Large Scale Analysis of Student Workbooks: What Can We Learn About Learning?

Nicole Shechtman, SRI International, nicole.shechtman@sri.com
Jeremy Roschelle, SRI International, jeremy.roschelle@sri.com

Abstract: As Learning Science–based innovations are studied at scale, traditional Learning Sciences methods such as video analysis and classroom observation become impractical. Yet Learning Scientists want to know more about student misconceptions, the connections between writing and conceptual understanding, and classroom practices than pretests and posttests can reveal. In this paper, we discuss an exploratory posthoc analysis of 765 workbooks from 48 classrooms that implemented SimCalc. These classrooms participated in a large-scale experiment in which we found that students learned more advanced mathematics in classrooms that implemented SimCalc. A team of three master teachers coded the workbooks for completeness, correctness, and other impressions. We found characteristics of students' work that predict gain scores, including the first evidence that classic SimCalc activities—writing stories and drawing graphs about motions—impact student learning. We discuss potential implications for large-scale Learning Sciences research.

Introduction

Although much Learning Science research is conducted at a small scale (e.g., within a handful of classrooms per research study), most Learning Scientists want to have a large impact on potentially hundreds or thousands of schools. When going from one scale to another, more variance will inevitably occur. Classrooms are complex places and differ in myriad ways. Some variations in classrooms and in implementation may have dramatic consequences for the effectiveness of a design that was successful in smaller scale trials. Ann Brown, for example, famously highlighted the potential for "lethal mutations" that could undermine the effectiveness of an otherwise promising innovation (Brown & Campione, 1996). Thus, a key part of the scaling-up process is research on implementation (e.g., Cohen & Hill, 2001)—seeking to understand what kinds of variation in implementation occur and which variations matter for desired outcomes.

Research on implementation can entail the collection and analysis of a variety of different types of data, including video recordings, field observations, teacher logs, surveys, and artifacts from the classroom. At the scale considered in this work, involving implementation in 48 classrooms in the same school year, researchers face difficult tradeoff decisions in choosing what data to collect. One key tradeoff is between the *cost* of collecting data and the *quality of inferences* the resulting data can support. For example, video recordings are very expensive to collect and analyze but support in-depth analysis of a type that many Learning Scientists prefer. It can be much less expensive to collect and analyze a survey, but inferences from such data may be more circumscribed.

In this paper, we consider the potential value of collecting student workbooks in implementation research at scale. Student workbooks provide a potentially useful tradeoff. On the one hand, they are quite inexpensive to collect; we have found that teachers are willing to put them in a prepaid mailer, eliminating the need for a researcher to visit the classroom. On the other hand, workbooks directly exhibit student work, which can be a rich source of insights into students' misconceptions, for example. We report what we were able to learn from an exploratory, posthoc analysis of 765 workbooks.

We consider two kinds of research questions, some that are general to the method and some that are specific to the SimCalc intervention.

We asked three *method questions*:

1. How long does it take to code 48 classroom sets of workbooks for a 2–3 week intervention?
2. Can suitable interrater reliability be achieved?
3. Could raters accurately detect indicators of high-achieving and low-achieving classrooms?

With regard to the SimCalc intervention, the program developers espoused the beliefs that students need to be actively engaged in using the materials in order to learn (e.g., rather than just watching a teacher lecture with the materials). It is difficult to collect evidence of active engagement of all students in 48 classrooms. However, workbook completion does reveal the extent to which students were engaged in doing the work, especially because the curricular workbooks required students to use the software in order to answer various questions. Furthermore, because students drew graphs and wrote stories in their workbooks, we were able to consider the impact of student engagement with these forms of representation on learning.

Consequently, we asked an additional two *research questions*:

4. Does the completeness of student workbooks, as a proxy for students' active engagement in doing mathematics during the course of the replacement unit, predict student learning gains?

5. Does the correctness of details in student explanations, drawings of graphs, and telling of stories about graphs predict student learning gains of advanced mathematical concepts?

In addition to a specific focus on these research questions, we will report some additional anecdotal insights about what we learned from the analysis of workbooks.

Background: Analyses of Student Work

In some ways, analysis of student work is ubiquitous in the Learning Sciences. Many published articles show artifacts from the classroom—including students drawings, calculations, graphs, and tables—to provide evidence for their arguments. Student work is often useful, for example, for identifying student misconceptions in graphing (Clement, 1989). However, most analyses are in the context of case studies. Few involve quantitative comparisons of students' prior knowledge and learning outcomes.

