# ZDM - The International Journal on Mathematics Education


## Cornerstone Mathematics: Designing Digital Technology for Teacher Adaptation and Scaling

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | ZDMI-D-13-00013R2 |
| **Full Title:** | Cornerstone Mathematics: Designing Digital Technology for Teacher Adaptation and Scaling |
| **Article Type:** | Issue 2013 - 8 |
| **Keywords:** | design-based research; digital technology; dynamic representations; professional development; instrumentalisation; adaptation; scaling |
| **Corresponding Author:** | Celia Hoyles, PhD<br>London Knowledge Lab, Institute of Education, University of London<br>London , UNITED KINGDOM |
| **Corresponding Author Secondary Information:** | |
| **Corresponding Author's Institution:** | London Knowledge Lab, Institute of Education, University of London |
| **Corresponding Author's Secondary Institution:** | |
| **First Author:** | Celia Hoyles, PhD |
| **First Author Secondary Information:** | |
| **Order of Authors:** | Celia Hoyles, PhD |
| | Richard Noss |
| | Phil Vahey |
| | Jeremy Roschelle |
| **Order of Authors Secondary Information:** | |
| **Abstract:** | We report the results of a design-based research project in England that embeds digital technology. The research followed from two phases in the United States: (1) a design phase that used dynamic representations to foster conceptual understanding of hard-to-teach mathematical ideas, and (2) a research phase that measured the efficacy of the resulting technology-based curriculum units as implemented in Texas schools. The goal of the third phase in England was initially to "scale up" the U.S. approach. We determined, however, that the materials had to be re-designed for adaptability by English teachers. We report how the features of the innovation—particularly its technological infrastructure—could be leveraged, not only to achieve positive learning outcomes, but also to lay the foundations for change in pedagogy and learning at scale. We identify an emergent framework of design affordances for teacher adaptability that are particularly salient when technology is a critical element. |

**Modifications undertaken in response to the recommendations of the reviewers re article:**

Cornerstone Mathematics: Designing Digital Technology for Teacher Adaptation and Scaling    C. Hoyles, R. Noss, P. Vahey,  & J. Roschelle

Thank you and the reviewers for the careful and perceptive remarks. Responding to their comments has provoked us to take another long and careful look at the paper.  We are trying to be ambitious in this publication: we report design research, which was, on the one hand derived from a prior project in another country thus raising issues and dilemmas, while on the other hand was intended as a basis for scaling, something which is itself a challenge for mathematics innovation in general and maybe even more so for innovation embedding technology.

We have made both of these points more clearly in the paper, and, we hope you will find the paper improved and satisfactory now for publication in ZDM. In particular we have:

1. Re-written the abstract and indeed much of the paper (and modified its structure) to include clarifying the relationship of the UK design research reported to the prior US work and how one of its aims is to leverage the technological infrastructure to achieve change at scale (this was at least partly provoked by the comments of reviewer 2 last two sentences).
2. Reframed the quantitative results from the United States so that the point of reporting on the US study is more clear: to ensure that the changes made to the materials and CPD did not reduce the amount of student learning found in the US
3. Made the idea of teachers' instrumentalisation more central to the overall argument
4. More fully described our data and analysis methods, and adjusted the order of the paper to include
    - details of the approach taken and samples in both the original US study  and the design research in England
    - specific research questions that were investigated using the qualitative data
5. Clarified how we derived the qualitative results and included some general evaluations rather than only positive ones
6. Recast the conclusions and we hope clarifying their warrants as well as where they will lead in terms of the scaling of the project
7. Adjusted the Figure.

Please do not hesitate to contact us if you have any additional questions or comments.

# Cornerstone Mathematics: Designing Digital Technology for Teacher Adaptation and Scaling[1]

C. Hoyles*, R. Noss*, P. Vahey**, J. Roschelle**

**London Knowledge Lab, Institute of Education, University of London, UK*

*** SRI International, Menlo Park, CA, USA*

Email: c.hoyles@ioe.ac.uk

Phone: +44 (0) 20 7911 5577

Fax: +44 (0) 20 7763 2138

---

[1] We gratefully acknowledge funding by the Li Ka Shing Foundation. The research has entailed intensive collaboration with the following researchers in the Center for Technology in Learning, SRI International, Menlo Park, California, USA: Jennifer Knudsen, Ken Rafanan, Teresa Lara-Meloy, Anna Werner, Gucci Estrella, and Nicole Shechtman.

Abstract

We report the results of a design-based research project in England that embeds digital technology. The research followed from two phases in the United States: (1) a design phase that used dynamic representations to foster conceptual understanding of hard-to-teach mathematical ideas, and (2) a research phase that measured the efficacy of the resulting technology-based curriculum units as implemented in Texas schools. The goal of the third phase in England was initially to "scale up" the U.S. approach. We determined, however, that the materials had to be re-designed for adaptability by English teachers. We report how the features of the innovation—particularly its technological infrastructure—could be leveraged, not only to achieve positive learning outcomes, but also to lay the foundations for change in pedagogy and learning at scale. We identify an emergent framework of design affordances for teacher adaptability that are particularly salient when technology is a critical element.

*Keywords*

*design-based research; digital technology; dynamic representations; professional development*

# 0 Introduction

The first two authors have argued that changes in the computational domain open up only the potential for change, not actual change in the didactical field (Hoyles and Noss, 2003). We cited Kaput (1992), who attributed the continuing marginalisation of technology in mathematics education to the complex issues that surround its use:

- Technology requires continually rethinking pedagogical and curricular motives and contexts.
- Classroom-based research is difficult, because exploiting the real power of the technology requires such innovative approaches that comparison to a traditional class is inappropriate.
- The practical complications of student access to computers, cost of software, and development of curricular materials often prohibit research.
- Given rapid changes in technology, research is often out-dated by the time it is complete.

Despite substantial developments in theory, and massive changes in technology, the core challenge of 'implementation' remains: how to ensure that digital technology is used at all in mathematics classrooms, and, if it is used, how to enhance mathematical thinking rather than simply reiterating current practice or, as is sometimes the case, circumventing mathematics altogether. How can research inform efforts to embed technology in transformational ways that support scalable models of classroom change?

This paper reports on research undertaken as part of the Cornerstone Mathematics (CM) project, a collaboration between researchers in the United States and England for a moderately large-scale, design-based implementation project in

England. U.S. researchers completed two phases of work before CM came to England: (1) design of the core use of dynamic representations in 'modules' that include curriculum workbooks and teacher professional development; and (2) efficacy trials that established causality linking the modules to improved student conceptual understanding of challenging mathematics. The goal of the project in England was eventually to reach 100 or more schools using the US-developed materials and approach. However, the team determined that before going directly to scale, a further phase focused on designing materials and processes to support teacher adaptation and instrumentalisation was needed.

As well as presenting results from research focused on these design concerns, we also present hypotheses about the key features required for scaling a technologically-based innovation in the field of mathematical learning.

# 1 Theoretical background

Sinclair et al.'s (2010) model of the development of digital technologies in mathematics education describes a shift in research attention from "Wave 1" with an exclusive focus on the relationship between individual learners and mathematics, to "Wave 2," which involves the broader context of learning, the teacher, and the curriculum. The authors argue that this trajectory is almost inevitable. Any project that seeks to scale demand must, in addition to assessing individual students' interactions with technology and teacher/peers, address the curriculum and the teacher's role in using and deploying the digital technologies. Hoyles and Lagrange (2009) provide an overview of these trends.

Two challenges for the current research were to identify existing or evolving theories that (1) provided an adequate underpinning for such a complex, intertwined set of components (technology, teachers and students, curricular context, materials), and (2) supported the generalisability of research findings. Drijvers, *et al.* (2010), who provide an overview of the theoretical frames used in technology-related research in mathematics education, argue that integrative frameworks allow for articulating different theoretical perspectives. This paper represents a step towards developing such a framework. It takes as its starting point the curricular activity system framework described in Roschelle *et al.* (2010) and Vahey *et al.* (2013). This framework is based on the recognition that an instructional activity includes a learning objective, available materials, the intended use of tools, and the roles of diverse participants. However, focus on such activities is not sufficient. An activity must fit into the structure of those classrooms that are expected to engage in the activity. The curricular activity systems approach integrates learning requirements, teacher professional development (PD), curriculum materials, and technology; the approach recognizes that these elements are situated in an educational context that includes people, conventions, and policy considerations.

We also look to design-based implementation research (Penuel, Fishman, Cheng, and Sabelli, 2011), which expands on traditional design of an artefact or intervention by developing the capacity for sustained change. Central to our scaling approach is building teachers' capacity as they both strive to adopt the CM materials, and participate in adapting and instrumentalising those materials. By building teachers' capacity to use the materials effectively and still maintain the core pedagogical approach, we optimise our chances of scaling up to large numbers of teachers, classrooms, and students.

We first provide a brief outline of the U.S. studies that formed the basis for the research in England.

## 2 Preliminary Phases in the United States: the SimCalc Intervention

Over the course of about 6 years, the U.S. team developed integrated technology, curriculum, and professional designs to address big ideas in the algebra and pre-calculus strand of mathematics. The *technology design* emphasized linked dynamic representations such as graphs to motions, tables, and equations (Kaput *et al.* 2007; Vahey *et al.* 2013; see Figure 1). Many evaluations show, at best, small positive effects for technology (Cheung and Slavin, 2011; Dynarski *et al.*, 2007). However, dynamic representations have been shown to be powerful in developing conceptual understanding (Heid and Blume 2008). Consistent with the Multimedia Principle (Fletcher and Tobias 2005), those representations can help students make important connections between intuitive and formal ideas, and between graphical and linguistic understandings. The *curriculum modules design* is based on organizing key mathematical ideas and important mathematical practices into integrated paper- and technology-based materials that are straightforward to use in classrooms. Within the modules, students solve increasingly challenging problems in one content area organized around a high-interest theme such as 'Designing Mobile Games'. Teachers focus on the key mathematical ideas, valuing conceptual understanding and guiding development of mathematical practices. The *professional development design* emphasizes practical 'mathematical knowledge for teaching' (Hill *et al.,* 2008)—the knowledge teachers need to make sense of and extend their students' mathematical reasoning. This design also emphasizes technological pedagogical content knowledge (Mishra and Koehler, 2006), which is the knowledge teachers need to teach effectively with technology. Modules emphasize a small set of teaching practices (Ball and Forzani, 2009; Hammerness, 2006), with a focus on supporting argumentation in the classroom (Kim, 2012).

This integrated design of technology, curriculum, and PD was formalized as an instructional package that could be implemented at scale. Subsequently, a randomized controlled trial (RCT) in Texas investigated whether a wide variety of teachers, when provided with appropriate PD and a replacement unit that integrated curriculum and software, (the "SimCalc intervention"), could increase student learning of important and complex mathematics. In this RCT, teachers were randomly assigned to participate in the treatment or the control group. The treatment group received the integrated intervention, which began with a 3-day teacher PD workshop in which teachers learned to teach using the technology-based unit. Treatment teachers were then asked to teach the replacement unit in place of their usual unit on linear function, and most teachers completed the unit in 2-3 weeks. Teachers in the control group received PD on the integration of technology into mathematics teaching, but the PD addressed different content.

Working with five Texas Education Service Centers (the primary support and PD providers in Texas), teacher volunteers were recruited whose students reflected the regional, ethnic, and socioeconomic diversity of the state. Complete data were returned by 56 teachers and 825 students. At intake, the Treatment group (33 teachers) and the Control group (23 teachers) did not differ in any important way (e.g., with respect to teaching experience, ethnicity, gender, mathematical content knowledge, or by socioeconomic status as indicated by percent of students eligible for free or reduced-price lunch in school). The greater number of teachers in the Treatment group was an artefact of teachers' scheduling conflicts with the workshops to which they were assigned. Because teachers were not informed about the workshop type until the workshop occurred, the consequences for randomization and thus the validity of the experiment were minimal. The attrition rate was comparable to other large experiments with educational technology (Dynarsky e*t al.*, 2007), and no evidence suggested differential attrition, which would be the principal threat to validity.

The primary outcome measure in this study was student learning of core mathematical content. The research team employed a bespoke assessment instrument. Working with a panel of mathematicians and mathematics education experts, the assessment encompassed both simple (M1) and more complex (M2) aspects of linear function. The simpler items (M1) were based on those used in existing standardised tests in Texas; for example, students were asked to calculate using a linear relationship represented in different ways. More complex M2 items required comparing multiple rates or finding average rates, and typically the rate information was not provided directly; for instance, students had to infer it from slopes in graphs. The assessment was administered in a single class period to students immediately before and after their linear function unit was taught.

The analysis of student gain scores from pre-test to post-test showed a large and significant main effect, with an effect size of 0.56 (Roschelle *et al.*, 2010). This effect was robust across a diverse set of student demographics. Students who used the SimCalc materials outperformed students in the control condition regardless of gender, ethnicity, teacher-rated prior achievement, and socioeconomic status. This finding provides evidence that the use of dynamic representations, when embedded in a set of replacement units designed within a curriculum activity system framework, can result in substantial learning gains.

# 3 The Cornerstone Mathematics Project

CM seeks, as did its U.S. predecessor, to exploit the dynamic and multi-representational potential of digital technology to enhance learners' engagement and understanding of mathematical ideas. CM consists of four units, each focused on key mathematical topics in middle school (students aged 11-14 years). Each unit embeds activities in a quasi-realistic digital context in which students *need* to use mathematical knowledge to achieve their—and our—goals. This is an important design decision, because although mathematics is a high status and compulsory subject, students still need to be motivated to think mathematically and to do so by tapping into their digital lives and making the work "realistic" (e.g., Confrey *et al.*, 2009).

Unit 1, the focus of the research reported here, concerns piecewise-linear functions, with the work in England replicating the SimCalc intervention (see Roschelle et al. 2010; Roschelle and Shechtman 2013). However, for success at scale in England, two levels of changes were required. For the initial, and simpler, level, the mathematical content and language were adjusted to fit the English setting. For example, numerous changes in vocabulary had to be made to account for differences between U.S. and UK English usage and spelling. Changes were also made to conform to local school mathematical conventions (e.g., $y$=m$x$+c rather than $y$=m$x$+b). Third, we sought to align the unit to the English National Curriculum, which is statutory and places greater emphasis on 'mathematical processes'.

