, 235 students did not reach the final item in the test compared with 102 students in the post-test. The results show that more students got further in the post-test before dropping out than in the pre-test, as expected.

Table 3: M1 items showing the largest pre- to post- test differences (with item 2a shown for reference)

| Item | Pre-test correct (%) ⁹ | Post-test correct (%) | Difference between pre- and post- test correct (percentage points) |
|------|-----------------------------------|-----------------------|--|
| 11a | 32.51 | 55.74 | 23.23 |
| 8b | 54.11 | 74.94 | 20.83 |
| 10b | 9.18 | 28.28 | 19.10 |
| 5 | 3.33 | 21.76 | 18.43 |
| 13 | 41.95 | 59.90 | 17.95 |
| 2a | 59.30 | 52.99 | -6.31 |

Table 4: M2 items showing the largest pre- to post- test differences

| Item | Pre-test correct (%) | Post-test correct (%) | Difference between pre- and post- test correct (percentage points) |
|------|----------------------|-----------------------|--|
| 16c | 14.90 | 62.43 | 47.53 |
| 15c | 26.69 | 72.09 | 45.40 |
| 15d | 20.68 | 64.13 | 43.45 |
| 15b | 34.36 | 75.07 | 40.71 |
| 17a | 3.76 | 37.30 | 33.54 |

4.3 Effectiveness for Different Groups

This section addresses the sixth research question:

RQ6. What are the perceived impacts on learning using these units in terms of different groups of students (e.g. ability, gender, ethnicity or SES)?

Perceptions

There were mixed perceptions about the impact of the unit on different groups of students. Teachers tended to differ in the groups they perceived to be gaining most from the unit, while the test data showed very few significant differences in terms of pre- to post-test gains between groups. This difference may have been affected by the small sample size. However, the test outcomes are useful as indicative findings and SRI will no doubt intend to explore such differences (or absence thereof) on a larger scale in future trials.

For four teachers, their perception was that the impact of the unit on learning applied equally for all students. Others suggested that outcomes were better for:

⁹ Percentages are calculated from the total number of students who answered each item (and excluding those who omitted the item).

- those with prior experience of the topic (because they can make links with what they already know);
- the less able (because they are supported by paired work and that helps those who are not confident);
- higher ability students because they learn more quickly (although this teacher acknowledged that they may not necessarily learn *more* than the less able); the unit could be made shorter and harder for higher ability students;
- boys (because of the technology);
- girls (because, in one school, they are more able than the boys; in another, they are seen as just doing better with the unit);
- those who like working on computers (some prefer pen and paper).

These views are not necessarily contradictory. Rather, they might represent individual rather than group effects, relating to specific features of particular contexts. It is also important to remember that they are perceptions and that some perceived effects might be conflated (as in the example above where gender and ability might be conflated in one class). Analysis of the test results can help to clarify the extent to which perceptions might relate to group effects.

Test outcomes

Subgroup analysis of the test results was carried out where feasible, according to key variables. Due to the small sample size, the results of the subgroup analysis should be treated with some caution but may suggest areas where further investigation might be required. Hierarchical linear modelling (HLM) was used to examine subgroup differences among the following subgroups:

- student gender;
- student level (3-7);
- student year;
- school percentage of A*-C mathematics¹⁰;
- school level of progress¹¹;
- school percentage of students eligible for FSM (free school meals).

'Gain scores' were generated for each group as an outcome measure (pre-test score subtracted from post-test scores). The difference in the pre-test results for each group was also analysed.

In addition, school type (boys/girls/mixed) and category (state/academy/private school) were analysed. Because of the particularly small sample size for some sub-

¹⁰ The percentage of students in each school who achieve the highest grades (A* to C) in their end-of-statutory-schooling mathematics exams (GCSEs) at age 15-16.

¹¹ A measure of how many students in a school achieve the required number of levels of progress, as defined in government targets, across the key stage.

samples, however, data for these groups must be treated with caution and so have not been included or summarised in the report.