One notable use of student work in larger scale research was in the work of the Consortium on Chicago School Research, which collected samples of teacher assignments and student work in 12 Chicago schools at three grade levels and both in mathematics and in language arts (Newman, Lopez, & Bryk, 1998). The researchers found that when teachers gave more challenging assignments (i.e., that required a higher level of intellectual work), students were more likely to exhibit high level work and to have higher achievement outcomes.

Analysis of student work is also featured as a method of teacher professional development and improvement of instruction. For example, in Cognitively Guided Instruction, the entry point for teachers into students' cognition is through examining examples of student work (Fennema et al., 1996). Student work has also been the subject of research related to the benefits of portfolio assessments (Wolf, 1989). The purpose of such work is to form an alternative outcome measure to conventional testing. Although this is a different purpose from the present research—we seek to use student work as an implementation measure, not an outcome measure—one useful commonality is the technique of developing rubrics to score student work.

Background: Scaling Up SimCalc

The Scaling Up SimCalc Project (Roschelle et al., in press) implemented two randomized controlled experiments (with one embedded quasi-experiment) designed to address the broad research question, "Can a wide variety of teachers use an integration of technology, curriculum, and professional development to increase student learning of complex and conceptually difficult mathematics?" There were two interventions, one for seventh grade and one for eighth grade. Each intervention integrated the representational technology SimCalc MathWorlds, curriculum workbooks, and teacher professional development organized around a 2-3 week replacement unit on rate, proportionality, and linear function. The replacement units incorporated the following hallmarks of the SimCalc approach to the mathematics of change and variation:

1. Anchoring students' efforts to make sense of conceptually rich mathematics in their experience of familiar motions, which are portrayed as computer animations.
2. Engaging students in activities to make and analyze graphs that control animations.
3. Introducing piecewise linear functions as models of everyday situations with changing rates.
4. Connecting students' mathematical understanding of rate and proportionality across key mathematical representations (algebraic expressions, tables, graphs) and familiar representations (narrative stories and animations of motion).
5. Structuring pedagogy around a cycle that asks students to make and compare their predictions.

The main effects of the treatment in the three main studies were positive, with student-level effect sizes of .63, .50 and .56; classrooms that used a SimCalc replacement unit had students who learned more advanced mathematics. This article focuses on data from the 48 treatment classrooms in the first experiment, which centered around seventh-grade students, teachers, and mathematical content.

SimCalc's developers have always espoused the view that students must be actively engaged with doing mathematics in order to learn; the software is a representational infrastructure that supports doing mathematics (Kaput, Hegedus, & Lesh, 2007). We confirmed this view, albeit weakly, in the larger quantitative data set by a finding that the number of days in the computer lab (a proxy for the level of student use of the software) correlated with overall learning gains. Case studies conducted within the Scaling Up SimCalc research provided further opportunities to examine this view within varying classroom implementations. Empson, Greenstein, Maldonado, and Roschelle (2009) examined differential student learning in three different classrooms, each with a different style of implementation. The analysis highlighted access to or blocking of learning resources as critical to students' ability to participate in the classroom as cognitively engaged learners. When a teacher significantly blocked students' autonomous use of SimCalc software and workbooks, by frequently interrupting the students' independent work, the classroom learning gains were lower. Likewise, Dunn (2009) found one low performing teacher spent much more time than the average teacher introducing and demonstrating ideas and much less time allowing students to work with the computer and workbooks. Analysis of student workbooks provides an opportunity to go beyond these case studies by examining connections between student completion of workbooks (and hence their level of engagement in doing the mathematics) with student learning.

SimCalc's developers have also always emphasized the activity of asking students to construct stories about graphs and motions (Kaput & Roschelle, 1998). These stories are seen as valuable because they evoke student engagement when they focus students on mathematically relevant details. For example, whether a students' story is about a race or a soccer game is irrelevant; whether a students' story describes speed and direction of motion is relevant. An emphasis on connecting linguistic and graphical forms of meaning is broadly consistent with the "Multimedia Principle" (Fletcher & Tobias, 2005). However, to date, there has been no empirical confirmation of this principle specific to SimCalc. Likewise, SimCalc developers believe that asking students to draw graphs and to explain motions are valuable activities. Analysis of student workbooks provides an opportunity to code these aspects of student work and to examine correlations to learning gains.