This paper focuses on more complex changes. In an RCT, experimenters try to hold the 'inputs' fixed and attempt to have teachers adopt the materials with high fidelity. However, in scaling up, innovators recognize that teachers and schools are not identical, and the metaphor of teacher adoption will not suffice. On the contrary, it is inevitable that each new classroom, teacher, and school context introduces new features that could knock the innovation off course. A key challenge, therefore, was to move beyond adoption and design for *robustness* (Roschelle *et al.*, 2008). We did so by encouraging teachers to appreciate the goals of the innovation, and to make the innovation their own through a process of adaptation and instrumentalisation. Thus the more complex change, and the focus of this paper, involved designing the materials to better support this intended process.

We wanted all CM teachers to develop pedagogical and technological confidence through PD and supporting their teaching of the unit. We also wanted teachers to work together in a mutually supportive community through a blend of face-to-face and virtual interaction: the latter is becoming recognised as important for scaling, especially when teachers are widely distributed (e.g., Baker-Doyle, 2011). To check that the ways we supported teacher adaptation and instrumentalisation of CM Unit 1 did not undercut the effectiveness of the materials (as established earlier), we conducted an evaluation of impact that was designed to ensure that student learning gains were preserved.

### 3.1 From the SimCalc Intervention in the United States to CM Unit 1 in England

CM Unit 1 started from the U.S.-based unit called *Designing Cell Phone Games*, which has been described in prior research (e.g., Roschelle *et al*., 2010; Vahey, Roy & Fueyo, 2013). In this unit's constructionist design, users learn through interaction with and feedback from digital tools that enable them to explore, build, and learn (Papert 1980). Constructionism has been the subject of extensive research and development and continues to result in innovative ways of designing tools and in work with learners worldwide (e.g., Kynigos *et al.,* 2012, Noss & Hoyles, submitted: 2013). At the same time, a complementary strand of research emphasizes the importance of 'instrumental genesis' for both students and teachers, with artefacts transformed into 'instruments'; that is, systems with which the user gains fluency and expressive competence (e.g., Drijvers and Trouche, 2008; and, in the context of SimCalc research, Roschelle *et al.,* 2008). Digital technologies not only add new representations (or link old ones), but research has increasingly found that digital representations change the epistemological map of what it is intended for teaching and learning (Noss and Hoyles, 1996; Kaput and Roschelle, 1998).

CM Unit 1 highlighted the following mathematical concepts: coordinating algebraic, graphical, and tabular representations of linear functions; $y$=m$x$+c as a model of constant velocity motion; the meaning of m and c in the motion context; and velocity as speed with direction. All activities were set in the context of using mathematics to design computer games for mobile phones, where functions must be used to make game characters move in appropriate ways. At the heart of the software environment (SimCalc in this Unit) was a simulation, or a 'journey', of an object that could be tracked in a graph and a table, as well as captured in algebraic or narrative form (see Figure 1). Students therefore receive feedback on any journey they have constructed by visually 'seeing it happen'; the mathematics plays out in terms of motion and vice versa. Students can control their object's journey by manipulating the position-time graph or its algebraic representation. The constructionist key in CM Unit 1 is that students themselves can intervene in the 'system' by constructing journeys and exploring them alone and with others.

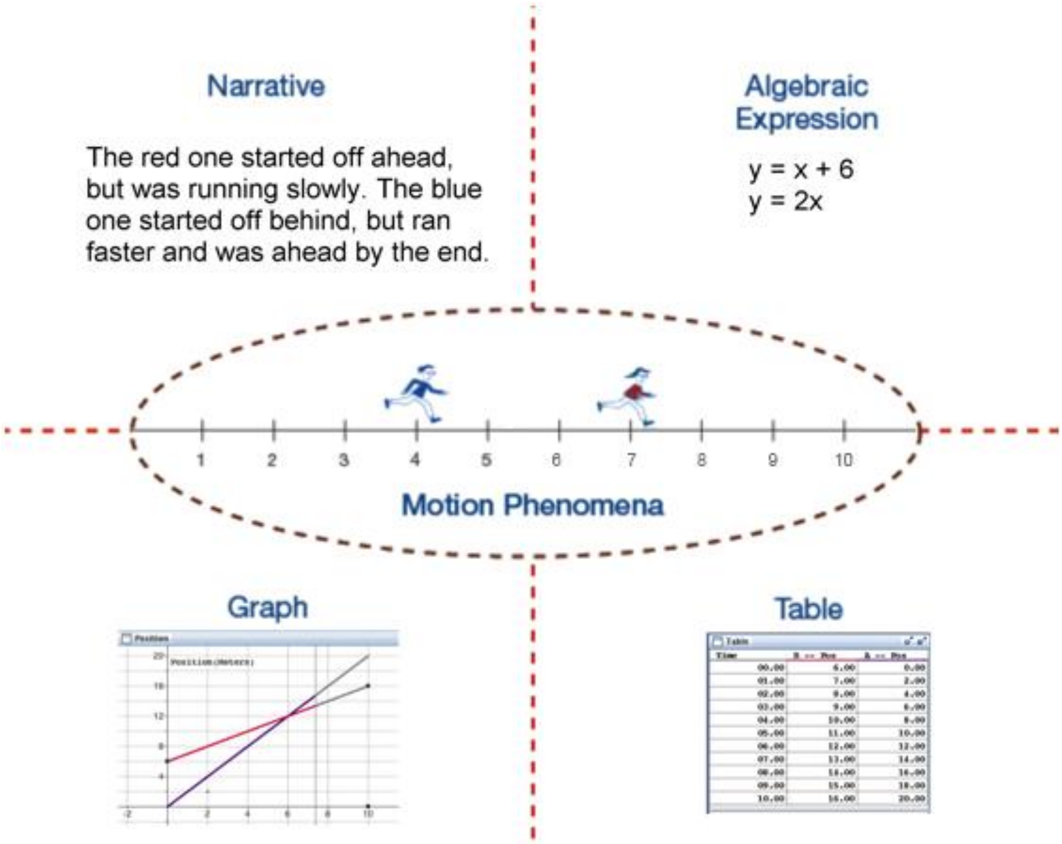


Figure 1: The SimCalc system links algebraic expressions, graphs, tables, and narratives through the phenomenon of motion

A first priority for the design research in England was revisiting the style and emphasis of the PD needed, not least because middle-school students in England are taught by secondary trained mathematics teachers, in contrast to the U.S. teachers who were mainly primary (elementary) trained. Teachers in England, and particularly subject leaders, expect a degree of autonomy in approaching given mathematical topics.

As a second priority, teacher support in England required a different character. In the U.S. efficacy trial, the researchers wanted to control precisely the PD, and thus they and PD experts provided it. However, in England, teacher adaptation and instrumentalisation were emphasised. For example, at the request of the teachers, more group-work was introduced in the PD and teachers orchestrated this themselves as a means to share experience and reflections. In addition, PD was extended through orchestrated peer interaction among the CM teachers in an online community, which was part of the technical national infrastructure provided by the National Centre for Excellence in the Teaching of Mathematics (NCETM), a nationwide professional learning community for mathematics teachers (www.ncetm.org.uk). Through the CM online community, teachers could share their experiences, the prompts and probes they had planned or used spontaneously, and the activities they designed themselves.

### 3.2 Sample

The initial sample for CM research consisted of 10 schools, with 2 teachers recruited from each school. One school was unable to complete its participation, however, because of a fire in the school buildings, and one teacher took maternity leave. The final sample comprised 9 schools, 17 teachers and classes, and 429 students drawn from Years 7 through 9 (students aged 11-14). We were careful to ensure that the schools and the classes exhibited a wide diversity of school contexts and prior student achievement. Although we can hardly claim that these schools were representative of all schools in England, the sample schools had a range of: academic strength (as measured by public examination results), student intake in terms of socioeconomic status, technology infrastructure, and teachers' experience of mathematics teaching. The school sample included one "public" school (an exclusive private school) with a privileged intake, and one state-run school whose students were economically disadvantaged and in which English was an 'additional language' for a high proportion of students. The teachers, too, were diverse in their teaching experience, including one new to the profession, most with 3-5 years of experience, and some highly experienced mathematics subject leaders in their school. Some of the teachers had excellent mathematics qualifications and some less so. The 429 students who engaged with Unit 1 were distributed across different year groups (according to teacher choice): Year 7: 179 (42%), Year 8: 227 (53%), Year 9: 23 (5%).

The intervention as a whole, consisted of 2 days of PD for all teachers, 0.5 day for a reflection/feedback session after implementation, and ongoing discussion through the online forum. The unit was taught over 2 to 3 weeks (spread over a month for the different schools) with students using computers in most lessons (a practice that was unusual in mathematics class). The hardware used varied across the sample, ranging from desktops in computer labs, through laptops in classrooms, to the extensive use of the interactive whiteboard.

### 3.3 Data Sources

The sources of data on the processes and outcomes of the study were collected by the following means:

a) A *proforma* that a researcher completed in discussion with each school's main contact, generally before a school visit. The proforma was designed to establish a baseline 'context for teaching', indicating the reasons why staff in these schools had chosen to participate in the study, background information about mathematics teaching in the school, expectations for the study, and any challenges anticipated.

b) *Lesson observations* in one or two participating classes in each of the 9 schools. We observed 16 lessons across 9 schools. During each observation, we took notes using a semi-structured protocol around the following main themes:

- Manageability: in particular how teachers orchestrated the lesson content and pace

- Engagement/behaviour (attitude, engagement and understanding, behaviour and interaction)
- Teaching style/approach (interactions and interventions).

c) *Post-observation interviews with the teacher(s)* of the observed class(es) in 9 schools. The 16 observed teachers in the 9 schools participated in an interview, covering the participation of 17 classes (one teacher talked about his own and a colleague's experience). The interviews served to collect relevant information in the following categories: the teachers and students observed; preparation for teaching the unit; the implementation of the unit; the use of the technology; and the impact of the unit on students in terms of their engagement and learning.

d) *Students' questionnaire responses* completed at or towards the end of the unit (eight schools completed and returned the responses). Students in the observed classes were asked to complete a questionnaire about their views of mathematics lessons generally and about CM Unit 1 in particular. Most teachers administered the questionnaire to their classes after the day of the observation, close to the end of the unit.

**e**) *Focus-group discussion* with a sample of students from nine schools. Teachers selected six-eight students for each focus group and were advised to provide a mix of students (e.g., of different ability levels or different language backgrounds).

f) *Pre-test and post-test data* Students completed two identical tests to evaluate their learning: a pre-test, before the unit was taught, and a post-test, taken at the end of the unit. These tests had been used in the Texas study and incorporated both M1 and M2 questions.

Interview and observation schedules were consistently followed, and the interviews and focus-group discussions were audio-transcribed.

# 4 Results

We begin with our quantitative results in which the analyses followed that undertaken in the U.S. study. This quantitative analysis did not seek to establish the efficacy of the intervention; doing so would have required a much more expensive and controlled design, such as that previously conducted in the United States. Instead, the research question was: 'Is the pattern of student learning gains in England consistent with the pattern of learning gains observed in the United States?' A consistent pattern would reassure the team that the changes to the design for the England scale-up were preserving the effectiveness of the materials and approach.

The following sections look at three qualitative research questions:

1. What changes supported teacher adaptation?
2. How did teachers come to greater ownership of the materials?
3. As the approach started to scale up, were teachers engaged in a community around it?

Finally, we summarize results from an external evaluator.

### 4.1 Quantitative Results: Consistent Pattern of Learning Results

Although we had no comparison group in England, the underlying unit, assessment and assessment method used in the Texas RCT was closely followed to allow a quasi-experimental comparison with pupils in that study. Pre- and post-tests for each student generated gain scores that were used for a quantitative analysis of learning gains, which were then compared with the Texas student results.

The pre- and post-tests were scored in London by a group of 12 pre-service mathematics teachers recruited from the Institute of Education. Using a web-based data entry and verification form, scorers entered their marks and comments online; this procedure allowed instant access to the data while ensuring rigorous methods consistent with prior methods used in the United States. All scorers were trained on the scoring key and rubrics, and were permitted to begin scoring actual assessments only after they had attained a minimal level of gold-standard marking of sample assessments. To measure inter-rater reliability during the scoring process, 10% of all assessments were scored by a second scorer. The inter-rater reliability was found to exceed 90%, well above the acceptable level.

Using the pre-test as a proxy for equivalence of prior background knowledge, we found that the groups in the two studies were similar. The researchers fitted a series of hierarchical linear models (HLM: Raudenbush and Bryk, 2002) to examine statistical equivalence at pre-test. These analyses showed no statistically significant differences for the CM and U.S. samples for the total score [$\beta = -0.64$, $p = 0.69$], M1 subscale [$\beta = -0.50$, $p = 0.52$], or M2 subscale [$\beta = -0.15$, $p = 0.86$]. The variation among pupil pre-test scores in both groups was also sufficient to rule out floor effects as a possible explanation for the apparent equivalence of prior knowledge. In a non-significant trend, CM pupils had slightly lower pre-test scores, which may be attributable to age differences. The determination that the groups were equivalent at pre-test is important because in a quasi-experimental comparison such as this, the primary threat to internal validity is the possibility of non-equivalence of groups at baseline.

Figure 2 shows the learning gains for pupils in the Texas Control group (who did not use the materials), the Texas Intervention group (who did use the materials), and the CM pupils in England. The dark part of the bars shows learning for M1, ‘simple’ linear functions, and the light part of the bars shows learning for M2, ‘complex’ linear functions. The learning gains for CM pupils were similar to the learning gains for pupils in the Texas Intervention group, and were significantly higher than for pupils in the Texas Control group.

These data show that the CM approach successfully met the goal of increasing pupil learning of important mathematics in England. Further analyses indicated that the materials were equally effective for students with different levels of prior mathematics achievement: schools with higher prior General Certificate of Secondary Education (GCSE) scores performed higher on pre-test, which was to be expected; however, *learning gains* were not correlated with school achievement level, indicating that the materials were effective for pupils from a variety of school contexts.

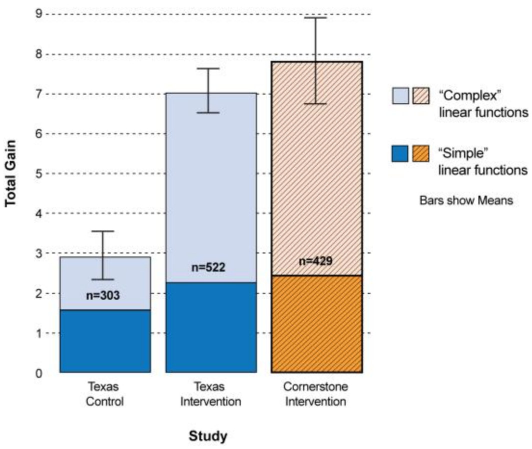


Figure 2 Learning gains of pupils in the Texas and England studies showed similar learning for both sets of students.