The headline results from these analyses are as follows:

- Gender¹²:
 - Overall, girls scored significantly lower than boys in their pre-test scores: their average score on the pre-test was 8.77 compared with 10.04 for boys.
 - However, girls' gains were significantly higher than those of boys. They made a gain of 7.43 points on average, compared with a 6.93 point gain for boys. Although small, this difference is significant at the 0.05 level. This might be a spurious effect or it might indicate that the Cornerstone materials may be helpful in narrowing the gender gap in this topic. It would be useful to investigate this further with a larger sample.
 - In M2, girls achieved a significantly lower score than boys in their pre-test scores, but the difference in the girls' average gain score compared with the boy's average gain score was not significant.
 - In M1, no significant results were found when comparing results for girls and boys (pre-test score and gain score).
- Student level¹³:
 - Overall, students who were working at a higher level were significantly more likely to have a higher pre-test score, as expected. However the difference in gain score between students of different levels was not significant.
 - In M1, level 5 students achieved significantly higher gains than students at other levels. A ceiling effect for higher level students may have contributed to lower gains for the level 6 and 7 students.
 - In M2, students who were at a higher level were significantly more likely to achieve a higher pre-test score and gain score. A ceiling effect may also exist for M2 questions; however, because there is a greater distance for students to go in order to achieve top results, a ceiling effect is less likely to be a restricting factor for most students attempting M2 questions.
- Student year:
 - Overall, Y8 students scored significantly higher than Y7 students on the pre-test scores.
 - The Y8 and Y9 scores were also compared, but, because of the especially small sample size in Y9 (23 students in one Y9 class, compared with 227 students in the Y8 classes), the results are not reported.
 - No significant results were reported for gain scores across years 7, 8 or 9, indicating that students in all years made comparable gains in understanding.

¹² Analyses for gender only used data from mixed-gender schools, excluding three single sex schools. Including these schools would have risked confounding pupil-level and school-level factors.

¹³ Students' attainment is assessed by their teachers according to national curriculum 'levels'. The levels represented in the analysis ranged from level 3 to 7. The expected level at the end of primary school in England (Y6, age 10-11) is level 4 and at the end of key stage 3 (age 13-14, level 5 or 6).

- Percentage A*-C mathematics¹⁴:
 - Overall, students who attended schools with a higher percentage of A*-C mathematics attainment were significantly more likely to achieve a higher pre-test score, as might be expected.
 - There were no significant results reported for gain scores in relation to schools' percentage of A*-C mathematics attainment.
- Percentage levels of progress¹⁵:
 - Significant results were only found in the pre-test scores for M2 questions: the higher the percentage of students in a school achieving at least three levels of progress, the higher the pre-test score for M2. M1 and overall results were not significant.
- Percentage of students eligible for FSM (free school meals):
 - Significant results were only found in the pre-test scores for M2 questions: the higher the percentage of students qualifying for FSM in a school, the lower the pre-test score. M1 and overall results were not significant.

4.4 Curriculum Fit

This section addresses the fourth research question:

RQ4. Do teachers believe that the materials help them meet the requirements of the national curriculum and their own mathematics learning goals, and prepare students for their assessments?

4.4.1 Curriculum

Most teachers who commented felt that the unit fits well with the demands of the English national curriculum. They also tended to feel that the unit covers more than their own scheme of work (although one teacher of a Y8 class felt that it does not go far enough) and that it gets students started on abstract concepts earlier and in a more supported way than their own scheme would. Teachers liked the connections the unit makes between different elements of the topic, although one noted that it takes longer than their own scheme (a view echoed by other teachers elsewhere in the interviews) and another suggested breaking the unit into smaller parts.

4.4.2 Assessment

Some teachers found it difficult to predict how well the unit had prepared students for their next assessment of their knowledge of linear functions. Four said they could not predict: one was waiting to see the results on the school's usual assessment, while the other three did not know or were not sure. Three teachers were not asked because of pressure of time on the interview and the remaining nine teachers all thought that their students would do better or well. These teachers commented on

¹⁴ This variable is the percentage of students in each school who achieve the highest grades (A* to C) in their end-of-statutory-schooling mathematics exams (GCSEs) at age 15-16.

¹⁵ This is a measure of how many students in a school achieve the required number of levels of progress, as defined in government targets, across the key stage.

their students' solid understanding, and expressed expectations of longer-term recall, less need for revision, and greater willingness to attempt questions. Again, the pre-test and post-test data seem to bear out the latter view, since omission rates on the post-test were consistently lower than those for the pre-test.

4.5 Engagement and Motivation

This section addresses the fifth research question:

RQ5. Do the materials increase student engagement/motivation compared with their general attitude to learning mathematics?