SimCalc Workbooks

The seventh-grade curriculum, entitled *Managing the Soccer Team*, was designed to be used daily over a 2- to 3-week period to address central concepts of proportionality. Speed as rate is developed through a sequence of increasingly complicated simulation-based activities. The workbook is 59 pages long and has a total of 20 lessons. Lessons progress through representations—from graphs to tables to equations—aiming to teach students to translate among representations and to connect each concept to verbal descriptions of motion or other real-world contexts.

Data Collection

Teachers were recruited for the Scaling Up SimCalc experiments from several regions of Texas. Recruitment and sampling in the overall experiment is described in Roschelle et al., in press. The seventh-grade treatment sample consisted of 796 students in 48 teachers' classes (for each teacher we collected data for one randomly selected class). The student ethnic breakdown was 48.5% White, 44.3% Hispanic, 4.2% African American, and 1.5% Asian. The mean campus-level percentage of students qualifying for free or reduced-price lunch was 54%, with a range of 1% to 94%, indicating a wide variation in campus poverty levels.

We asked all teachers participating in the Scaling up SimCalc experiment to return student workbooks to us and provided a prepaid express mail shipping box for them to do so. For the seventh-grade treatment teachers, we received boxes of workbooks from all 48 teachers, resulting in a collection of 765 workbooks (overall, 3.9% of the student workbooks were missing). Before analysis, each workbook was assigned the identification number that cross-referenced all other student-level data.

In addition to the workbooks, other data collected included: (1) unit pretests and posttests, (2) student and teacher demographics, (3) teacher mathematical knowledge, and (4) teacher daily implementation logs. Instrumentation and findings for these data are described extensively elsewhere (e.g., Roschelle et al., in press).

Coding Protocol and Process

Based on a pilot workbook analysis, we designed a coding protocol that would be implemented by experienced math teachers. Teachers coded workbooks grouped by class so that they could observe patterns that may emerge within a class. There were three parts to the protocol.

1. For each workbook, the coder reviewed each of the 20 lessons and applied criteria to determine the level of completeness of the activities in the lesson (no attempt, low, medium, high). Coders did not make judgments about the quality of the work, but rather the presence of any attempted work.
2. For each workbook, the coder reviewed a subset of six key lessons to code the written work for correctness. The research team selected the six activities that seemed most central to the development of students' conceptual understanding. Four of these lessons entailed students' construction of stories from linear and/or piecewise linear graphs or construction of linear and/or piecewise linear graphs from stories. To capture the specificity of mathematical detail for each of the constructed stories and graphs, the coding protocol entailed a checklist of all possible correct mathematical details.
3. After coding a classroom for completeness and correctness, the coders made a holistic guess about the prior achievement level and learning gains in the classroom. The coders made written documentation of the indicators they used to make these judgments and recorded misconceptions that they observed.

We hired three master teachers who had not participated in the experiment to serve as coders during their school summer break. All were mathematics teachers, with one working at the high school level and two working at the middle school level. They all had little or no experience in conducting research.

The coders were trained on a set of 20 workbooks. They coded 5 workbooks together with one of the researchers and then 15 on their own. Interrater reliability on Parts 1 and 2 was sufficiently high for coders to then begin coding the full corpus. Within each class, two workbooks were randomly selected to be double-coded for Parts 1 and 2 to check for the maintenance of reliability.

Analysis

We report our data analysis by each research question.

1. How long does it take to code workbooks?

We found it took the three teachers about 2 months working roughly half time to code 765 workbooks, including 3 days for training and 1 day for debriefing. The approximate time per workbook was 10–20 minutes, which translates to an average of about 4 hours per classroom, for classrooms implementing a 2–3 week long curriculum workbook. This was a relatively inexpensive task in the context of a multimillion dollar project.

2. Were we able to achieve adequate interrater reliability?

Interrater reliability was high. The index that we used was the intraclass correlation between each of the coder's ratings. Among the completeness and correctness ratings, there was a total of 26 ratings. The intraclass correlations for 20 of these ratings was over .90, 5 were between .80 and .90, and 1 was .66.