These analyses provided important evidence about the effectiveness of the materials. Although we would not advise drawing any conclusions from cross-country comparisons (i.e., success in the United States vs. success in England), comparison between the magnitude of gains among the U.S. control, U.S. treatment, and CM groups provides evidence for feasibility of effectiveness in the English context.

It is noteworthy that although gains were similar in 'simple' linear functions for all groups (although the gains for the control were slightly lower than those for the other groups, that difference was not statistically significant), the difference in groups was predominantly for complex concepts.

Figure 3 shows an example of a complex M2 item, and a sample student response in the pre-test. The student had been provided with the questions and gridded graph areas, and the student drew the lines on the graphs. We note, in this one pupil's response, several well-known conceptual difficulties, including treating a position-time graph as a velocity graph, and representing backwards motion as a line that goes back toward zero on the time axis.

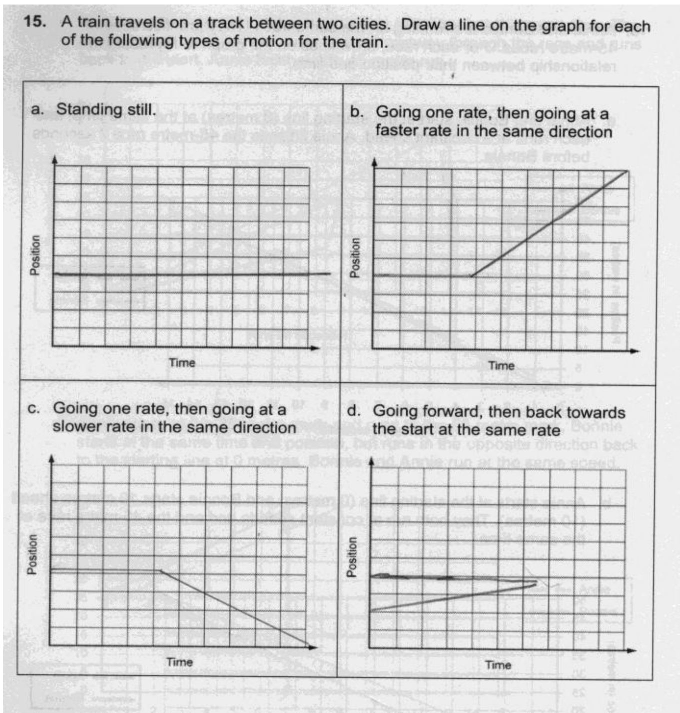


Figure 3: Example of a complex linear function item, Question 15, and a student response in pre-test.

| % Correct | Q15a pre | Q15a post | Q15b pre | Q15b post | Q15c pre | Q15c post | Q15d pre | Q15d post |
|---|---|---|---|---|---|---|---|---|
| CM Year 7 | 21.2 | 56.4 | 11.7 | 51.4 | 9.5 | 46.4 | 7.3 | 41.9 |
| CM Year 8 | 45.4 | 77.1 | 28.2 | 73.6 | 20.3 | 71.8 | 14.5 | 63.9 |
| Texas Control | 40.3 | 52.8 | 30.7 | 44.6 | 24.8 | 35 | 21.1 | 30.4 |

Table 1: Scores on Question 15 showed increased learning for CM Pupils between the pre- and post-test.

As shown in Table 1, the changes in student scores for this question are substantial, and compare very well with the (older) control group students. In the next section, we track how these changes may have been catalysed by the design and structure of the CM unit, and its deployment.

### 4.2 Qualitative Results: External Evaluation.

To assess the validity and reliability of the qualitative research, an external evaluation of the implementation of Unit 1 was undertaken (Sturman and Cooper, 2012). The main findings of this evaluation were, in summary:

1. The unit was generally manageable to implement and it helped students to learn the difficult concepts covered.
2. Perceptions of the groups that the unit affected most were varied, but test results suggest that the unit was appropriate and effective for the whole

range of students, with no group conclusively progressing better than others.

3. The unit had a good curriculum fit and covered more than most teachers' own schemes of work. When teachers could make a prediction, they felt that students would do well in their next assessment, based on their solid understanding and willingness to attempt questions.
4. Teachers were positive about the unit's ability to engage students, although students themselves expressed some reservations. Despite this, students acknowledged that the unit helped them to learn.
5. The unit was generally seen as useful, with most of its parts more likely to be rated 'helpful' than 'unhelpful'. Some technical difficulties were encountered, and some changes and additions to resources were suggested.
6. Teachers were generally positive about the impact of the unit on students' understanding of mathematics in the real world, although they identified different ways of how the unit achieved this. Students also perceived real-life benefits.

These general findings provide a backdrop for the more detailed analyses we present below.

### 4.3 Qualitative Results: Teacher Adaptation.

The themes around which the observation, interview, and student focus group data were analysed included the extent of teachers' preparation for teaching the unit; the implementation of the unit; technology instrumentation; and the impact on students in terms of their engagement and learning. To identify relevant episodes, data were provisionally coded, assigned to one of these themes. and sub-divided into the appropriate categories .

In a major adaptation of the CM unit most, (15 out of 18) of the teachers modified each lesson to conform to the common '3-part lesson' format (starter, student activity, plenary consolidation), a format that the UK government's "National Strategy" for education had effectively mandated during the years 2003-11. Although this adaptation was common, the specific adaptations took different forms: some teachers devised new starter activities, while others creatively varied the kinds of orchestration they adopted (e.g., student groups sharing their work on the interactive whiteboard).

Teachers also adapted the pace of implementation. Teachers had to adapt learning activities to a variety of student groups (year, set, level of 'English as Additional Language') and lesson formats (single or double periods, a PC room, or laptops in the classroom), and they broadly managed to do so. We found, however, that teachers' planning of timing was not precise (all teachers were given time for planning, and all created an initial plan of around 8-10 hours of lesson time, which in practice ranged between 8 and 20 hours). Such inconsistencies may be inevitable for novel curriculum activities, given that teachers can learn how to use the unit only by teaching it. It is hardly surprising that a disruptive technology would lead to unpredictable timing of lessons. But it is surprising that the teachers were so positive about this disruption and the extent to which it encouraged—perhaps forced—new thinking about pedagogic strategies. They chose to elaborate on what the actions on the software meant mathematically (couched in the language of the unit), with some adding new activities for groups to work on together. We believe these adaptations were crucial to the success of the unit—the students were not simply drawing the graphs or changing the slope, but were engaging with digital mathematical objects as they thought deeply about the mathematical concepts.

In summary, all teachers adopted the CM activity sequence, with 13 out of the 18 adapting it by some subset of:

- Compressing/stretching activities, according to the teacher's expectations of how students would react, with adjustments in the course of the lesson
- Switching on an *ad hoc* basis between pair and whole-class work if interesting ideas or points of explanation arose
- Introducing activities to 'consolidate' the mathematical ideas (e.g., one 'matching' activity in the activity booklet was represented as a card-sort activity for a starter in the next lesson)
- Adding short activities at the ends of lessons that could be used for consolidation in class or as homework.

Some teachers were more open to adaptation than were others. For some teachers with students with English as an Additional Language, or students needing more exercise in numeracy/arithmetic, the activities opened up extension opportunities to work on developing students' capacities with mathematical language/argument, or numeracy skills, which was not part of the core design. In these ways, the extended time spent on the unit was judged as highly valuable.

### 4.4 Qualitative Results: Teacher Ownership and Instrumentalisation

Teachers in England have, over the last decade, been encouraged through Government strategy and inspection to adopt a surprising level of uniformity in their pedagogic strategies. Doing so has, to a large extent, militated against innovation in teaching style, the adoption of different media, and significant exploration in terms of pace, lesson style, etc. What became evident in this study was the extent to which the technology disrupted this conformity, and in doing so, led the teachers to adopt strategies that were new to them, and which they recognised as being successful in evolving student learning. We identify some examples below.

One example was in lesson style. Some students reported a change in lesson style, which was also evident from the observations and teacher interviews. The students, in particular, noted that typical mathematics lessons were usually dominated by 'copying from the board'. Teachers tended to recognise that students arrived at a deeper understanding mathematics when using CM, where they interacted with technology rather than using their traditional approaches. The curriculum replacement approach thus provided a highly structured and relatively short-term (and therefore safe) setting for doing more with technology, particularly moving away from whole-class demonstrations (by the teacher, or by selected students) via an interactive whiteboard.

Another example of teacher ownership can be observed in the ways in which technology was deployed. All of the teachers managed to use digital technology in almost every lesson, which led to significant changes in their initial reservations about using technology. Indeed, many of the teachers immediately decided to re-use parts of the unit as revision material for older pupils. Thus, teachers considered the materials useful for meeting their own needs and the needs of their pupils beyond the use specified by the study.

More generally, 15 of the 18 teachers used their professional judgement to repackage the material by choosing to teach disparate pieces together, or to decompose one idea into many. Although we recognise that these examples of teacher ownership fall short of the re-design of tasks themselves to exploit the use of the technology to give sense to mathematical concepts (in the manner discussed by Laborde, 1995), we nonetheless interpret these changes as expressions of epistemological autonomy on the teachers' part, a finding that subsequently informed our own revisions of this unit and the design of future units.

More broadly, multiple sources of evidence indicate the ways in which the teachers instrumentalised the innovation considered as a whole. This finding underlines the limitation associated with thinking of an innovation as something to 'implement' or 'deploy'. Considerably more than that is required: the presence of the technology (and of carefully designed workbook-based activities) highlights how teachers implementing the innovation need time and support to make the innovation their own, to reshape it, and to use it to create novel strategies as well as new epistemologies for themselves and their students.

A key focus for teacher instrumentalisation was the use of multiple representations to reveal and address student misconceptions. In our interviews teachers consistently reported that that feedback for actions was a learning driver when successful learning occurred. As one teacher put it:

> 'Having the things that move, the real world simulation, glued … [the other representations] together in a more meaningful way. Normally, [students] would be fine, they would be able to repeat to me that the gradient is the steepness of the line and the y intercept is where it crosses, but somehow having it linked with the simulation really brought that home in a way that I have not seen before.

Another teacher reflected on her students' learning as follows:

> The best thing was [that the software] could visually contradict their misconceptions … the average speeds in Red Riding Hood [example], when the journey was not split into equal parts, it was really noticeable the

> students still wanted to add the speeds and divide by two. But as soon as they put that in as the wolf's motion, it was so clear that the wolf was not arriving where it needed to.

This quote refers to the notoriously hard concept of average velocity, framed in the context of Red Riding Hood (RRH) and the Wolf. The Wolf, which can move only at constant velocity, is trying to arrive at Grandma's house at the same time as RRH, whose speed can vary in a piecewise linear way. Students were asked to create a trip where RRH travels at two velocities and arrives at Grandma's house at exactly the same time as the Wolf. They could then check the graph and animation, and revise if it did not achieve their goal. They were asked to write a story describing RRH's trip.

As the quote indicates, the teacher used the 'multiple mathematical representational' approach—written/verbal, graph, simulation—and the coordination between the context world and mathematical representations in a predict/check/explain cycle that invited students to coordinate their thinking across representations. Students were offered ownership of the journey, as they could build their own narrative about the story as depicted in its mathematical representations. Through observations and interviews we found that teachers were quick to accept the responsibility of addressing student misconceptions that were revealed by this activity, and readily saw the utility of instrumentalising the affordances of the technology to aid them in addressing those misconceptions.

Quantitative evidence from pre- and post-tests corroborated this claim. Figure 4 shows a question related to average speed, and a sample student pre-test response. Although this item was difficult for students, even at post-test, gains were large and significant: at pre-test 2% of pupils got part (a) correct, and 12% got part (b) correct; at post-test 28% and 47%, respectively, gave correct answers. Pupils' qualitative understanding of linear function also increased: at pre-test approximately 20% of pupils answered item 15 correctly (see Figure 3), whereas at post-test 63% answered correctly. Although we are not fully satisfied with the results, we are encouraged by the large number of Key Stage 3 students (i.e., Years 7 through 9 students) whose understanding of these complex and subtle linear function ideas increased, given that many students do not exhibit such understanding even at a much older age.

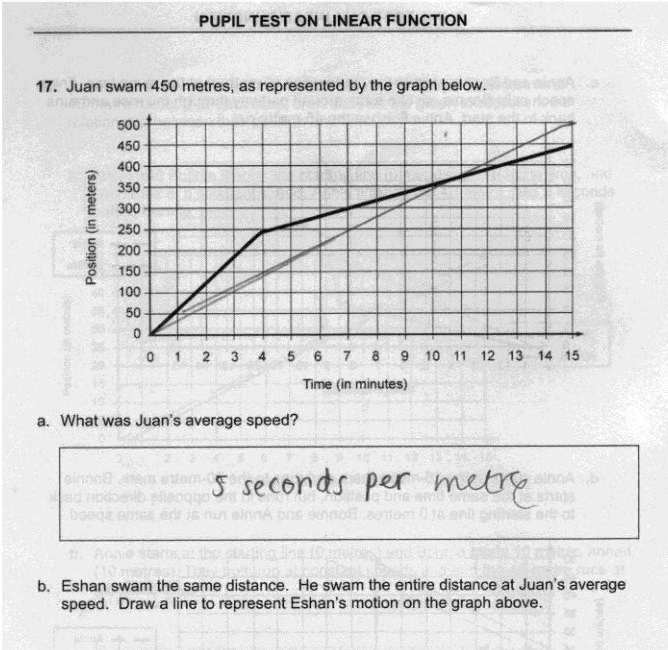


Figure 4: Example of an *average speed* item, and a sample pre-test response.

## 4.5 Additional Results: Sustainability and Community

We have reported here the CM design research for Unit 1. Work is ongoing on CM Unit 2, which concerns geometric similarity. Out of the 9 schools and 17 teachers in our sample, *all* have continued with CM and implemented the second unit of work. As we move to the next phase of the project, 2013-15, all 9 schools have requested to continue their participation. Additionally, two schools have reported that they have already integrated Unit 1 into their *schemes of work.* This is a major step towards sustainability. In England, the mathematics curriculum is institutionalised at the school level through schemes of work planned by the mathematics team, led by the subject leader. Subject leaders negotiate the school mathematics curriculum with their colleagues and instantiate mathematics schemes of work year by year. The schemes must, of course, be aligned with the statutory National Curriculum and pupils progress through the scheme, usually assessed by the school annually to monitor progress. So when CM is part of the scheme, it will have become integrated into the routine of mathematics teaching in the school and taught by the whole mathematics department. In the two cases mentioned above, where Unit 1 has already been integrated, this was instigated by two CM participants who were subject leaders (middle-rank executive school leaders).