4.5.1 Engaging students in SandCircle

Motivating students was one of the factors that prompted teachers to participate in the pilot and efforts to engage students were evident in many of the lessons observed. Some of the views expressed in this section may have been expressed by a single teacher only. However, that does not necessarily invalidate the response. It is important to bear in mind that, whilst it is not possible to draw conclusions on the basis of a single response, that single response might raise a question that is worth considering further or investigating more broadly in the future.

As might be expected, researchers observed a range of classroom environments, and these were influenced by the culture of the school and the pedagogical style of the teacher. Some classes gave the impression of being more focussed than others because the class settled down quickly and the teacher did not have to work hard to get the students to listen. Other classes observed were lively, and maintained a level of chatter throughout the lesson. In many of the observed classes, students' focus began to flag towards the end of the lesson, particularly in one case when the observed lesson was the second lesson of a 'double period' back-to-back mathematics lesson.

Teachers generally approached the materials in a positive and 'business-like' manner. Many made a special effort in the lesson introduction to engage students, for example, by using short starter activities they had created themselves. Some attempted to link the mathematics principles to everyday life outside Cornerstone (for example, one teacher used the example of driving a car to link to speed). In this extract the teacher reminds the student to link the bank account graphs to SandCircle's business:

S [reading their account of the bank balance graph]: "we started from zero and increased by 10,000"

T: "OK, very good, you're describing the graph but we need to know in the world what's making the changes happen? What's causing it?"

S: "We made a new game and a lot of people bought it, but then we had to buy new equipment for the games..." etc

Teachers encouraged students to participate in a number of ways. They asked questions, sometimes asking students with hands up, and sometimes asking those without hands up. Mini whiteboards were also used as a way of trying to engage all students and were a good way of checking students' understanding. Some teachers asked students to draw graphs on their whiteboard, or to draw a graph on the IWB. Students across the sample really focussed when their classmates were asked to demonstrate their answers. Another effective method of engaging all students was to get students to read out passages from the workbook. Watching simulations on the IWB was also very engaging for students.

In the observed classes, many students seemed very enthusiastic and engaged. There was lively discussion between partners and a number of examples of pairs explaining to other pairs. However, some observed students were less engaged with the lesson, and were distracted by playing around on the laptops or doodling on their mini whiteboards. Some students did not seem to find the topic interesting and observers reported that they appeared 'bored' throughout the lesson. Some students did not seem to understand what they were meant to be doing. There was often a lot of demand for the teacher's attention and some students did not get to ask their questions until late in the lesson.

4.5.2 Teachers' views of SandCircle in relation to usual mathematics lessons

Of course, it is not uncommon for mathematics classes to contain both students who are highly engaged and those who are less so. For that reason, teachers were asked to report on how students' engagement with the SandCircle unit compared with their usual level of engagement with similar mathematics topics, whilst students were asked how much they enjoyed their usual mathematics lessons and their SandCircle lessons.

Most teachers (nine) felt that SandCircle was more engaging for students but several believed that the novelty of using the software was a potential factor in that engagement. In addition, another noted that the novelty of working with the software means that students persevere for longer. However, two teachers noted that motivation was already waning: that the novelty was wearing off and students had decided that the unit was repetitive and boring. This is discussed further in section 4.5.4 below.

Other factors that teachers consider motivated students were the elements of creativity, collaboration, discussion and visual hands-on approach. Some negative views were reported, however. Several teachers mentioned students' disappointment at not being able to create their own games, while others noted that: girls seemed more resistant to the change to using technology; reading demands made it difficult for some; and students are either motivated or not so that the approach is immaterial. One teacher thought that motivation was not affected by the unit; that the unit simply promoted a different type of motivation because of the software and the games

context. Whilst these views might have been isolated (i.e. not necessarily shared by other teachers), this does not necessarily prevent them having validity. Individual views might be idiosyncratic or might be worthy of further investigation with a larger sample.

4.5.3 Students' views of SandCircle in relation to usual mathematics lessons

Students' views provide another perspective. Seventy per cent of students reported enjoying learning mathematics at school, in general, and nearly all students (95 per cent) think that they need to get good grades in mathematics, so it is clearly a subject in which many students have a motivation to do well and that many enjoy. These percentages are higher than those found for similar questions in the TIMSS 2007 survey of eighth grade (Y9) students¹⁶. This may reflect an age difference or might be related to the relative enthusiasm of the pilot group of teachers creating a more engaging environment for the students.