3. Could raters accurately detect indicators of high-achieving and low-achieving classrooms?

As shown in Figure 1, we found that coders were somewhat accurate in their holistic guesses about classrooms' average prior achievement levels [$t(33) = 2.5$, $p < .05$]; however they were not accurate in their holistic guesses about classrooms' average learning *gains*. We assume this is indicative of the characteristics of the student work, not these particular teachers' ability to detect learning gains.

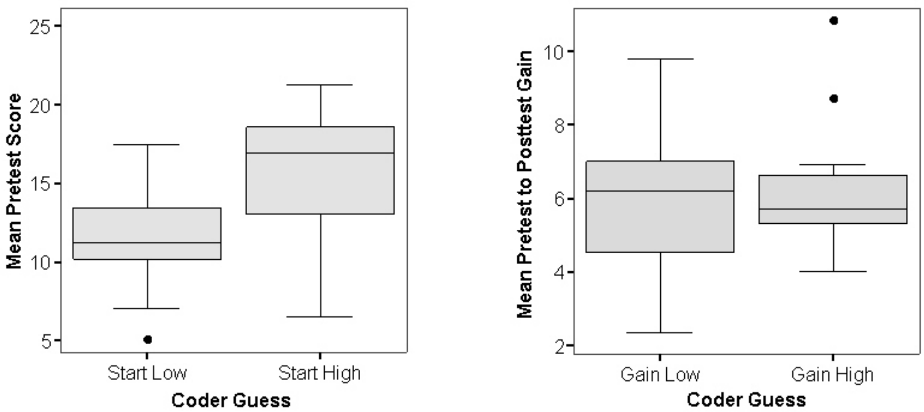

Figure 1. Actual classroom-level means distributed over coders' holistic guesses of achievement levels.

As summarized in Table 1, we debriefed the coders after completing this task to ascertain the qualities of student work that were salient to coders as they made these judgments. In addition, teachers were asked to collect misconceptions they observed in the workbooks. We found they detected a wide range of common misconceptions for this mathematical subject matter, including:

- Confusion between miles and mph—not paying attention to the fact that these are different.
- Distance interpreted as speed and speed interpreted as distance.
- Interpreting an object moving at a constant rate as increasing in speed.
- Confusion regarding which runner is the winner of a race based on reading a position graph.
- Confusion between dollars and decimals (e.g., "0.80 cents").
- In a position graph representing two races, the graphed line "on top" is more.
- Iconic interpretation of graphs as a picture (e.g., "the runner went up" describes a positively sloped position graph).
- In a piecewise linear function, two segments interchanged or confused on graph.
- Incorrect use of conventions (e.g., "13-50" as the ordered pair (13,50)).

Table 1. Qualities of student work that were salient to coders in judging student prior achievement and learning.

Indicative of High Achievement	Indicative of Low Achievement
• Used pencil and erased when they made a mistake. • Attention to correcting their work. • Students were allowed to work on their own. The stories within a classroom were very different from each other, compared to classes where all the kids had the same stories verbatim. • Indication of testing predictions with the software. Say that they saw their prediction was wrong. • Quantified statements, rather than just saying "faster." • Graphing was detailed. For every hour they would plot a point. Paid attention to precision, making sure they ended at the right point on the graph. Used straight edges. • Stories included mathematical observations from the graph. Wrote a lot and were very specific (not simply a creative story). • Evidence of teacher feedback and modeling.	• Sloppy handwriting, incomplete answers, poor spelling, reverse letters, graphing without a straight-edge, wrote in pen, did not correct answers. • Lack of detail, not paying attention to answering questions completely. • A lot of absences—large numbers of pages with missing work. • Struggled with math vocabulary and writing skills. Could not accurately describe motion because they had poor word choice. "The van caught up" instead of "the bus slows down" was very important. • Stories tended to be more creative but were missing a lot of mathematical detail. • In some classrooms, all students had the same incorrect answers. This suggested a teacher error or misconception.

4. Does completeness of student workbooks predict learning gains?

Overall, we found that the mean completion of student workbooks was 75% (SD = 17%), which suggests that most teachers really did implement the SimCalc intervention. As Figure 2 shows, there was considerable variation in workbook completion, both within and across classrooms. The figure shows classrooms in rank order of the classroom median for student completion of workbooks. About two-thirds of the classrooms completed 75% or more of the workbooks. Of the remaining third, three classrooms completed 50% or less of the workbook.