However, it proved challenging to spark productive conversations on the online community: the research team made repeated attempts throughout the implementation of Unit 1 to provoke discussion, but was only partially successful in doing so. Only 11 teachers posted; 45 topics were introduced with 104 posts in all, but most topics had only a handful of posts, which mainly consisted of answers to questions or factual statements, and little discussion took place. It seems that online professional discussion was too removed from teacher practice and judged as perhaps tangential to established professional practice.

In response, in the scaling work we are now undertaking, we are building and promoting the online community from the start, setting a norm in the PD sessions of using the online community as a place to engage in asynchronous discussions, and stressing the need to exploit the support that a vibrant community can provide.

# 5 Conclusions

This section elaborates an emergent framework of design affordances for teacher adaptation, ownership, and community that will inform and shape the next phase of our work. Our aim is to scale our work to 100 schools, with a minimum of 200 teachers and 5,000 pupils. Although we address these affordances under three headings, they are, of course, interrelated.

First, in the research context, technology was critical in driving the process of *conceptual and affective change* in teachers, particularly as the technology revealed some mathematical ideas that may not have been initially evident to all of them. In the PD sessions, for example, we noted several instances where the mathematics was unclear to certain participants because of its novel representation, and where the use of the technology encouraged them to reconstruct or enhance their understanding. At the same time, some of the most interesting aspects of the discussions and activities were catalysed by often unsuspected links between representational forms.

Ongoing debate in educational circles has addressed the precise characteristics that technology might best play in helping teachers to exploit its potential in mathematics classrooms. From the qualitative data we derived the role of the technology, although not without problems (e.g., software downloads, computer room availability) fostered teacher (re-)consideration of how they taught, and in some cases *what* they taught. This finding is hardly surprising: the tendency of new technologies to engender re-evaluation of what is possible is well documented (e.g., Noss & Hoyles, 1996).

The leverage of technology reaches beyond such re-evaluations, however, because the use of digital technologies is fast becoming mandatory in England: it confers high status on head teachers, who want to be seen to participate in the latest trends; moreover, policy-makers and inspectors check on whether technology is being used in useful ways. For once we are cutting with the grain.

The second affordance for adaptation is to shift from *scripting* to *steering*. We charted a path through the students' activities, a path that we wanted *all* the teachers to follow. Yet, in a show of autonomy at a PD session, one teacher decided: 'OK, I've got the idea. I'm going to use [the CM materials] but I'm going to do it my way, completely differently'. It is tempting to attribute this to the feistiness of one teacher, and that may indeed be part of the explanation. But we can, perhaps, also take some credit. We organised the CM Unit as one in which the teacher's role was not so much scripted as *steered*. We charted the course, showed them the 'big ideas', but we were at pains to leave room for teachers to 'get the idea' and adapt it in their own way.

In part, the tension between adoption (do as I say) and adaptation (find ways to fit what I say into what you think and do) is exacerbated by the constraints of the system. In England, close alignment with the *statutory* National Curriculum is imperative. And the centralized control of teachers (through the only recently discontinued National Strategies, recommended lesson structure, and inspection, which has deeply influenced most educators in England) is delicately balanced with teacher professionalism.

A major challenge for the future is to encourage teachers to move from adoption to adaptation. The innovation began to develop momentum at the point when teachers started to buy into tool use, to think beyond 'curriculum fit', and to adapt the innovation for themselves in ways that enhanced the tool's epistemic value. Clearly, no rigid dichotomy exists between routine use and epistemological tool use, and we wanted the 'tools to travel between' these poles. This endeavour was in fact successful not least because of the ways in which the teachers adapted the materials. Inevitably, there is a dialectical relationship between what is being fitted and what it is being fitted into. We provided scaffolding and guidance so that teachers could make the innovation their own and make it fit their setting, without undermining the intended epistemic student experience. Because teachers clearly valued the mathematical experience, we are optimistic that, in our next phase of work, we will be able to build on what we have learned to scale up to a much larger number of schools.

A note of caution: positive responses were not universal among the sample teachers. We know that 3 of the 18 teachers simply 'taught the unit' and, despite our best endeavours, might not have been aware of the powerful ideas informing it. They made no adaptions of the materials of any sort. However, these teachers have made efforts to continue with the work and teach the CM units with other groups of students. This seems to be to consolidate their new practice but again maybe for different motives (gleaned from interviews with them) —esteem from being part of a project or even as a way of enhancing career opportunities.

The third affordance is finding the *right grain-size for the manipulable elements of innovation*. Where on the spectrum between closed simulations and completely open programmable systems should we design? What is the relationship between what we would like students to learn, and the technology tools they have at their disposal? How can we design computational interfaces that align with what we are trying to teach? The software design field in education has yet to provide definitive conclusions for the many questions of this type. In our CM research, we have found it useful to generalise these questions beyond software design to the design of other elements of the activity system: the tasks, written materials, and the formative assessment techniques can all benefit from consideration of the precise objects to which we want students, teachers, and researchers to attend. These questions must be resolved before we can be confident of scaling to 100 schools. Our current work, for example, is seeking to determine whether open systems (like SimCalc, Sketchpad, or Cabri) should be replaced by tools that are co-designed alongside the 'curriculum'. Doing so, however, will inevitably limit what is possible in order to enhance what is probable.

This is the strength of the curriculum activity system, but it poses a challenge and tension for a constructionist programme. The teachers' reaction to the innovation consistently stressed the importance of the constructive element (making things happen and attending to them) and the discursive element (building things to share them and create a shared basis for discussion and reflection). However some moved to class demonstration on the interactive whiteboard and in the process sometimes curtailed the time for students to engage in hands-on activity

The most favourable examples of school take-up of CM should be considered against the background of the high stakes associated with mathematics achievement and with the 'grading' of schools; heads (principals) are under pressure *not* to innovate, continuing to use established, conservative pedagogies that minimise the risk of failing to attain expected achievements. Yet our PD sessions revealed a surprising reluctance to view the 'replacement' module as a single entity as we have noted. Instead most teachers adapted its use in unanticipated ways.

We surmise that fostering this autonomy can be designed into our next phase of work. The key seems to have been the integration of the Unit into existing schemes of work. The most notable instances of integration occurred when pressure from 'above' (i.e., from the research team with the agreement of the Head of the school) was combined with active participation from 'below' (i.e., from the teachers themselves) and with executive leadership from the 'middle' (e.g., subject leaders' enthusiastic endorsement). Finding ways to support teachers from different schools to share online how they might achieve this integration will be a key design challenge in the next phase of the scaling process.

# References


Baker-Doyle, K.J. (2011). *The networked teacher: How new teachers build social networks for professional support*. New York, NY: Teachers College Press.

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497–511.

Confrey, J., Hoyles, C., Jones, K., Kahn, K., Maloney A., Nguyen, K., & Pratt D. (2009). Designing software for mathematical engagement through modelling. In C. Hoyles & J.-B. Lagrange. (Eds.) *Digital technologies and mathematics teaching and learning: Rethinking the terrain* (pp. 19–46). New York, NY: Springer.

Drijvers, P., & Trouche, L. (2008). From artefacts to instruments: A theoretical framework behind the orchestra metaphor. In G.W. Blume & M.K. Heid (Eds.), *Research on technology and the teaching and learning of mathematics: Vol. 2. Cases and perspectives* (pp. 363–392). Charlotte, NC: Information Age.

Drijvers, P., Kieran, C., & Mariotti, M.A. (2010). Integrating technology into mathematics education: Theoretical perspectives. In C. Hoyles, & J.-B. Lagrange (Eds.), *Mathematics education and technology – Rethinking the terrain* (pp. 89–132). New York, NY/Berlin, Germany: Springer.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., . . . et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington DC: National Center for Educational Evaluation

Fletcher, J., & Tobias, S. (2005). The multimedia principle. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117–133). New York, NY: Cambridge University Press.

Hammerness, K. (2006). *Seeing through teachers' eyes: Professional ideals and classroom practices*. New York, NY: Teachers College Press.

Heid, M. K., & Blume, G. W. (2008). Algebra and function development. In M. K. Heid & G. W. Blume (Eds.), *Research on technology and the teaching and learning of mathematics: Research syntheses* (Vol. 1, pp. 55–108). Charlotte, NC: Information Age.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., et al. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511

Hoyles. C., & Lagrange, J-B. (2009). *Mathematics education and technology-Rethinking the terrain*. New York, NY: Springer

Hoyles, C., & Noss, R. (2003). What can digital technologies take from and bring to research in mathematics education? In A.J. Bishop, M.A. Clements, C. Keitel, J. Kilpatrick, & F.K.S. Leung (Eds.), *Second international handbook of research in mathematics education* (pp. 323–349) Dordrecht, the Netherlands: Kluwer.

Kaput, J. (1992). Technology and mathematics education. In D. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (pp. 515–556). New York, NY: Simon and Schuster Macmillan.

Kaput, J., Hegedus, S., & Lesh, R. (2007). Technology becoming infrastructural in mathematics education. In R. Lesh, E. Hamilton, & J. Kaput (Eds.), *Foundations for the future in mathematics education* (pp. 173–192). Mahwah, NJ: Lawrence Erlbaum Associates.

Kaput, J., & Roschelle, J. (1998). The mathematics of change and variation from a millennial perspective: New content, new context. In C. Hoyles, C. Morgan, & G. Woodhouse (Eds.), *Rethinking the mathematics curriculum*. London, England: Falmer Press.

Kim, H. J. (2012). *An exploratory study of teacher's use of mathematical knowledge for teaching to support mathematical argumentation in middle-grades classrooms*. Unpublished doctoral dissertation. University of Texas at Austin.

Kynigos, C., Clayson, E., & Yainnoutso, N. (2012). *Constructionism 2012, theory, practice and impact: Conference proceedings.* Athens, Greece: Educational Technology Lab

Laborde, C. (1995). Designing tasks for learning geometry in a computer-based environment. In L. Burton & B. Jaworski (Eds.), Technology in Mathematics Teaching - a bridge between teaching and learning (pp.35-68). London: Chartwell-Bratt.

Mishra & Koehler (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings: Learning cultures and computers*. Dordrecht, the Netherlands: Kluwer Academic.

Noss. R., & Hoyles C. (submitted 2013). Constructionism & microworlds. In R. Duval, R. Sutherland, & M. Sharples (Eds.) *The Technology-Enhanced Learning Reader*. New York, NY: Springer.

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. London, England: Harvester Press.

Penuel, W.R., Fishman, B.J., Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Second Ed). Newbury Park, CA: Sage Publications.

Roschelle, J., & Shechtman, N. (2013). SimCalc at scale: Three studies examine the integration of technology, curriculum and professional development for advancing middle school mathematics. In J. Roschelle & S. Hegedus (Eds.), *The SimCalc vision and contributions: Democratizing access to important mathematics* (pp. 125–44). Berlin, Germany: Springer.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S. . . . Gallagher, L. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics. *American Educational Research Journal 47*(4), 833–878.

Roschelle, J., Tatar, D., Shechtman, N., & Knudsen, J. (2008). The role of scaling up research in designing for and evaluating robustness *Educational Studies in Mathematics*, *68*(2), 149–170.

Sinclair, N., Arzarello, F., Gaisman, M., Lozano, M., Dagiene, M., Behrooz, E., & Jackiw, N. (2010). In C. Hoyles & J-B. Lagrange (Eds.) *Mathematics education and technology-Rethinking the terrain* (pp. 61–78). New York, NY: Springer.

Sturman, L., & Cooper, L., (2012). Evaluation of Cornerstone Mathematics pilot in England: Unit 1, linear functions. Slough, England: National Foundation for Educational Research.

Vahey, P., Knudsen, J., Rafanan, K., & Lara-Meloy, T. (2013). Curricular activity systems supporting the use of dynamic representations to foster students' deep understanding of mathematics. In C. Mouza & N. Lavigne (Eds.). *Emerging technologies for the classroom: A learning sciences perspective* (pp. 15–30). New York, NY: Springer.

Vahey, P., Roy, G., & Fueyo, V. (2013). Sustainable use of dynamic representational environments: Toward a district-wide adoption of SimCalc-based materials. In S. Hegedus & J.

Roschelle (Eds.). *The SimCalc visions and contributions: Democratizing access to important mathematics* (pp. 183-202). New York, NY: Springer.

# Cornerstone Mathematics: Designing Digital Technology for Teacher Adaptation and Scaling[1]

C. Hoyles*, R. Noss*, P. Vahey**, J. Roschelle**

**London Knowledge Lab (LKL), Institute of Education, University of London, UK*

*** SRI International, Menlo Park, CA, USA*

Email: c.hoyles@ioe.ac.uk

Phone: +44 (0) 20 7911 5577

Fax: +44 (0) 20 7763 2138

[1] We gratefully acknowledge funding by the Li Ka Shing Foundation. The research has entailed collaboration among, in LKL, Philip Kent; and in SRI International, Jennifer Knudsen, Ken Rafanan, Teresa Lara-Meloy, Anna Werner, Gucci Estrella, and Nicole Shechtman.

Abstract

We report the results of a design-based research project in England that embeds digital technology. The research followed from two phases in the United States: (1) a design phase that used dynamic representations to foster conceptual understanding of hard-to-teach mathematical ideas, and (2) a research phase that measured the efficacy of the resulting technology-based curriculum units as implemented in Texas schools. The goal of the third phase in England was initially to “scale up” the U.S. approach. We determined, however, that the materials had to be re-designed for adaptability by English teachers. We report how the features of the innovation—particularly its technological infrastructure—could be leveraged, not only to achieve positive learning outcomes, but also to lay the foundations for change in pedagogy and learning at scale. We identify an emergent framework of design affordances for teacher adaptability that are particularly salient when technology is a critical element.

# 0 Introduction

The potential for technologies to transform mathematics education is well-established. However, the process of getting to scale from initial proof-of-concept research has been problematic (Hoyles & Noss, 2003; Roschelle, Tatar & Kaput, 2008). Kaput, (1992) attributes the continuing marginalisation of technology in mathematics education to the complex issues that surround its use:

- Technology requires continually rethinking pedagogical and curricular motives and contexts.
- Classroom-based research is difficult, because exploiting the real power of the technology requires such innovative approaches that comparison to a traditional class is inappropriate.
- The practical complications of student access to computers, cost of software, and development of curricular materials often prohibit research.
- Given rapid changes in technology, research is often out-dated by the time it is complete.