Students say that they like to be challenged, they like problem-solving, they like to learn new things and they like to learn things in different ways, all of which they can do in mathematics. Criticisms about the subject mentioned by a few students include that they do not like to “go over the same stuff”, and they do not like copying from the board, or when the mathematics gets too hard.

When asked to rate their usual mathematics lessons out of 10¹⁷, the most frequent score that pilot students gave was 8, with three quarters of students rating their lessons between 6 and 10. Researchers also asked students to rate their SandCircle lessons out of 10. They gave a spread of responses, indicating mixed opinion about how much they enjoy SandCircle lessons compared with their regular mathematics lessons (see Table 5). A larger proportion of students gave ratings of 5 or below for their SandCircle lessons.

The average score given for enjoyment of regular mathematics was just under seven; the average score for SandCircle mathematics was just under six. Although this is an apparently small difference (of 0.932), it is highly significant. A t-test comparing the results for these two questions showed a difference at the 0.001 level.

¹⁶ Mullis, I.V.S., Martin, M.O. and Foy, P. with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., and Galia, J. (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: Boston College, TIMSS & PIRLS International Study Center.

¹⁷ where 1 = I don't like them and 10 = I like them a lot

Table 5: On a scale of 1 to 10, how much do you usually like your mathematics lessons/did you like the SandCircle lessons?

| | Frequency of rating given by students 1 = I don't/didn't like them; 10 = I like/liked them a lot | | | | | | | | | | Average rating |
|---|---|----|----|----|----|----|----|----|----|----|----------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Usual mathematics N=388 | 7 | 7 | 14 | 23 | 44 | 54 | 75 | 89 | 48 | 27 | 6.8 |
| SandCircle mathematics N=390 | 26 | 24 | 19 | 36 | 57 | 52 | 60 | 57 | 46 | 13 | 5.9 |

| | Frequency of rating given by students 1 = I don't/didn't like them; 10 = I like/liked them a lot | | | | | | | | | | |
|--|---|---|---|---|---|-----|---|---|---|----|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| % rating usual mathematics 1-5 | 24% | | | | | 76% | | | | | % rating usual mathematics 6-10 |
| % rating SandCircle mathematics 1-5 | 42% | | | | | 58% | | | | | % rating SandCircle mathematics 6-10 |

The questionnaire also reveals that 42 per cent of students held a relatively negative perception of SandCircle lessons (those students who rated enjoyment between 1 and 5 inclusive), compared with 24 per cent of students rating their regular mathematics lessons between 1 and 5. This was backed up by comments made during the focus group sessions and in the open response question on the questionnaire. A frequent point made in the focus group was that students would have given higher scores for their enjoyment of SandCircle at the start of the unit, but too much perceived repetition made SandCircle less appealing over time. This is despite the fact that the vast majority of students in the focus groups gave scores of eight, nine or ten when asked about their overall enjoyment of SandCircle mathematics lessons. Their scores might have been higher than those of the questionnaire responses because the students who were selected¹⁸ for the focus group were particularly positive students. Alternatively, it could be that students thought more positively about SandCircle after they had discussed different aspects of the unit amongst themselves in the interview, or that they felt less able to express their true feelings in a group environment. It is also possible that timing was an issue: the focus groups were held during the teaching period, whereas the questionnaires were completed nearer to the end of the unit.

¹⁸ Teachers selected students for the focus group. They were advised to provide a mix of students (e.g. of different ability levels or different language backgrounds).

Whatever the reason, there is evidence that a sizeable number of students did not rate their SandCircle lessons as highly as their usual mathematics lessons. There might be an element of ‘disliking change’ about this finding, but the scale of the effect suggests that other factors might be at play. Of course, this does not negate the fact that students *learned* from the unit. Nevertheless, it merits further investigation, both in Unit 2 and during a wider trial, to establish whether this is widespread or an effect particular to this group of students and, if the former, what guidance can be given to teachers and, through them, to students to minimise the effect.