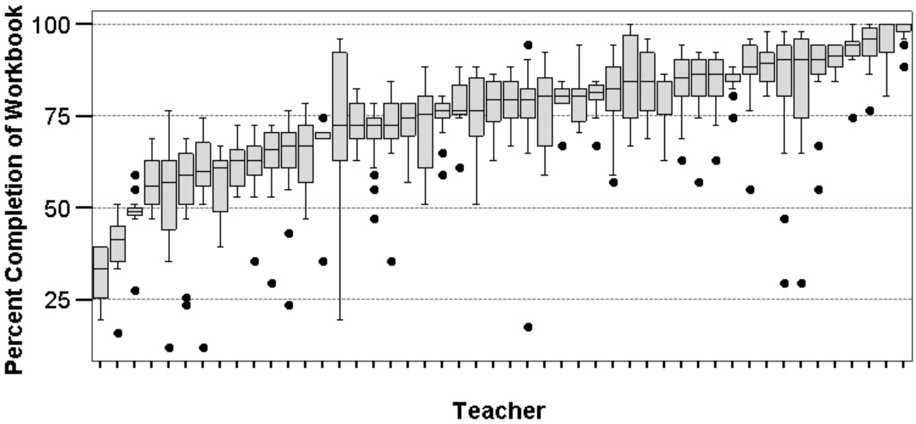

Figure 2. Distributions across classrooms of completion of workbook (teachers ordered by class median).

We then tested whether workbook completion predicted pretest to posttest learning gains, using a hierarchical learning model (HLM) to correct for the nesting of students within classrooms. We examined the "M_2" subscale of our assessment, which focuses on more advanced mathematics. We chose to focus on this scale because classrooms that used SimCalc had particularly large gains on this scale; hence, it is sensible to look at this scale when considering the value added by the SimCalc intervention. As Figure 3 shows, there was an overall correlation between student completion of workbooks and mean classroom learning gains [$z = 4.3$, $p < .001$]. Using HLM, we determined that on the 21-point M_2 scale, students on the average gained 1 M_2 point for every 13.3% of the workbook that was completed.

We then examined completion of each of the 20 lessons individually. We found that the completion rate was uniformly high for the first half of the workbook (with the exception of a section intended for homework); declines in workbook completion occurred in some classrooms in the second half of the workbook. We found a high alpha (0.83) among completion rates for all activities, suggesting that whether a student

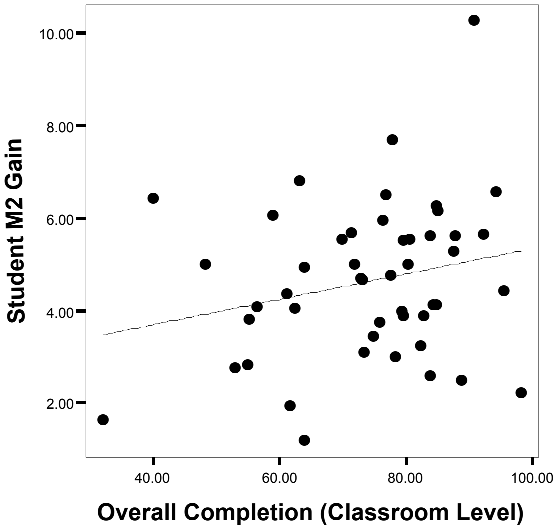

Figure 3. Workbook completion and student learning gains.

completes one activity is a good indicator of whether they will complete other activities. With respect to learning gains, we found correlations between completion of most activities in the second half of the curriculum and learning gains. This may be due to the fact that the second half of the workbook covered the material most relevant to the M_2 scale and/or had the most variation in completion.

We then examined factors that might predict completeness of the workbooks. The best predictor of workbook completeness was students' pretest score. The pretest had 30 items. For every 10 additional items correct, students on average completed an additional 6% of the workbook [z = 5.92, p < .001]. The other statistically significant predictor was gender. On average, girls completed 1.8% more of the workbook than boys did. The other variables that were not significant included: geographic location, student ethnicity, school poverty level, teachers' years-of-experience, teachers' level of mathematical knowledge, how often class was conducted in the computer lab, and the number of days spent implementing the unit. Clearly workbook completion is not a proxy for computer use.