Despite substantial developments in theory, and massive changes in technology, the core challenge of ‘implementation’ remains: how to ensure that digital technology is used at all in mathematics classrooms, and, if it is used, how to enhance mathematical thinking rather than simply reiterating current practice or, as is sometimes the case, circumventing mathematics altogether. How can research inform efforts to embed technology in transformational ways that support scalable models of classroom change?

This paper reports on research undertaken as part of the Cornerstone Mathematics (CM) project, a collaboration between researchers in the U.S. and England for a moderately large-scale, design-based implementation project in England. U.S. researchers completed two phases of work before CM came to England: (1) design of the core use of dynamic representations in ‘modules’ that include curriculum workbooks and teacher professional development; and (2) efficacy trials that established causality linking the modules to improved student conceptual understanding of challenging mathematics. The goal of the project in England was eventually to reach 100 or more schools using the US-developed materials and approach. However, the team determined that before going directly to scale, a further phase focused on designing materials and processes to support teacher adaptation and instrumentalisation was needed: we elaborate below.

As well as presenting results from research focused on these design concerns, we also present hypotheses about the key features required for scaling a technologically-based innovation in the field of mathematical learning.

# 1 Theoretical background

Sinclair *et al.'s* (2010) model of the development of digital technologies in mathematics education describes a shift in research attention from "Wave 1" with an exclusive focus on the relationship between individual learners and mathematics, to "Wave 2," which involves the broader context of learning, the teacher, and the curriculum. The authors argue that this trajectory is almost inevitable. Any project that seeks to scale must, in addition to assessing individual students' interactions with technology and teacher/peers, address the curriculum and the teacher's role in using and deploying the digital technologies. Hoyles & Lagrange (2009) provide an overview of these trends.

Two challenges for the current research were to identify existing or evolving theories that (1) provided an adequate underpinning for such a complex, intertwined set of components as technology, teachers and students, curricular context, materials, and (2) supported the generalisability of research findings. Drijvers, *et al.*, (2010), who provide an overview of the theoretical frames used in technology-related research in mathematics education, argue that integrative frameworks allow for articulating different theoretical perspectives. This paper represents a step towards developing such a framework. It takes as its starting point the curricular activity system framework described in Roschelle *et al.* (2010) and Vahey, *et al.* (2013a). This framework is based on the recognition that an instructional activity includes a learning objective, available materials, the intended use of tools, and the roles of diverse participants. However, focus on such activities is not sufficient. An activity must fit into the structure of classrooms that are expected to engage in the activity. The curricular activity systems approach integrates learning requirements, teacher professional development (PD), curriculum materials, and technology; the approach recognises that these elements are situated in an educational context that includes people, conventions, and policy considerations.

We also look to design-based implementation research (Penuel *et al.*, 2011), which expands on traditional design of an artefact or intervention by developing the capacity for sustained change. Central to our scaling approach is building teachers' capacity as they both strive to adopt the CM materials, and participate in *adapting* the curriculum materials and *instrumentalising* the technology. *Instrumentalisation* is part of the process of *instrumental genesis*, whereby artefacts are transformed (by both students and teachers), into 'instruments', systems with which the user gains fluency (becomes *instrumented*) and acquires expressive competence, (is *instrumentalised*) (see Artigue, 2002; Drijvers & Trouche, 2008; and, in the context of SimCalc research, Roschelle *et al.*, 2008). So by instrumentalisation we mean the process by which students and teachers come to use the potential of the technology (or other digital artefact) for their own purposes, transforming it as they do so.

A complementary strand of research concerns constructionist design, whereby users learn through interaction with and feedback from digital tools that enable them to explore, build, and learn (Papert 1980). Constructionism has been the subject of extensive research and development and continues to result in innovative ways of designing tools and in work with learners worldwide (e.g., Kynigos *et al.*, 2012, Noss and Hoyles, submitted: 2013). Digital technologies not only add new representations (or link old ones), but research has increasingly shown that digital representations change the epistemological map of what it is intended for teaching and learning (Noss and Hoyles, 1996; Kaput and Roschelle, 1998).

By building teachers' capacity to use materials effectively and still maintain the core pedagogical approach, we optimise our chances of scaling up to large numbers of teachers, classrooms, and students. Related approaches to scaling emphasize the importance of "fit" to broader reforms in a school system (Blumenfeld *et al.*, 2000); that scale involves a shift to increased ownership by teachers (Coburn, 2003); and that scale involves new kinds of empirical evidence (Schneider & McDonald, 2007). (Additional approaching to scaling could include working with a publisher or technology company to incorporate research-based insights into established products (Roschelle & Jackiw, 2000).

We first provide a brief outline of the U.S. studies that formed the basis for the research in England.

## 2 Preliminary Phases in the United States: the SimCalc Intervention

Over the course of about 6 years, the U.S. team developed integrated technology, curriculum, and professional designs to address big ideas in the algebra and pre-calculus strand of mathematics. The *technology design* emphasized linked dynamic representations such as graphs to motions, tables, and equations (Kaput *et al.* 2007; Vahey, *et al.* 2013a; see Figure 1). Many evaluations show, at best, small positive effects for technology (Cheung & Slavin, 2013; Dynarski *et al.*, 2007). However, dynamic representations have been shown to be powerful in developing conceptual understanding (Heid & Blume 2008). These representations can help students make important connections between intuitive and formal ideas, and between graphical and linguistic understandings. The *curriculum modules design* is based on organising key mathematical ideas and practices into integrated paper- and technology-based materials that are straightforward to use. Within the modules, students solve increasingly challenging problems in one content area organised around a high-interest theme such as ‘Designing Mobile Games’. Teachers focus on the key mathematical ideas, valuing conceptual understanding and guiding development of mathematical practices. The *professional development design* emphasises practical ‘mathematical knowledge for teaching’ (Hill *et al.,* 2008)—the knowledge teachers need to make sense of and extend their students’ mathematical reasoning.

This integrated design of technology, curriculum, and PD was formalized as an instructional package that could be implemented at scale. Subsequently, a randomized controlled trial (RCT) in Texas investigated whether a wide variety of teachers, when provided with appropriate PD and a replacement unit that integrated curriculum and software, (the “SimCalc intervention”), could increase student learning of important and complex mathematics. In this RCT, teachers were randomly assigned to participate in the treatment or the control group. The treatment group received the integrated intervention, which began with a 3-day teacher PD workshop in which teachers learned to teach using the technology-based unit. Treatment teachers were then asked to teach the replacement unit in place of their usual unit on linear function. Most teachers completed the unit in 2-3 weeks. Teachers in the control group received PD on the integration of technology into mathematics teaching, but the PD addressed different content.

Working with five Texas Education Service Centers (the primary support and PD providers in Texas), teacher volunteers were recruited whose students reflected the regional, ethnic, and socioeconomic diversity of the state. Complete data were returned by 56 teachers and 825 students. At intake, the Treatment group (33 teachers) and the Control group (23 teachers) did not differ in any important way (e.g., with respect to teaching experience, ethnicity, gender, mathematical content knowledge, or by socioeconomic status as indicated by percent of students eligible for free or reduced-price lunch in school). The greater number of teachers in the Treatment group was an artefact of teachers’ scheduling conflicts with the workshops to which they were assigned. Because teachers were not informed about the workshop type until the workshop occurred, the consequences for randomisation and thus the validity of the experiment were minimal. The attrition rate was comparable to other large experiments with educational technology (Dynarsky e*t al.*, 2007).

The primary outcome measure in this study was student learning of core mathematical content. The team employed a bespoke assessment instrument. Working with a panel of mathematicians and mathematics education experts, the assessment encompassed both simple (M1) and more complex (M2) aspects of linear function. The simpler items (M1) were based on those used in existing standardised tests in Texas; for example, students were asked to calculate using a linear relationship represented in different ways. More complex M2 items required comparing multiple rates or finding average rates, and typically the rate information was not provided directly; for instance, students had to infer it from slopes in graphs. The assessment was administered in a single class period to students immediately before and after their linear function unit was taught.

The analysis of student gain scores from pre-test to post-test showed a large and significant main effect, with an effect size of 0.56 (Roschelle *et al.*, 2010). This effect was robust across a diverse set of student demographics. Students who used the SimCalc materials outperformed students in the control condition regardless of gender, ethnicity, teacher-rated prior achievement, and socioeconomic status. This finding provides evidence that the use of dynamic representations, when embedded in a set of replacement units designed within a curriculum activity system framework, can result in substantial learning gains.

# 3 The Cornerstone Mathematics Project

CM seeks, as did its U.S. predecessor, to exploit the dynamic and multi-representational potential of digital technology to enhance learners' engagement and understanding of mathematical ideas. CM consists of four units, each focused on key mathematical topics in middle school (students aged 11-14 years). Each unit embeds activities in a quasi-realistic digital context in which students *need* to use mathematical knowledge to achieve their—and our—goals. This is an important design decision: students need to be motivated to think mathematically and this can be achieved by tapping into their digital lives, making the work "realistic" (e.g., Confrey *et al.,* 2009).

Unit 1, the focus of the research reported here, concerns piecewise-linear functions, with the work in England replicating the SimCalc intervention (see Roschelle *et al.*, 2010; Roschelle & Shechtman, 2013). However, for success at scale in England, two levels of changes were required. For the initial, and simpler, level, the mathematical content and language were adjusted to fit the English setting. For example, numerous changes in vocabulary had to be made to account for differences between U.S. and UK English usage and spelling. Changes were also made to conform to local school mathematical conventions (e.g., $y$=m$x$+c rather than $y$=m$x$+b). Third, we sought to align the unit to the English National Curriculum, which is statutory and places comparatively greater emphasis on 'mathematical processes'.

This paper focuses on more complex changes. In an RCT, experimenters try to hold the 'inputs' fixed and attempt to have teachers adopt the materials with high fidelity. However, in scaling up, innovators recognise that teachers and schools are not identical, and the metaphor of teacher adoption will not suffice. On the contrary, it is inevitable that each new classroom, teacher, and school context introduces new features that could knock the innovation off course. A key challenge, therefore, was to move beyond adoption and design for *robustness* (Roschelle *et al.*, 2008). We did so by encouraging teachers to appreciate the goals of the innovation, and to make the innovation their own through a process of adaptation of the materials and instrumentalisation of the technology. Thus the more complex change, and the focus of this paper, involved designing the materials to better support this intended process.

We wanted all CM teachers to develop pedagogical and technological confidence through PD and supporting their teaching of the unit. We also wanted teachers to work together in a mutually supportive community through a blend of face-to-face and virtual interaction: the latter becoming increasingly recognised as important for scaling, especially when teachers are widely distributed (e.g., Baker-Doyle, 2011). To check that the ways we supported teacher adaptation and instrumentalisation of CM Unit 1 did not undercut the effectiveness of the materials (as established earlier), we conducted design research to evaluate the impact of the redesigned unit, to assess the extent to which student learning gains were preserved.

### 3.1 From the SimCalc Intervention in the United States to CM Design Research in England

CM Unit 1 started from the U.S.-based unit called *Designing Cell Phone Games* as described earlier (see Roschelle *et al*., 2010; Vahey, *et al*., 2013b). The Unit highlighted the following mathematical concepts: coordinating algebraic, graphical, and tabular representations of linear functions; $y$=m$x$+c as a model of constant velocity motion; the meaning of m and c in the motion context; and velocity as speed with direction. All activities were set in the context of using mathematics to design computer games for mobile phones, where functions must be used to make game characters move in appropriate ways. At the heart of the software environment is a simulation, or a ‘journey’, of an object that could be tracked in a graph and a table, as well as captured in algebraic or narrative form (see Figure 1). Students therefore receive feedback on any journey they have constructed by visually ‘seeing it happen’; the mathematics plays out in terms of motion and vice versa. Students can control their object’s journey by manipulating the position-time graph or its algebraic representation. The constructionist key in the Unit is that students themselves can intervene in the ‘system’ by constructing journeys and exploring them alone and with others.

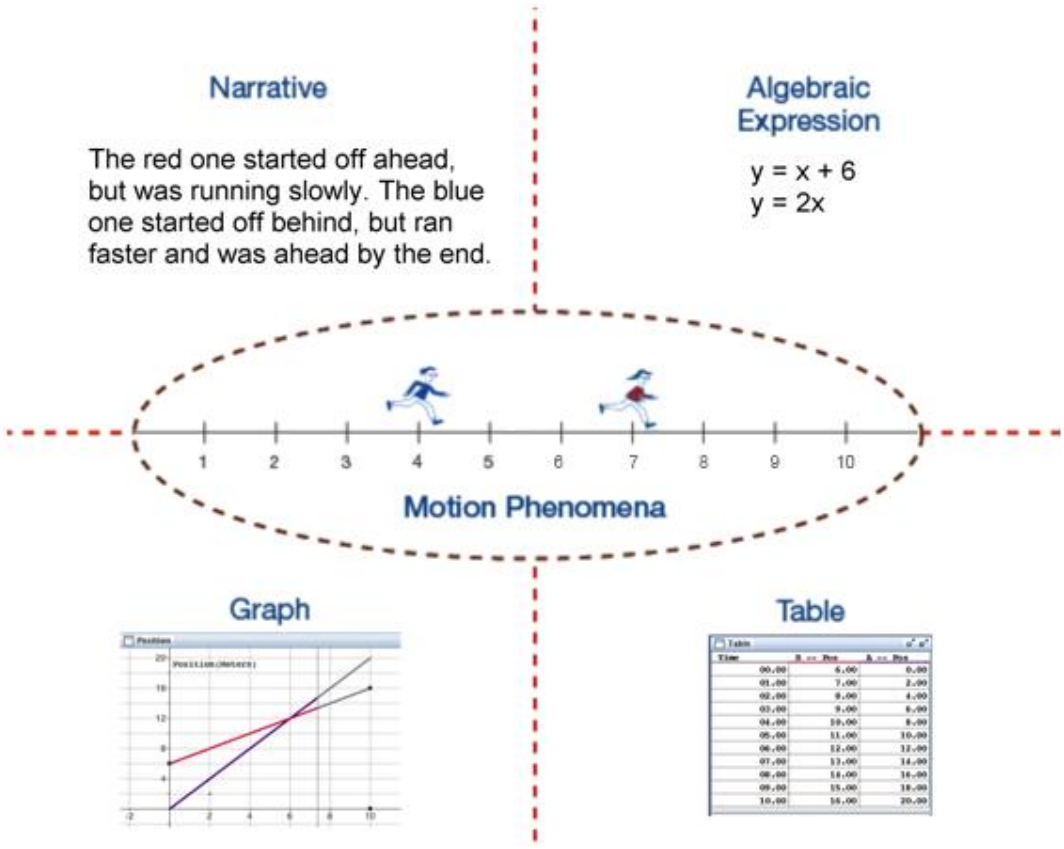


Figure 1: The SimCalc system links algebraic expressions, graphs, tables, and narratives through the phenomenon of motion

A first priority for the design research in England was revisiting the style and emphasis of the PD needed, not least because middle-school students in England are taught by secondary trained mathematics teachers, in contrast to the U.S. teachers who were mainly primary (elementary) trained. Teachers in England, and particularly subject leaders, expect a degree of autonomy in approaching given mathematical topics.