4.5.4 What students think helps them engage

Students’ specific feedback about their engagement in the unit included comments on the software. The introduction of intensive laptop/computer work means that students work in pairs and small groups, and for several classes this is a different arrangement compared with their regular mathematics classes, when students often work on their own. This has an influential effect both on students’ perception of learning and their engagement. For the majority of students, paired work is fun, and the extra discussion that arises from sharing a laptop/computer *“improves understanding”*. As one student put it, *“explaining to someone else helps learning”*. More group discussion often means less teacher-talk, and several students pointed out that this is also beneficial for them. However, a small number of students do not enjoy paired work, particularly when paired with someone of a different ability level. It is therefore difficult at this stage to disentangle the benefits to learning of changing working arrangements from the benefits of the SandCircle unit.

Seventy-six per cent of students responding to the questionnaire said that they enjoy using computers to help them solve mathematics problems; however, the majority of students only rarely use laptops or computers in their mathematics lessons. For this reason, the use of laptops/computers in the lessons clearly has a novelty value for nearly all students and this should be taken into account when considering students’ reactions to the software. This also mirrors comments made by some teachers in section 4.5.2. Whilst perceived novelty appears to be supporting the unit’s positive effect on learning for most students, it will be important to consider the possible impact of that effect wearing off. SRI is developing some units that do not use the SimCalc software and it will be important to evaluate the extent to which these impact on learning, to confirm that such units have similar impact on learning for students in England as the units which do use the software. Equally, it will be important to compare outcomes for the SandCircle unit when taught before other Cornerstone units, with outcomes for the SandCircle unit when taught after other Cornerstone units. This would confirm that learning is solid and not affected by the ‘novelty factor’ of the approach.

Many students in the focus groups expressed the view that the SandCircle software improves their lessons, making them more varied (*“it’s better than just writing or doing worksheets”*), making the students more involved (more *“interactive”*), and making the topic easier to understand (*“you can check the results straightaway”*).

These findings are in keeping with the tendency for focus group students to give more positive ratings for SandCircle compared with the larger group of their peers. A few students in the focus groups criticised the computer graphics, saying that they are not as sophisticated as real game graphics. A couple of students indicated that they prefer to work on worksheets, and others mentioned occasional problems with getting the laptops/computers to work.

Context was also relevant. The majority of students enjoyed the topic of mobile phone game design, finding it relevant to their everyday life (*"it is not just maths – it is about real things"*). Students enjoyed the characters and found the storylines *"motivating"* (especially more developed storylines such as Red Riding Hood or Wendella the dog), although some thought that the context was fabricated – *"if it was a real game no-one would play it as it would be so boring"*. Other criticisms, voiced by a small number of students, included that the characters are *"babyish"* (from year 7, 8 and 9 students) and that there is a gender-bias towards boys (the unit includes robots and cars as characters). Some of the girls interviewed said that they do not often play games on their phones so the topic was not particularly relevant to them. Many students were also disappointed that the unit did not actually involve designing or developing their own game.

Some students found the workbook easy to understand, and liked the fact that they had all the information they needed in one place (and that they did not have to copy questions from the board). However, other students criticised the style (*"dry/boring"*, *"needs more pictures"*); and the lack of differentiation activities for higher ability students (*"tedious for fast workers"*). A summary statement from one student was that, *"the workbooks are not 'wow' – they don't have a big impact"*.

4.5.5 Differential impact on engagement

Teachers expressed a range of views regarding whether the unit impacted on the engagement of any group of students more than others and there were no obvious patterns according to the year group taught or the extent of the teacher's experience. Findings were as follows.

- Five teachers felt that there were no differences in the engagement of different groups of students: three because all students were engaged, and two because their students responded with the usual level (or lack) of motivation.
- Three teachers perceived gender differences. Two considered that girls were engaging more (although in one case, the teacher felt that this might have been because the boys in the class were generally weaker at mathematics). One considered that boys were engaging more, because they engage with the technology.
- Six teachers reported ability differences. Four felt that weaker students were engaging more, in some cases because increased discussion and less writing helps them to show their understanding, and because repetition helps to consolidate their learning. In contrast, two felt that more able students engaged better because they persist and need less teacher intervention.

As with the previous section regarding differential learning, these mixed views might represent individual rather than group effects, relating to specific features of particular contexts. Equally, they might represent conflated effects. Further evidence is required.

4.6 Elements of the Unit

This section addresses the third research question:

RQ3. Which parts/aspects of the unit were found to be useful? Which were not so useful?