We also used the number of days the unit was implemented in conjunction with the completeness to examine the *pace* with which the class did the unit. Overall, workbook completion was independent of the time teachers' spent doing the unit. However, we did find that classrooms with a higher mean pretest score completed the workbooks faster [r (48) = .42, p < .01]. Teachers may be purposefully taking more or less time with the workbooks depending on the incoming level of their students, which suggests that most teachers managed implementation to substantially complete the workbooks rather than to occupy a fixed amount of time.

5. Does correctness of details in workbook answers predict learning gains?

Figure 4 shows examples of two students' written stories in response to the graph depicted in the workbook. As we found throughout the workbooks, students varied in how much mathematical detail they provided and whether the detail was correct. In the top story, the student provides little detail about speed and describes the wrong vehicle as stopped. In the bottom story (which is still relatively incomplete compared to other stories we saw), the student describes the constant speed of the bus until it stopped. (The scores "-30" and "-10" were written by the teacher and not part of our coding scheme.) We see that the student in the top story has also made notes on the task: "explain lines using speed, location, time; explain motion." One lesson we learned from reviewing workbooks is that the task may not have been clear to most students and teachers; this teacher may have given the students more information on what they were supposed to do. In a revision of the materials, we would consider more specifically prompting students to focus on mathematical detail.

Overall, we found that the level of mathematical detail and correctness in all six activities predicted student learning, either significantly or marginally significantly. Also, as with the completion rates, we found high intercorrelations between correctness for each lesson (alpha = 0.70). Hence, our ability to argue that doing any specific learning activity is important to learning is limited. However, on a whole this finding does suggest that the activities that focus on drawing graphs and writing stories are valuable for student learning.

As was the case for completeness, student pretest scores predicted correctness in the workbook. No other variable predicted correctness, including student gender, geographic location, student ethnicity, school poverty level, teachers' years-of-experience, teachers' level of mathematical knowledge, how often class was conducted in the computer lab, and the number of days spent implementing the unit.

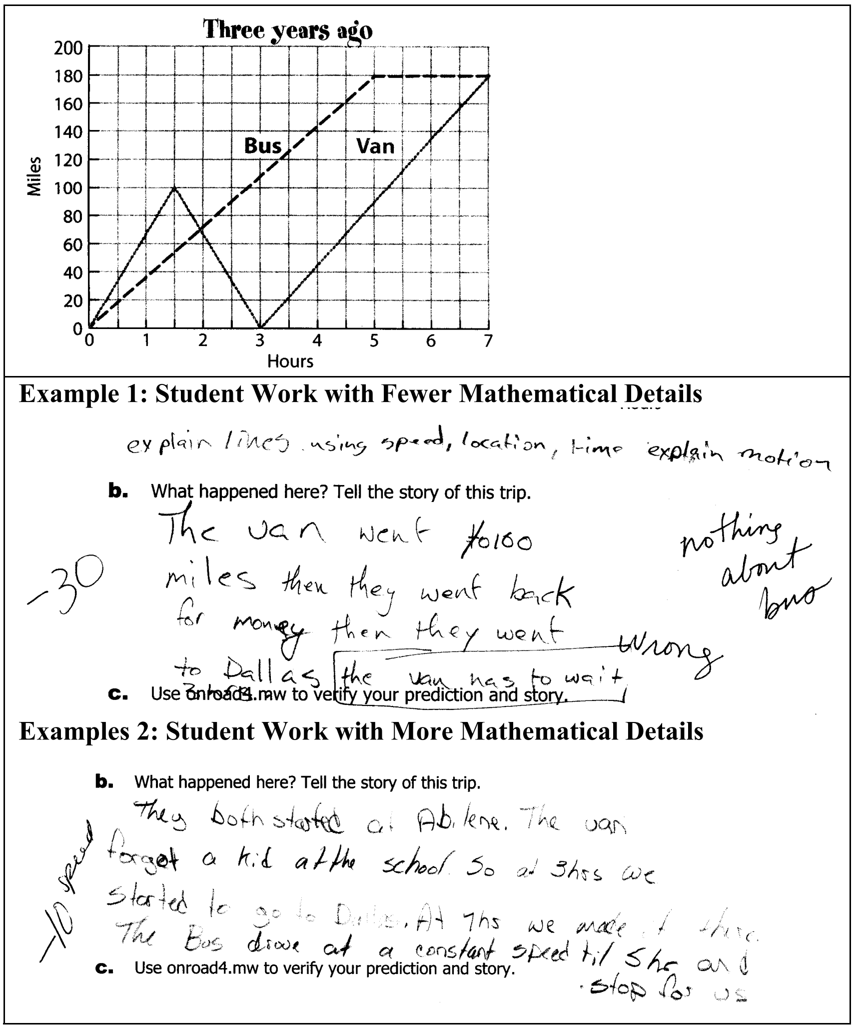

Figure 4: Two contrasting student responses to a story writing task.