As a second priority, teacher support in England required a different character. In the U.S. efficacy trial, the researchers wanted to control precisely the PD, and thus they and PD experts provided it. However, in England, teacher adaptation and instrumentalisation were emphasised. For example, at the request of the teachers, more group-work was introduced in the PD and teachers orchestrated this themselves as a means to share experience and reflections. In addition, PD was extended through orchestrated peer interaction among the CM teachers in an online community, which was part of the technical national infrastructure provided by the National Centre for Excellence in the Teaching of Mathematics, a nationwide professional learning community for mathematics teachers (www.ncetm.org.uk). Through the CM online community, teachers could share their experiences, the prompts and probes they had planned or used spontaneously, and the activities they designed themselves.

### 3.2 Research Questions

Our study did not seek to establish the efficacy of the intervention as compared to some other mathematics program. Instead, our initial research question focused on investigating if the materials, as revised for England, maintained their effectiveness in terms of students learning complex math. Hence, the research question was:

> RQ1: Is the pattern of student learning gains in England consistent with the pattern of learning gains observed in the United States?

A consistent pattern would reassure the team that the changes to the design for the England scale-up preserved the effectiveness of the materials and approach, whereas if we found that there was significantly less learning in England we would call into question either the suitability of the materials or the design decisions made in our revisions. To determine if the pattern of learning gains was similar, the assessment and assessment method used in the Texas RCT was closely followed to allow a quasi-experimental comparison with pupils in that study. Pre- and post-tests for each student generated gain scores that were used for a quantitative analysis of learning gains, which were then compared with the Texas student results.

Even if the learning gains were consistent, our goal of scale-up would only be met if teachers in England considered the materials usable and useful. To provide insight into teacher perceptions of the materials, an external evaluation was undertaken, focusing on the overarching question of:

> RQ2: What were teachers' perceptions of the materials, and in particular did they find that the materials were manageable to implement, addressed important mathematical topics, and were effective for a variety of pupils?

To provide insight into teacher adaptation and ownership of the materials, we investigated three qualitative research questions:

> RQ3: What changes supported teacher adaptation?
>
> RQ4: How did teachers come to greater ownership of the materials and technology?
>
> RQ5: As the approach started to scale up, were teachers engaged in a community around the materials and approach?

### 3.3 Sample

The initial sample for CM research consisted of 10 schools, with 2 teachers recruited from each school. One school was unable to complete its participation, however, because of a fire in the school buildings, and one teacher took maternity leave. The final sample comprised 9 schools, 17 teachers and classes, and 429 students drawn from Years 7 through 9 (students aged 11-14). We were careful to ensure that the schools and the classes exhibited a wide diversity of school contexts and prior student achievement. Although we can hardly claim that these schools were representative of all schools in England, the sample schools had a range of: academic strength (as measured by public examination results), student intake in terms of socioeconomic status, technology infrastructure, and teachers' experience of mathematics teaching. The school sample included one private school with a privileged intake, and one state-run school whose students were economically disadvantaged and in which English was an 'additional language' for a high proportion of students. The teachers, too, were diverse in their teaching experience, including one new to the profession, most with 3-5 years of experience, and some highly experienced mathematics subject leaders in their school. Some of the teachers had excellent mathematics qualifications and some less so. The 429 students who engaged with Unit 1 were distributed across different year groups (according to teacher choice): Year 7: 179 (42%), Year 8: 227 (53%), Year 9: 23 (5%).

The intervention consisted of 2 days of PD for all teachers, 0.5 day for a reflection/feedback session after implementation, and ongoing discussion through the online forum. The unit was taught over 2 to 3 weeks (spread over a month for the different schools) with students using computers in most lessons (a practice that was unusual in mathematics class). Hardware varied across the sample, ranging from desktops in computer labs, through laptops in classrooms, to the extensive use of the interactive whiteboard.

### 3.4 Data Sources

The sources of data on the processes and outcomes of the study were collected by the following means:

a) A *proforma* that a researcher completed in discussion with each school's main contact, generally before a school visit. The proforma was designed to establish a baseline 'context for teaching', indicating the reasons why staff in these schools had chosen to participate in the study, background information about mathematics teaching in the school, expectations for the study, and any challenges anticipated.

b) *Lesson observations* in one or two participating classes in each of the 9 schools. We observed 16 lessons across 9 schools. During each observation, we took notes using a semi-structured protocol around the following main themes:

- Manageability: in particular how teachers orchestrated the lesson content and pace
- Engagement/behaviour (attitude, engagement and understanding, behaviour and interaction)
- Teaching style/approach (interactions and interventions).

c) *Post-observation interviews with the teacher(s)* of the observed class(es) in 9 schools. We interviewed 16 teachers in the 9 schools. The interviews served to collect relevant information in the following categories: the teachers and students observed; preparation for teaching the unit; the implementation of the unit; the use of the technology; and the impact of the unit on students in terms of their engagement and learning.

d) *Students' questionnaire responses* completed at or towards the end of the unit (eight schools completed and returned the responses). Students in the observed classes completed a questionnaire about their views of mathematics lessons generally and about CM Unit 1 in particular.

**e**) *Focus-group discussion* with a sample of students from nine schools. Teachers selected 68 students for each focus group and were advised to provide a mix of students (e.g., of different ability levels or different language backgrounds).

f) *Pre-test and post-test data* Students completed two identical tests to evaluate their learning: a pre-test, before the unit was taught, and a post-test, taken at the end of the unit. These tests had been used in the Texas study.

Interview and observation schedules were consistently followed, and the interviews and focus-group discussions were audio-transcribed.

# 4 Results

We begin with our quantitative results in which the analyses followed that undertaken in the U.S. study. We then summarise results from the external evaluator, and next present the results of the qualitative analysis that was guided by the research questions listed previously.

## 4.1 Quantitative Results: Consistent Pattern of Learning Outcomes

We replicated the effect found in the earlier Texas results by (a) establishing that the English classrooms performed equivalently to the Texas students on the pre-test and (b) that the English classrooms gained at least as much as the Texas treatment classroom did on the post-test. Although conducting a randomized controlled trial in England with a control group would have been a stronger test of replication, it was simply not feasible to do so. The main threat to validity of this one-condition study is the possibility that English classrooms learn difficult concepts without new materials, whereas Texas classrooms do not. If this were so, the gains documented in England could be unrelated to Cornerstone materials. Based on our extensive knowledge of challenges in English mathematics education, it seems unlikely, that English students would find it easier to learn these concepts, and thus a one condition trial is a reasonable approach to replicating and extending the prior result. The pre- and post-tests were scored in London by a group of 12 pre-service mathematics teachers recruited from the Institute of Education. Using a web-based data entry and verification form, scorers entered their marks and comments online; this procedure allowed instant access to the data while ensuring rigorous methods consistent with prior methods used in the United States. All scorers were trained on the scoring key and rubrics, and were permitted to begin scoring actual assessments only after they had attained a minimal level of gold-standard marking of sample assessments. To measure inter-rater reliability during the scoring process, 10% of all assessments were scored by a second scorer. The inter-rater reliability was found to exceed 90%, well above the acceptable level.

Using the pre-test as a proxy for equivalence of prior background knowledge, we found that the groups in the two studies were similar. The researchers fitted a series of hierarchical linear models (HLM: Raudenbush & Bryk, 2002) to examine statistical equivalence at pre-test. These analyses showed no statistically significant differences for the CM and US samples for the total score [$\beta = -0.64$, $p = 0.69$], M1 subscale [$\beta = -0.50$, $p = 0.52$], or M2 subscale [$\beta = -0.15$, $p = 0.86$]. The variation among pupil pre-test scores in both groups was also sufficient to rule out floor effects as a possible explanation for the apparent equivalence of prior knowledge. In a non-significant trend, CM pupils had slightly lower pre-test scores, which may be attributable to age differences. The determination that the groups were equivalent at pre-test is important because in a quasi-experimental comparison such as this, the primary threat to internal validity is the possibility of non-equivalence of groups at baseline.

Figure 2 shows the learning gains for pupils in the Texas Control group (who did not use the materials), the Texas Intervention group (who did use the materials), and the CM pupils in England. The dark part of the bars shows learning for M1, ‘simple’ linear functions, and the light part of the bars shows learning for M2, ‘complex’ linear functions. The learning gains for CM pupils were similar to the learning gains for pupils in the Texas Intervention group, and were significantly higher than for pupils in the Texas Control group.

These data show that the CM approach met the goal of increasing pupil learning of important mathematics in England. Further analyses indicated that the materials were equally effective for students with different levels of prior mathematics achievement: schools with higher prior General Certificate of Secondary Education (GCSE) scores performed higher on pre-test, which was to be expected; however, *learning gains* were not correlated with school achievement level, indicating that the materials were effective for pupils from a variety of school contexts. (GCSE is an academic qualification awarded in a specific subject, usually taken in a number of subjects by students aged around 15-16 years).

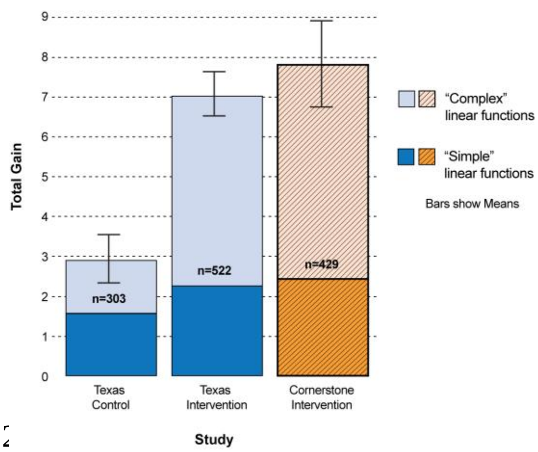

Figure 2. Learning gains of pupils in the Texas and England studies showed similar learning for both sets of students.

These analyses provided important evidence about the effectiveness of the materials. Although we would not advise drawing any conclusions from cross-country comparisons (i.e., success in the U.S. vs. success in England), comparison between the magnitude of gains among the U.S. control, U.S. treatment, and CM groups provides evidence for feasibility of effectiveness in the English context.

It is noteworthy that although gains were similar in ‘simple’ linear functions for all groups (although the gains for the control were slightly lower than those for the other groups, that difference was not statistically significant), the difference in groups was predominantly for complex concepts.

Figure 3 shows an example of a complex M2 item, and a sample student response in the pre-test. The student had been provided with the questions and gridded graph areas, and the student drew the lines on the graphs. We note, in this one pupil’s response, several well-known conceptual difficulties, including treating a position-time graph as a velocity graph, and representing backwards motion as a line that goes back toward zero on the time axis.

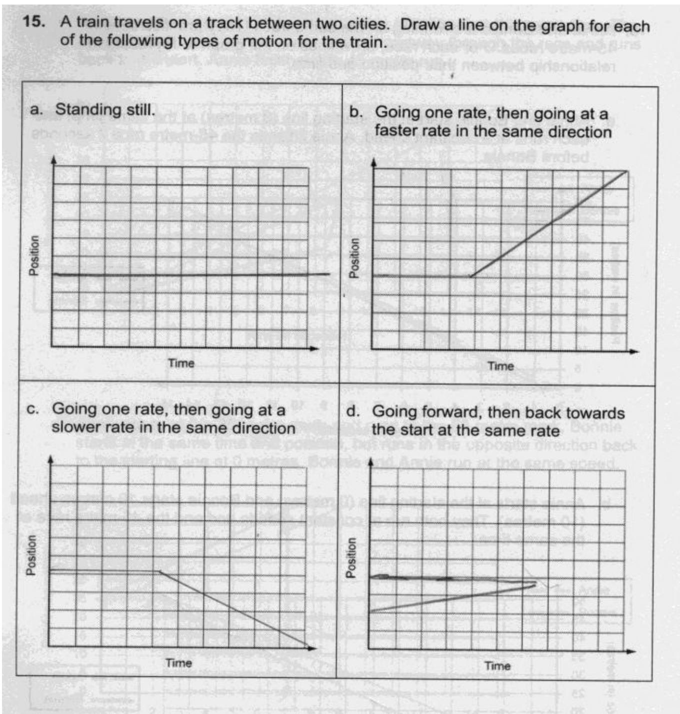


Figure 3: Example of a complex linear function item, Question 15, and a pre-test student response.

Table 1 shows that the changes in student scores for this question are substantial, and compare well with the (older) control group students. In the next section, we track how these changes may have been catalysed by the design and structure of the CM unit, and its deployment.

| % Correct | Q15a pre | Q15a post | Q15b pre | Q15b post | Q15c pre | Q15c post | Q15d pre | Q15d post |
|---|---|---|---|---|---|---|---|---|
| CM Year 7 | 21.2 | 56.4 | 11.7 | 51.4 | 9.5 | 46.4 | 7.3 | 41.9 |
| CM Year 8 | 45.4 | 77.1 | 28.2 | 73.6 | 20.3 | 71.8 | 14.5 | 63.9 |
| Texas Control | 40.3 | 52.8 | 30.7 | 44.6 | 24.8 | 35 | 21.1 | 30.4 |

Table 1: Scores on Question 15 showed increased learning for CM Pupils between the pre- and post-test.

### 4.2 Qualitative Results: External Evaluation

To assess the validity and reliability of the qualitative research, an external evaluation of the implementation of Unit 1 was undertaken (Sturman & Cooper, 2012). The main findings of this evaluation were, in summary:

1. The unit was generally manageable to implement and helped students to learn the difficult concepts covered.
2. The unit had a good curriculum fit and covered more than most teachers' own schemes of work. When teachers could make a prediction, they felt that students would do well in their next assessment, based on their solid understanding and willingness to attempt questions.
3. The unit was generally seen as useful, with most of its parts more likely to be rated 'helpful' than 'unhelpful'. Some technical difficulties were encountered, and some changes and additions to resources were suggested.
4. Teachers were positive about the unit's ability to engage students, although students themselves expressed some reservations. Despite this, students acknowledged that the unit helped them to learn.
5. Perceptions of the groups that the unit affected most were varied, but test results suggest that the unit was appropriate and effective for the whole range of students, with no group conclusively progressing better than others.
6. Teachers were generally positive about the impact of the unit on students' understanding of mathematics in the real world, although they identified different ways of how the unit achieved this. Students also perceived real-life benefits.