4.6.1 Best aspects of SandCircle

During the interviews, teachers were asked to identify the best three aspects of the SandCircle unit. Some named only two aspects, but many comments were broad, relating to the software or resources, for example. Elements identified among the 'best' were as follows.

- The most common first answer was the software (seven teachers) and this was also mentioned elsewhere in the 'top three' by another four teachers.
- The materials and resources were listed first by five teachers and received another 10 mentions overall.
- Two specific activities were mentioned (the Red Riding Hood and Wendella activities, although reasons were not given; students also identified that they liked these activities – see section 4.5.4).
- Teachers also praised the general approach of the unit, sometimes in relation to the materials (e.g. the fact that the activities are exploratory, support students in making connections, and require higher order thinking skills) and sometimes in general (e.g. the collaborative approach to working).
- Three teachers noted that the target level of the unit was good and appropriately scaffolded and another that outcomes for students were positive.
- One teacher felt that the unit was good in terms of forcing the use of technology in teaching and therefore taking the teacher out of their 'comfort zone'.

4.6.2 Specific changes teachers would like made to the unit

When asked what they would want changed if teaching the unit again (up to three aspects), many teachers mentioned only one or two things, although one mentioned four. The materials were mentioned most often (13 mentions, seven first and six second comments). Comments about materials are summarised below. Some are observations and may not require action; others may have potential associated action. Most are accompanied by a constructive suggestion of something that could be amended to improve the unit beyond the current, generally positively-rated, standard. In most cases, each specific comment was made by only one or a small number of teachers. Nevertheless, all are informative and worthy of consideration.

- Provide access to an electronic version of the booklet so the teacher can tailor the selection of activities and print pages as worksheets and extension activities.
- The activities feel repetitive to the students. Combine similar activities to limit this and change the homework accordingly; change some contexts and/or scenes;

perhaps add activities that do not use the software (example given: working outside, drawing axes on the ground and modelling in different ways).

- Include contexts/graphs with negative coefficients of x .
- The equation $x=0$, was found to be confusing; the teacher acknowledged that the best GCSE/A level students can struggle with that so it is hard for these younger groups.
- Some sections of the booklet were perceived to be 'wordy' and, thus, confusing for students, Red Riding Hood in particular (although the students became used to it).
- Include more material in the teachers' booklet and replicate the student booklet information in the teachers' booklet (for example, the teachers' slides do not have the graphs that are in the students' books).
- Additional homework activities would be useful.
- Provide blank graphs to help teachers prepare homework.
- There is too big a leap in some activities (e.g. it becomes abstract when looking at direct proportion). Some activities jump too quickly from one concept to another without much explanation.
- Include a 'key vocabulary' list at the beginning of the unit so students have knowledge of it in advance.

Suggestions for changes to lessons were also frequent (10 mentions). These comments related mainly to timing but also referred to differentiation, prediction skills, and the importance of knowing the students' ability. Ideas for further refining lessons were as detailed below.

- Four teachers commented on time: the amount of material overall is fine but too intensive; spread the activities over a longer time; allow more time to introduce a concept; allow time for group discussion so that learning is remembered; break down the middle activities (e.g. cover intercept in one context and intercept/change in another).
- Provide differentiation activities (consolidation and extension) and more starter activities.
- Put more emphasis on prediction: students were too keen to press 'play' and did not do enough prediction; enable them to record predictions, not necessarily in writing (students were not always willing to write predictions in their booklet in case they were 'wrong'); consider separating out the files so that the 'predict' file is separate from the 'check and explain' file.

One teacher commented that the pre-test was too hard and that this was upsetting for some students. This teacher suggested calling it something else, not a test: the alternative title suggested was 'What do you already know?'.

Another teacher commented that it is important that teachers know the ability of the students before starting the unit. This teacher had had limited contact with these students beforehand so could not anticipate how they would cope with the unit. Given the inevitable level of challenge of the unit, it might be useful to add guidance about this to the teacher materials.

Other themes raised related to software issues.

- The software was less 'glitzy' than one teacher expected (although the students became used to the format). Another referred to it as basic and user friendly, while a third called it 'archaic' but noted that it met the purpose. It was apparently seen by these three teachers as fit for purpose, despite being less sophisticated in appearance than they had expected.
- One asked if files could be made available to students outside of school.
- One teacher wanted to be able to open multiple files simultaneously.