Discussion

Our analysis of the first two research questions suggests that analysis of student workbooks is a viable technique for measuring implementation of a Learning Sciences–based innovation at scale. Workbook coding is reasonably cost effective and time efficient, especially relative to methods requiring field observations or video recordings of classrooms. Furthermore, we were able to achieve suitable interrater reliability. Workbook coding has the advantage that it makes it possible to examine the work of the majority of the students in a classroom, whereas field observations or video recordings often have to be focused on just a few students.

The third research question looked at the connections between coders' holistic judgments and students' actual knowledge and learning. We found that coders could detect with some accuracy which classrooms had lower mean pretest scores. However, they did not tend to detect which classrooms had lower or higher mean learning gains. This suggests that coding specific activities for completeness and correctness provides important analytical information that cannot necessarily be obtained through teachers' holistic judgments.

We did find that completeness of student workbooks predicted student learning. This is important because it suggests that having students do the mathematical work is important to student learning, a core tenet of SimCalc program developers. Although this may seem obvious, it is not obvious to all teachers who implement SimCalc—some prefer to focus on teacher demonstrations using the SimCalc software. Indeed, case studies provide evidence of some classrooms in which teachers blocked student independent work; this work confirms the conjecture arising from those case studies that the workbooks are a valuable learning resource and that low use of the workbooks was a factor in lower student learning gains. A corollary finding was that pretest scores predicted completeness. This is an indication of an advantage that students with higher prior knowledge have in learning the content of this unit. This suggests the necessity for additional supports for students who start the unit with lower prior knowledge, a point that could be addressed in teacher professional development to improve implementation.

We also found that the level of detail in student work and the correctness of the detail predicted student learning. We particularly focused on details in students' construction of graphs and stories about graphs. This is the first empirical confirmation that these activities, long part of the SimCalc approach, have an impact on learning.

Of course, because the workbook is designed for and used in a variety of ways as part of the many resources in the classroom (i.e., not as an assessment instrument), inferences about the work must be made with caution. For example, it is possible that students can be actively engaged in doing simulations with the software or having mathematical conversations with the teacher or peers yet not write extensively in their workbooks. However, while these limitations do exist, we believe the methodology does provide an important window into students' cognitive engagement across a large-scale study that would otherwise be unavailable.

Overall, we would recommend that Learning Scientists who are scaling up their curricular designs consider designing workbooks and other student work artifacts that can be collected and analyzed for evidence of the quality of classroom implementations. Relative to field observation or video recording, analysis of student work is relatively cost effective, while preserving the opportunity for detailed insight into how teachers and students are using materials. This form of data collection may be particularly useful in examining the relationship of constructive activities, such as drawing graphs and writing stories, to students' learning gains.

Acknowledgements

This material is based in part upon work supported by the National Science Foundation under Grant Number 04-37861. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also acknowledge and deeply appreciate our three master teachers for all of their hard work in doing this coding.

References

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. *Innovations in learning: New environments for education*, 289-325.

Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics, 11(*1-2), 77-87.

Cohen, D. K., & Hill, H. (2001). *Learning policy: When state education reform works.* New Haven, CT: Yale University Press.

Dunn, M. B. (2009). *Investigating variation in teaching with technology-rich interventions: What matters in training and teaching at scale?* Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

Empson, S. B., Greenstein, S., Maldonado, L., & Roschelle J. (2009). Scaling up innovative mathematics in the middle grades: Case studies of "good enough" enactments. Manuscript submitted for publication.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. *The Cambridge handbook of multimedia learning*, 117–133.

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* (pp.).. London, UK: Falmer Press.

Newman, F. M., Lopez, G., & Bryk, A. S. (1998). *The quality of intellectual work in Chicago schools: A baseline report*. Chicago IL: Consortium on Chicago School Research.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J. & Gallagher, L. (in press). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal.*

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.

Wolf, D. P. (1989). Portfolio assessment: Sampling student work. *Educational Leadership, 46*(7), 35-39.