These general findings provide a backdrop for the more detailed analyses we present below.

### 4.3 Qualitative Results: Teacher Adaptation and Instrumentalisation

The themes around which the observation, interview, and student focus group data were analysed included the extent of teachers' preparation for teaching the unit; the implementation of the unit; technology instrumentation; and the impact on students in terms of their engagement and learning. To identify relevant episodes, data were provisionally coded, assigned to one of these themes. and sub-divided into the appropriate categories .

In a major adaptation of the CM unit 15 out of 18 of the teachers modified each lesson to conform to the common ‘3-part lesson’ format (starter, student activity, plenary consolidation), a format that the UK government’s “National Strategy” for education had effectively mandated during the years 2003-11. Although this adaptation was common, the specific adaptations took different forms: some teachers devised new starter activities, while others creatively varied the kinds of orchestration they adopted (e.g., student groups sharing their work on the interactive whiteboard).

Teachers also adapted the pace of implementation. Teachers had to adapt learning activities to a variety of student groups (year, set, level of ‘English as Additional Language’) and lesson formats (single or double periods, a PC room, or laptops in the classroom), and they broadly managed to do so. We found, however, that teachers’ planning of timing was not precise (all teachers were given time for planning, and all created an initial plan of around 8-10 hours of lesson time, which in practice ranged between 8 and 20 hours). Such inconsistencies may be inevitable for novel curriculum activities, given that teachers can learn how to use the unit only by teaching it. It is hardly surprising that a disruptive technology would lead to unpredictable timing of lessons. But it is surprising that the teachers were so positive about this disruption and the extent to which it encouraged—perhaps forced—new thinking about pedagogic strategies. As part of the process of instrumentalisation, they chose to elaborate on what the actions on the software meant mathematically (couched in the language of the unit), with some adding new activities for groups to work on together. We believe these adaptations were crucial to the success of the unit—the students were not simply drawing the graphs or changing the slope, but were engaging with digital mathematical objects as they thought deeply about the mathematical concepts.

In summary, all 17 teachers adopted the CM activity sequence, with 13 changing it by some subset of:

- Compressing/stretching activities, according to the teacher’s expectations of how students would react, with adjustments in the course of the lesson
- Switching on an *ad hoc* basis between pair and whole-class work if interesting ideas or points of explanation arose
- Introducing activities to ‘consolidate’ the mathematical ideas (e.g., one ‘matching’ activity in the activity booklet was represented as a card-sort activity for a starter in the next lesson)
- Adding short activities at the ends of lessons that could be used for consolidation in class or as homework.

Some teachers were more open to adaptation than were others. For some teachers with students with English as an Additional Language, or students needing more exercise in numeracy/arithmetic, the activities opened up extension opportunities to work on developing students' capacities with mathematical language/argument, or numeracy skills, which was not part of the core design. In these ways, the extended time spent on the unit was judged as highly valuable.

### 4.4 Qualitative Results: Teacher Ownership and Instrumentalisation

Teachers in England have, over the last decade, been encouraged through Government strategy and inspection to adopt a surprising level of uniformity in their pedagogic strategies. Doing so has militated against innovation. What became evident in this study was the extent to which the technology disrupted this conformity, and in doing so, led the teachers to adopt strategies that were new to them, and which they recognised as being successful in evolving student learning. We identify some examples below.

One example was in lesson style. Some students reported a change in lesson style, which was also evident from the observations and teacher interviews. The students, in particular, noted that typical mathematics lessons were usually dominated by 'copying from the board'. Teachers tended to recognise that students arrived at a deeper understanding mathematics when they allowed them to interact with the technology, and discuss the meanings they developed. The curriculum replacement approach provided a structured and relatively short-term (and therefore safe) setting for doing more with technology, particularly moving away from whole-class demonstrations (by the teacher, or by selected students) via an interactive whiteboard.

Another example of teacher ownership can be observed in the ways in which technology was deployed. All of the teachers used digital technology in almost every lesson, thus overcoming any initial reservations. Indeed, many of the teachers decided to re-use parts of the unit as revision material for older pupils. Thus, teachers considered the materials useful for meeting their own needs and the needs of their pupils beyond the use specified by the study.

More generally, 15 of the teachers used their professional judgement to repackage the material by choosing to teach disparate pieces together, or to decompose one idea into many. Although we recognise that these examples of teacher ownership fall short of the re-design of tasks themselves to exploit the use of the technology to give sense to mathematical concepts (as discussed by Laborde, 1995), we nonetheless interpret these changes as expressions of epistemological autonomy on the teachers' part.

More broadly, multiple sources of evidence indicate the ways in which the teachers instrumentalised the innovation considered as a whole. This finding underlines the limitation associated with thinking of an innovation as something to 'implement' or 'deploy'. Considerably more than that is required: the presence of the technology (and of carefully designed workbook-based activities) highlights how teachers implementing the innovation need time and support to make the innovation their own, to reshape it, and to use it to create novel strategies as well as new epistemologies for themselves and their students.

A key focus for teacher instrumentalisation was the use of multiple representations to reveal and address student misconceptions. In our interviews teachers consistently reported that that feedback for actions was a learning driver when successful learning occurred. Crucially, these teachers recognised the power of the simulation that drove the software. As one teacher put it:

> 'Having the things that move, the real world simulation, glued … [the other representations] together in a more meaningful way. Normally, [students] would be fine, they would be able to repeat to me that the gradient is the steepness of the line and the y intercept is where it crosses, but somehow having it linked with the simulation really brought that home in a way that I have not seen before'.

One activity concerned the notoriously hard concept of average velocity, and was framed in the context of Red Riding Hood (RRH) and the Wolf. The Wolf, which can move only at constant velocity, is trying to arrive at Grandma's house at the same time as RRH, whose speed can vary in a piecewise linear way. Students were asked to create a trip where RRH travels at two velocities and arrives at Grandma's house at exactly the same time as the Wolf. They could then check the graph and animation, and revise if it did not achieve their goal. They were asked to write a story describing RRH's trip.

One teacher reflected on her students' learning as follows:

> 'The best thing was [that the software] could visually contradict their misconceptions … the average speeds in Red Riding Hood [example], when the journey was not split into equal parts, it was really noticeable the students still wanted to add the speeds and divide by two. But as soon as they put that in as the wolf's motion, it was so clear that the wolf was not arriving where it needed to'.

As the quote indicates, the teacher recognised the power of the digital tool and the 'multiple mathematical representational' approach—written/verbal, graph, simulation—along with the coordination between the context world and mathematical representations in a predict/check/explain cycle that invited students to coordinate their thinking across representations. Students were offered ownership of the journey, as they could build their own narrative about the story as depicted in its mathematical representations. Through observations and interviews we found that teachers were quick to accept the responsibility of addressing student misconceptions that were revealed by this activity, and readily saw the utility of instrumentalising the affordances of the technology to aid them in addressing those misconceptions.

Quantitative evidence from pre- and post-tests corroborated this claim. Figure 4 shows a question related to average speed, and a sample student pre-test response.

Although this item was difficult for students, even at post-test, gains were large and significant: at pre-test 2% of pupils got part (a) correct, and 12% got part (b) correct; at post-test 28% and 47%, respectively, gave correct answers. Pupils' qualitative understanding of linear function also increased: at pre-test approximately 20% of pupils answered item 15 correctly (see Figure 3), whereas at post-test 63% answered correctly. Although we are not fully satisfied with the results, we are encouraged by the large number of Key Stage 3 students (i.e., Years 7 through 9 students) whose understanding of these complex and subtle linear function ideas increased, given that many students do not exhibit such understanding even at a much older age.

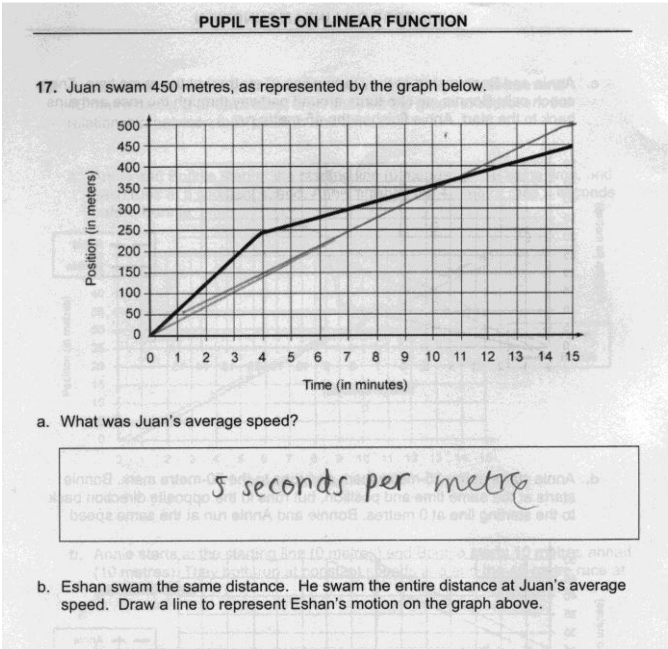


Figure 4: Example of an *average speed* item, and a sample pre-test response.

### 4.5 Additional Results: Sustainability and Community

We have reported here the CM design research for Unit 1. Work is ongoing on CM Unit 2, which concerns geometric similarity. Out of the 9 schools and 17 teachers in our sample, *all* have continued with CM and implemented the second unit of work. This provides evidence of considerable commitment, as teaching another CM unit was in no sense compulsory. As we move to the next phase of the project, 2013-15, again *all* the schools have requested to continue their participation. Additionally, two schools have reported taking one more significant step and have integrated Unit 1 into their *schemes of work*. Others are in the process of doing so. This is a major step towards sustainability as CM moves towards its next target of 100 schools. In England, the mathematics curriculum is institutionalised at the school level through schemes of work planned by the mathematics team, led by the subject leader. Subject leaders negotiate the school mathematics curriculum with their colleagues and instantiate mathematics schemes of work year by year. The schemes must, of course, be aligned with the National Curriculum and pupils progress through the scheme, usually assessed by the school annually, monitors progress. So when CM is part of the scheme, it will have become integrated into the routine of mathematics teaching in the school and taught by the whole mathematics department. In the two cases mentioned above, where Unit 1 has already been integrated, this was instigated by two CM participants who were subject leaders. In schools where the subject leaders were not CM participants, we expect the integration of CM into the schemes of work to take more time.

However, it proved challenging to spark productive conversations on the online community: the research team made repeated attempts throughout the implementation of Unit 1 to provoke discussion, but was only partially successful in doing so. Only 11 teachers posted; 45 topics were introduced with 104 posts in all, but most topics had only a handful of posts, which mainly consisted of answers to questions or factual statements, and little discussion took place. It seems that online professional discussion was too removed from teacher practice and judged as perhaps tangential to established professional practice.

In response, in the scaling work we are now undertaking, we are building and promoting the online community from the start, setting a norm in the PD sessions of using the online community as a place to engage in asynchronous discussions, and stressing the need to exploit the support that a vibrant community can provide.

## 5 Next Steps

This section elaborates a speculative framework of design affordances for teacher adaptation, ownership, and community that will inform and shape the next phase of our work. Our aim is to scale our work to 100 schools, with a minimum of 200 teachers and 5,000 pupils. Although we address these affordances under three headings, they are, of course, interrelated.

In this research context, the first affordance is the *technology itself*, which proved critical in driving the process of change in teachers, particularly as interaction with the technology revealed mathematical ideas that may not have been initially evident to all of them. In the PD sessions, we noted several instances where the mathematics was unclear to certain teachers because of its novel representation, and where the use of the technology encouraged them to reconstruct or enhance their understanding. An example was that several teachers did not themselves correctly how backwards motion on distance time graph would be represented. At the same time, some of the most interesting aspects of the discussions and activities were catalysed by often unsuspected links between representational forms.

Our qualitative data showed the standard problems in introducing technology (e.g., software downloads, computer room availability), but in this particular instance of a replacement unit, we can defend the idea that it fostered teacher (re-)consideration of how they taught, and in some cases *what* they taught. This finding is hardly surprising: the tendency of new technologies to engender re-evaluation of what is possible is well documented (e.g., Noss and Hoyles, 1996).

The leverage of technology reaches beyond such re-evaluations, however, because the use of digital technologies is fast becoming mandatory in England: it confers high status on head teachers, who want to be seen to participate in the latest trends; moreover, policy-makers and inspectors check on whether technology is being used in useful ways. For once we are cutting with the grain.

The second affordance for adaptation is to shift from *scripting* to *steering*. We charted a path through the students' activities, a path that we wanted *all* the teachers to follow. Yet, in a show of autonomy at a PD session, one teacher decided: 'OK, I've got the idea. I'm going to use [the CM materials] but I'm going to do it my way, completely differently'. It is tempting to attribute this to the feistiness of one teacher, and what she achieved remains unclear at this point, but she drafted ideas for her practice that appeared to be aligned with our goals as well as pragmatic from her school's perspective. We can, perhaps, also take some credit. We organised the CM Unit as one in which the teacher's role was not so much scripted as *steered*. We charted the course, showed them the 'big ideas', but we were at pains to leave room for teachers to 'get the idea' and adapt it in their own way.

In part, the tension between adoption (do as I say) and adaptation (find ways to fit what I say into what you think and do) is exacerbated by the constraints of the system. In England, close alignment with the statutory National Curriculum is imperative. And the centralized control of teachers (through the only recently discontinued National Strategies, recommended lesson structure, and inspection, which has deeply influenced most educators in England) is delicately balanced with teacher professionalism.

A major challenge for the future is to encourage teachers to move from adoption to adaptation, with a greater awareness, perhaps, of the potential of instrumentalisation. The innovation began to develop momentum at the point when teachers started to buy into tool use, to think beyond 'curriculum fit', and to adapt the innovation for themselves in ways that enhanced the tool's epistemic value. Clearly, no rigid dichotomy exists between routine use and epistemological tool use, and we wanted the 'tools to travel between' these poles. This endeavour was in fact successful not least because of the ways in which the teachers adapted the materials. Inevitably, there is a dialectical relationship between what is being fitted and what it is being fitted into. We provided scaffolding and guidance so that teachers could make the innovation their own and make it fit their setting, without undermining the intended epistemic student experience. Because teachers clearly valued the mathematical experience, we are optimistic that, in our next phase of work, we will be able to build on what we have learned to scale up to a much larger number of schools.