4.6.3 Students' views of aspects of the unit

In the questionnaire, students were asked whether various aspects of the unit were helpful or not to their learning. Their responses are summarised in Table 7, where the units are listed in the order in which they are given in the students' booklet. The various elements were generally seen as helpful.

Nearly a quarter to a half of students reported that they had not yet done some of the topics towards the end of the unit. Bearing this in mind, of the students who had been taught each topic, the majority in most cases thought that the topic was helpful. The exception was 'Problems from the SandCircle mobile lunchroom', which was rated as 'helpful' and 'not helpful' by similar proportions. Note that the percentage of students who ticked, 'Haven't done this yet' did not steadily increase across the unit. This may be because teachers did not always follow the order of the teaching materials as provided, or it may be that students did not recall some activities.

The top four results from this question, where the proportion of students finding the topic helpful was much bigger than the proportion of students who did not find the topic helpful, were (in order of size of difference):

- Wendella's journey (setting up the journey paths for Wendella the dog);
- Controlling characters with equations (including the robots Shakey and Roberta);
- Money matters (solving business problems for SandCircle);
- Better Games (making mathematical controls for Reynaldo, Bommakanti, and Geneva for *Better Games*).

The proportion finding an individual topic area 'not helpful' varied from almost one fifth (linear relationships) to more than one third (Yari).

Table 7: How helpful were the different parts of SandCircle to your learning?

| Topic area | Percentage ¹⁹ of student response | | |
|---|--|-------------|-----------------------|
| | Helpful | Not helpful | Haven't done this yet |
| Yari, the yellow school bus | 57.5 | 37.0 | 5.4 |
| Texas Road Rally | 67.2 | 30.2 | 2.6 |
| Controlling characters with equations (including the robots Shakey and Roberta) | 74.7 | 24.3 | 1.0 |
| Better Games (making mathematical controls for Reynaldo, Bommakanti, and Geneva for <i>Better Games</i>) | 71.4 | 26.5 | 2.1 |
| Wendella's journey (setting up the journey paths for Wendella the dog) | 74.6 | 17.0 | 8.2 |
| Money matters (solving business problems for SandCircle) | 66.1 | 18.6 | 15.0 |
| Crab velocity | 51.6 | 24.4 | 24.1 |
| Wolf and Red Riding Hood (exploring characters with different velocities) | 56.3 | 17.7 | 25.5 |
| From average to average rate (calculating average game scores, pay rate, and travelling rate of characters) | 52.3 | 20.5 | 26.9 |
| Problems from SandCircle mobile lunchroom | 28.3 | 28.0 | 43.4 |
| Linear Relationships: Proportional and non-proportional | 44.7 | 16.8 | 38.4 |
| Going full time (an audit of skills you have learned over the unit) | 31.8 | 20.6 | 47.6 |

4.6.4 The software

In addition to comments made earlier about aspects they liked about the unit or would like to see changed, teachers were asked questions specifically about the software. Benefits of the technology identified by teachers included:

- that students were able to model ideas for the rest of the class and display them on the interactive whiteboard;
- that this was the first time that students had used laptops in class (not in the computer suite); and

¹⁹ Percentages are calculated from the total number of students who answered each item (and excluding those who omitted it).

- that this was the first time that one of the experienced teachers had taught with laptops.

Comments related to technical issues included problems with perceived ‘clunky’ access and difficulties in opening and manipulating files (e.g. opening them up, moving windows, moving the graph lines up and down; a teacher suggested modifying the technology so that students simply have to move the end points of the graphs to draw the lines). Some of the teachers’ issues were caused by lack of familiarity, some by technical issues. The following specific factors were mentioned.

- The function box for ‘Crab Velocity’ was found to be confusing.
- Wendella’s journey would be better if she were moving through the different places in the mini-world not just through a green background.
- Labels floating on the graph lines would be useful, especially where there are many lines.
- Use resolutions that can fit on any screens (the smaller the screen, the less you can see).
- The license agreement is hidden behind what is on the screen when it is set up, so there can be initial problems in accessing it.
- Graphs and tables need to be visible immediately (sometimes they are not) when a student opens the software.
- A teacher was unable to link SimCalc into the school’s interactive whiteboard software (possibly a software license matter).
- The students want to play with the software; it takes discipline to do the graph first then play.
- Recording their ideas on paper means students have to record their ‘mistakes’; it would be better if they could record on screen, so that their mistakes are not seen.
- It is important that the teacher knows how to use the technology to avoid wasting time in set-up.
- Teachers need access to laptops for their classes so that students do not lose learning time in setting up computers.

Finally, students in one school thought that SandCircle might make money from their work (i.e. from the games to which the context of the unit related). It is possible that the context was too convincing for these students. However, in other cases, teachers argued that the context was not realistic enough and that their students were disappointed at not being able to design their own mobile phone games.

The points listed in this section are not necessarily major criticisms, but are seen by teachers as helpful amendments that might improve an already useful unit.

4.7 Mathematics in the 'Everyday' World

This section addresses the final research question:

RQ7. Do the materials help students see the role of mathematics in the 'everyday' world?

4.7.1 Teachers' views

Teachers were relatively positive about the impact of the unit on students' understanding of mathematics in the real world. Only one felt that the unit does not make real life links, whereas nine felt that the unit does this. However, teachers' answers revealed different interpretations of the term. Some answered in terms of 'real-life maths' as indicated through realistic contexts; others interpreted it in terms of students being able to apply learning elsewhere in their mathematics lessons or applying mathematics in real life (e.g. understanding a newspaper article).

Four teachers thought there was an impact on real life mathematics to some extent. They mentioned three specific contexts (Wendela, money, using graphs to balance finances) and highlighted another which was considered less functional (Crab Velocity). One teacher commented that the contexts encourage students to have a limited concept of 'a mathematician'. As mentioned earlier, several teachers also noted that the context misled some students, who thought they would be designing or making their own mobile phone games; some teachers considered this was a missed opportunity to engage students in real life mathematics.

4.7.2 Students' views

The majority of students in the focus groups thought that it is important for the mathematics they do at school to relate to everyday life, with most students rating the importance between 8 and 10.²⁰ Most students thought that SandCircle helped with this: 70 per cent of students who completed the questionnaire thought that SandCircle helped them understand how mathematics is used in everyday life. However the focus group discussion revealed that students found it difficult to explain or give specific examples of how SandCircle impacts on their daily life. When the focus group students were asked about this, the most popular responses were, jointly, improving computer skills and shopping (handling money).

While many students found it hard to express how SandCircle mathematics relates to their everyday activities, they found it easier to talk about application to their future adult life, and they thought that SandCircle helped them become more aware about future careers: 70 per cent of students agreed or strongly agreed that SandCircle broadens their understanding of how mathematics is used in different jobs. Students thought that banking and finance are both areas where their SandCircle experience would help them. Several students mentioned that SandCircle would help them read graphs showing financial profit. Many students believe that SandCircle mathematics

²⁰ 1 = not very important; 10 = very important

improved their IT skills, which can be useful in “*any job*”. In contrast, a handful of students could not imagine how they would be able to apply what they were learning in SandCircle to their future life: “*I don’t see myself looking at a graph in the future and saying, x equals y* ”.

Eighteen per cent of students agreed or strongly agreed with the statement (about mathematics generally) that ‘mathematics is not relevant for the career I want to do’. However most students thought that mathematics will be relevant to their career or they were not sure if it will: 43 per cent of them disagreed or strongly disagreed with the statement (i.e. these students believe that mathematics *will* be relevant for their careers), and 39 per cent of students were not sure whether mathematics will be relevant.

Making the relevance of the unit clearer to students (and teachers) in the SandCircle materials may help students to understand it better and to relate the unit more easily and more specifically to their own lives.

5. Conclusion

This report has outlined outcomes from the evaluation of the first Cornerstone Maths unit to be piloted in England: SandCircle Mobile Games. The key findings and recommendations arising from them were presented in sections 1 and 2 of this report and further detail about each finding is given in the body of the report.

Overall, the findings show that there is definite potential for Cornerstone Maths to be a success in England as there is evidence from this small group that significant progress is made with learning. At the same time, the evaluation has highlighted areas that might need some attention in ensuring a smooth transition from the US context.

Providing independent evidence to improve education and learning.

© 2012 National Foundation for Educational Research



**National Foundation for
Educational Research**
The Mere, Upton Park,
Slough, Berks SL1 2DQ

T: 01753 574123
F: 01753 691632
E: enquiries@nfer.ac.uk

www.nfer.ac.uk

CMP