A note of caution: positive responses were not universal among the sample teachers. We know that 3 of the teachers simply 'taught the unit' and, despite our best endeavours, might not have been aware of the powerful ideas informing it. They made no adaptions of the materials of any sort. However, these teachers have made efforts to continue with the work and teach the CM units again with other groups of students. This seems to be to consolidate their new practice but again maybe for different motives (gleaned from interviews with them) —esteem from being part of a project or even as a way of enhancing career opportunities.

The third affordance is finding the *right grain-size for the manipulable elements of innovation*. Where on the spectrum between closed simulations and open-ended toolkits should we design? What is the relationship between what we would like students to learn, and the technology tools they have at their disposal? How can we design computational interfaces that align with what we are trying to teach? The software design field in education has yet to provide definitive conclusions for the many questions of this type. In our CM research, we have found it useful to generalise these questions beyond software design to the design of other elements of the activity system: the tasks, written materials, and the formative assessment techniques can all benefit from consideration of the precise objects to which we want students, teachers, and researchers to attend. These questions must be resolved before we can be confident of scaling to 100 schools. Our current work, for example, is seeking to determine whether open systems (like SimCalc, Sketchpad, or Cabri) should be replaced by tools that are co-designed alongside the 'curriculum'. Doing so, however, will inevitably limit what is possible in order to enhance what is probable.

This is the strength of the curriculum activity system, but it poses a challenge and tension for a constructionist programme. The teachers' reaction to the innovation consistently stressed the importance of the constructive element (making things happen and attending to them) and the discursive element (building things to share them and create a shared basis for discussion and reflection). However some moved to class demonstration on the interactive whiteboard and in the process sometimes curtailed the time for students to engage in hands-on activity

The most favourable examples of school take-up of CM should be considered against the background of the high stakes associated with mathematics achievement and with the 'grading' of schools; heads (principals) are under pressure *not* to innovate, continuing to use established, conservative pedagogies that minimise the risk of failing to attain expected achievements. Yet our PD sessions revealed a surprising reluctance to view the 'replacement' module as a single entity as we have noted. Instead most teachers adapted its use in unanticipated ways.

We surmise that fostering this autonomy can be designed into our next phase of work. The key seems to have been the integration of the Unit into existing schemes of work. The most notable instances of integration occurred when pressure from 'above' (i.e., from the research team with the agreement of the

Head of the school) was combined with active participation from 'below' (i.e., from the teachers themselves) and with executive leadership from the 'middle' (e.g., subject leaders' enthusiastic endorsement). Finding ways to support teachers from different schools to share online how they might achieve this integration will be a key design challenge in the next phase of the scaling process.

## References


Artigue, M. (2002). Learning Mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, 7, 3, 245-274.

Baker-Doyle, K.J. (2011). *The networked teacher: How new teachers build social networks for professional support*. New York, NY: Teachers College Press.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating Usable Innovations in Systemic Reform: Scaling Up Technology-Embedded Project-Based Science in Urban Schools. *Educational Psychologist, 35(3),* 149-164.

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9, 88-113.*

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32(6),* 3-12.

Confrey, J., Hoyles, C., Jones, K., Kahn, K., Maloney A., Nguyen, K., & Pratt D. (2009). Designing software for mathematical engagement through modelling. In C. Hoyles & J.-B. Lagrange. (Eds.) *Digital technologies and mathematics teaching and learning: Rethinking the terrain* (pp. 19–46). New York, NY: Springer.

Drijvers, P., & Trouche, L. (2008). From artefacts to instruments: A theoretical framework behind the orchestra metaphor. In G.W. Blume & M.K. Heid (Eds.), *Research on technology and the teaching and learning of mathematics: Vol. 2. Cases and perspectives* (pp. 363–392). Charlotte, NC: Information Age.

Drijvers, P., Kieran, C., & Mariotti, M.A. (2010). Integrating technology into mathematics education: Theoretical perspectives. In C. Hoyles, & J.-B. Lagrange (Eds.), *Mathematics education and technology – Rethinking the terrain* (pp. 89–132). New York, NY/Berlin, Germany: Springer.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., . . . et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington DC: National Center for Educational Evaluation

Heid, M. K., & Blume, G. W. (2008). Algebra and function development. In M. K. Heid & G. W. Blume (Eds.), *Research on technology and the teaching and learning of mathematics: Research syntheses* (Vol. 1, pp. 55–108). Charlotte, NC: Information Age.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., et al. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511

Hoyles. C., & Lagrange, J-B. (2009). *Mathematics education and technology-Rethinking the terrain*. New York, NY: Springer

Hoyles, C., & Noss, R. (2003). What can digital technologies take from and bring to research in mathematics education? In A.J. Bishop, M.A. Clements, C. Keitel, J. Kilpatrick, & F.K.S. Leung (Eds.), *Second international handbook of research in mathematics education* (pp. 323–349) Dordrecht, the Netherlands: Kluwer.

Kaput, J. (1992). Technology and mathematics education. In D. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (pp. 515–556). New York, NY: Simon and Schuster Macmillan.

Kaput, J., Hegedus, S., & Lesh, R. (2007). Technology becoming infrastructural in mathematics education. In R. Lesh, E. Hamilton, & J. Kaput (Eds.), *Foundations for the future in mathematics education* (pp. 173–192). Mahwah, NJ: Lawrence Erlbaum Associates.

Kaput, J., & Roschelle, J. (1998). The mathematics of change and variation from a millennial perspective: New content, new context. In C. Hoyles, C. Morgan, & G. Woodhouse (Eds.), *Rethinking the mathematics curriculum*. London, England: Falmer Press.

Kynigos, C., Clayson, E., & Yainnoutso, N. (2012). *Constructionism 2012, theory, practice and impact: Conference proceedings.* Athens, Greece: Educational Technology Lab

Laborde, C. (1995). Designing tasks for learning geometry in a computer-based environment. In L. Burton & B. Jaworski (Eds.), Technology in Mathematics Teaching - a bridge between teaching and learning (pp.35-68). London: Chartwell-Bratt.

Mishra & Koehler (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings: Learning Cultures and Computers*. Dordrecht, the Netherlands: Kluwer Academic.

Noss. R., & Hoyles C. (submitted 2013). Constructionism & microworlds. In R. Duval, R. Sutherland, & M. Sharples (Eds.) *The Technology-Enhanced Learning Reader*. New York, NY: Springer.

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. London, England: Harvester Press.

Penuel, W.R., Fishman, B.J., Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Second Ed). Newbury Park, CA: Sage Publications.

Roschelle, J., & Jackiw, N. (2000). Technology design as educational research: Interweaving imagination, inquiry & impact. In A. Kelly & R. Lesh (Eds.), *Research design in mathematics & science education* (pp. 777-797). Mahwah, NJ: Lawrence Erlbaum Associates.

Roschelle, J., & Shechtman, N. (2013). SimCalc at scale: Three studies examine the integration of technology, curriculum and professional development for advancing middle school mathematics. In J. Roschelle & S. Hegedus (Eds.), *The SimCalc vision and contributions: Democratizing access to important mathematics* (pp. 125–44). Berlin, Germany: Springer.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S. . . . Gallagher, L. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics. *American Educational Research Journal 47*(4), 833–878.

Roschelle, J., Tatar, D., & Kaput, J. (2008). Getting to scale with innovations that deeply restructure how students come to know mathematics. In A. E. Kelly, R. Lesh & J.Y. Baek (Eds.), *Handbook of Design Research Methods in Education*. New York: Routledge. 369-395.

Roschelle, J., Tatar, D., Shechtman, N., & Knudsen, J. (2008). The role of scaling up research in designing for and evaluating robustness *Educational Studies in Mathematics*, *68*(2), 149–170.

Schneider, B., & McDonald, S.K. (Eds.). (2007). *Scale-up in Education: Issues in Practice (Vol. 2).* New York: Rowman & Littlefield Publishers, Inc.

Sinclair, N., Arzarello, F., Gaisman, M., Lozano, M., Dagiene, M., Behrooz, E., & Jackiw, N. (2010). In C. Hoyles & J-B. Lagrange (Eds.) *Mathematics education and technology-Rethinking the terrain* (pp. 61–78). New York, NY: Springer.

Sturman, L., & Cooper, L., (2012). Evaluation of Cornerstone Mathematics pilot in England: Unit 1, linear functions. Slough, England: National Foundation for Educational Research.

Vahey, P., Knudsen, J., Rafanan, K., & Lara-Meloy, T. (2013a). Curricular activity systems supporting the use of dynamic representations to foster students' deep understanding of mathematics. In C. Mouza & N. Lavigne (Eds.). *Emerging technologies for the classroom: A learning sciences perspective* (pp. 15–30). New York, NY: Springer.

Vahey, P., Roy, G., & Fueyo, V. (2013b). Sustainable use of dynamic representational environments: Toward a district-wide adoption of SimCalc-based materials. In S. Hegedus & J. Roschelle (Eds.). *The SimCalc visions and contributions: Democratizing access to important mathematics* (pp. 183-202). New York, NY: Springer.

Dear Prof. Hoyles, dear Celia.

Your revised version of the manuscript has now been reviewed.

We are highly satisfied with the changes made and only ask for minor changes. Below you will find the reviewers comments, which provide you with a few critical remarks, which I kindly ask you to address in the next and probably last version of your paper. Especially the more fundamental critical remarks by reviewer 2 on the necessity to extend and theoretically underpin the conclusion part are highly convincing.

Because your paper is already 60,000 characters long, I propose that you use additional 3,000 characters including space in order to meet the proposal of reviewer 2 for a more critical and theory-based conclusion part.

We have added in Section 1 reference to alternative approaches with appropriate references. We have clarified the curricula activity framework that underpins the US approach and extended the description of instrumentalisation (partly this is also a matter of differentiating it from adaptation, as requested. we have redrafted the Conclusions Next Steps that are more speculative and better aligned to the aims of the paper

May I add as minor remarks the following small issues:
- the introductory sentence is not very challenging, please use either "we" or rephrase the sentence; Done
- is it helpful to include only one control group in the quantitative part, does that really say something as long as there is no British control group? We have redrafted and argued why there is little likelihood of weakened validity in having just one control group

- please introduce CM as abbreviation for Cornerstone Mathematics Done
- check inconsistences, e.g. inconsistent usage of dots at the end of headings Done

Your revision is due by 04 Sep 2013. I hope very much that you can meet this tight deadline, we are under strong time constraints, because the "end of the year" is nearing (at least for publishers).

To submit a revision, go to http://zdmi.edmgr.com/ and log in as an Author. You will see a menu item call Submission Needing Revision. You will find your submission record there.If you have any questions, please do not hesitate to contact either the issue editors or me.

Yours sincerely,

Gabriele Kaiser
Editor-in-Chief
ZDM - The International Journal on Mathematics Education

Reviewers' comments:

Reviewer #1: This new version has been clearly improved, and the authors have taken into account most of the reviewers' comments. Only some minor changes could be made for the last version:
- some formal details (double space between two words, << et al. sometimes partially in italics, some coma missing, etc.) Done

- about the keywords, perhaps relevant to add << scaling , and << adaptation (or instrumentalisation) ? Done
- Cheung and Slaving 2011, evoked in the article, missing in the final list of references Done
- There are actually two Vahey et al.. 2013 references, to be distinguished with 2013a and 2013b ? Done
- Some authors missing in two references Dynarski et al. 2007, and Rochelle et al. 2010, perhaps because there is too many authors ? WE have made this consistent
- P. 21 << conceptual and affective change , the meaning of << affective in this context, is not clear Deleted both adjectives
- At several places, we find the expression << adaptation and instrumentalisation , perhaps necessary to situate these two concepts (synonymous ?) we have elaborated the meaning of instrumentalisation to make the distinction clearer
- Last remark, I think that a reference to the article Drijvers and Trouche (about orchestration) should be relevant at the first occurrence of the this notion. Done

Reviewer #2: This is a much improved paper and I enjoyed reading it very much. It is still largely a descriptive piece however. The article would be stronger if the concluding discussion were changed in two ways:

The concluding discussion is somewhat speculative and comes across as a little overly optimistic. The issue of moving from adoption to adaptation is certainly an interesting one, although the examples cited are largely of adaptation. I'm somewhat surprised by this. My own experience of such
initiatives is that many teachers claim to do more than they actually do both in terms of adoption and of adaptation. I think it might help if the adoption / adaptation issue could be framed in terms of the aspects of Cornerstone that were negotiable / non-negotiable (as in Dylan Wiliam's "tight but loose" approach.) Here it would be useful to return to the theory - and to use this to conceptualise and discuss adoption / adaptation. Currently, the theory is hardly used at all. It would be useful to consider when a modification is adoption and when it is adaptation.
I would also like some acknowledgement of alternative approaches to scaling up and some discussion of how one might evaluate the Cornerstone approach in relation to alternatives together with what might be learnt from such a comparison.es

we have addressed this concern. we have included reference to some alternative approaches and theories to scaling and clarified the theoretical underpinning instrumentalisation and evidence of it use. we are aware of Wiliam's work and re read it but at this point cannot see it helps our analyses - -

Minor issues:

1. Generally, I felt the article to be somewhat over-referenced. Better to focus in depth on a few key references. WE have deleted some references

p.9, Section 3.2: Tabulation of the sample would be helpful. We have tried, but it is difficult to present, so have decided against. However we clarified the numbers to make it more straightforward to understand the sample

p. 10, line 6: How many teachers? 16 or 17? Or 18 (p.14, line 37)? Done

p.12, line 1/2: Explain GCSE briefly in a footnote. Done (we have made it in a bracket as footnotes proved impossible – messed up the template!).

p.18, line 33: Two out of nine early adopters is not in my view a huge cause for optimism. We have clarified why it was worth a degree of optimism – putting into scheme of work etc.

p.19, line59: "OK. I've got the idea. I'm going to use Š but I'm going to do it my way, completely differently". So what exactly did the teacher change? A good point, and we have modified to say we will monitor the situation.

p.21, line 3/4: I don't understand the reference to open systems and the contrast with tools designed alongside the curriculum. I assume that the latter reference is to very specific "one function" software. If so, it is somewhat odd, since there has been no discussion of this earlier in the article. Good point. We have fixed this by deleting 'systems' and replacing with 'open toolkits'.