INCREASING SCIENCE LEARNING IN GRADES 3-8 USING COMPUTERS AND PROBES: FINDINGS FROM THE TEEMSS II PROJECT

Abstract. The Technology Enhanced Elementary and Middle School Science project (TEEMSS II), funded by the National Science Foundation, produced 15 inquiry-based instructional science units for teaching in grades 3-8. Each unit uses computers and probeware to support students' investigations of real-world phenomena using probes (e.g., for temperature or pressure) or, in one case, virtual environments based on mathematical models. TEEMSS units were used in more than 100 classrooms by over 60 teachers and thousands of students. This paper reports on cases in which groups of teachers taught science topics without TEEMSS materials in school year 2004-2005 and then the same teachers taught using TEEMSS materials in 2005-2006. There are eight TEEMSS units for which such comparison data are available. Students showed significant learning gains for all eight. In four cases (sound and electricity, grades 3-4; temperature, grades 5-6; and motion, grades 7-8) there were significant differences in science learning favoring the students who used the TEEMSS materials. The effect sizes are 0.58, 0.94, 1.54, and 0.49, respectively. For the other four units there were no significant differences in science learning between TEEMSS and non-TEEMSS students. We discuss the implications of these results for science education.

Andrew A. Zucker, The Concord Consortium
Robert Tinker, The Concord Consortium
Carolyn Staudt, The Concord Consortium
Amie Mansfield, Education Development Center
Shari Metcalf, Education Development Center

The TEEMSS II Project

The goal of the Technology Enhanced Elementary and Middle School Science (TEEMSS II) project was to address the need for inquiry-based instructional materials for elementary and middle school science teaching that make use of computers and probeware. TEEMS II created and the Concord Consortium is disseminating 15 easily implemented technology-based instructional units for teaching and learning science in grades 3-8. The project also supported associated teacher professional development, as well as research focusing on the effectiveness of this approach to learning science. TEEMSS II was funded by National Science Foundation (NSF) grant # 0352522 from the Instructional Materials Development program.

TEEMS II materials are designed to work with whatever computers and probeware schools choose to use. The project selected age-appropriate, standards-based topics for which technology offers real advantages. The units are modular so they can be integrated with existing curricula or used on their own. The TEEMS II learning strategy is based on student investigations of real phenomena using probes or, in one unit (Adaptation), a virtual environment based on mathematical models. To support student learning, the project produced software, curriculum materials, and an online course for teachers.

TEEMSS II was an extension of a prior TEEMSS pilot project NSF funded by the NSF IMD program (grant #9986419). This pilot project demonstrated the soundness of the overall approach and funded the development of a structure for the materials that was used in TEEMSS II. The pilot project focused on developing just two units and testing them with three groups of teachers, supporting one group of teachers with a face-to-face workshop and the other groups exclusively with online courses. Units that address the *National Science Education Standards* (*NSES*) middle school force and motion and energy transfer standards were developed. Because teachers report that these standards are difficult to achieve, they represented a challenging test of our approach. Pre- and post-tests aligned to the standards were given to students to measure their learning gains (Metcalf & Tinker, 2003).

The TEEMSS Units

TEEMSS II produced 15 units keyed to the *National Science Education Standards* that take advantage of computers, sensors, and interactive models. Five units each were developed for grade levels 3-4, 5-6, and 7-8, with each set of units targeting the five NSES standards: Inquiry, Physical Science, Life Science, Earth and Space Science, and Technology and Design, as well as the National Council of Teachers of Mathematics (NCTM) standards. All of the units are now available to download and use, free of charge, at http://teemss.concord.org/.

Every unit contains two one-week investigations, each with a discovery question, several trials, analysis, and further investigations. There is also a teacher's version of each investigation, which contains background material, answers to questions, and a discussion guide. Table 1 on the next page shows the 15 curriculum units and the technology used in each.

Technology and Science Education

Computers and other digital technologies are an essential part of modern science, but they are not widely used in elementary and middle school science education. For example, in 2000 fewer than one-third of science teachers in grades 5-8 reported ever collecting data using sensors or probes (Hudson, McMahon, & Overstreet, 2002). Yet there is widespread agreement that computer and information technologies should be an integral part of elementary and middle school science teaching, in ways that can greatly improve learning. The *National Science Education Standards* (NRC, 1995), the *Benchmarks for Science Literacy* (AAAS, 1993) and many state education standards require the integration of technology into science teaching and learning starting as early as first grade, as a way to facilitate student inquiry and enhance students' understanding of temperature, pressure, and many other phenomena.

At its core science is about investigating, exploring, asking questions, analyzing, and thinking. Technology is uniquely able to support observation and inquiry in ways that are largely lacking in elementary science teaching. A substantial body of research shows that probeware can facilitate student learning of complex relationships (Adams & Shrum, 1990; Beichner, 1990; Friedler, Nachmias, & Linn, 1990; Krajcik & Layman, 1993;

Laws, 1997; Linn, Layman, & Nachmias, 1987). Similarly, models and simulations allow students to understand through exploration the behavior of systems that are difficult or impossible to understand by other means (Beichner, 1990; Brassell, 1987; Mokros & Tinker, 1987; Thornton, 1997).

Table 1

The TEEMSS II Curriculum Units

Standard	Grades 3-4	Grades 5-6	Grades 7-8
Inquiry	**Sound** Explore sound and vibrations with the SoundGrapher	**Water and air temperature** Mix fluids and measure temperature changes with a temperature sensor	**Air Pressure** Explore soda bottle, balloons and lungs with a gas pressure sensor
Physical Science	**Electricity** Explore light bulbs, batteries, and wires using a voltage sensor	**Levers and machines** Design and test your own compound machine with a force sensor	**Motion** Graph, describe, and duplicate motion using a motion sensor
Life Science	**Sensing** Compare electronic and human sensing of your environment using temperature and light sensors	**Monitoring a living plant** Monitor a living plant in a plastic bag with relative humidity and light sensors	**Adaptation** Explore population, selection pressure and adaptation with a computer model
Earth & Space Science	**Weather** Observe and measure weather-related changes with temperature and relative humidity sensors	**Sun, Earth, Seasons** Connect planetary motion to day/night cycles and seasons with a light sensor	**Water cycle** Study water phase changes and relate to terrestrial phenomena with temperature and light sensors
Technology/ Engineering	**Design a playground** Study your playground and build models of several pieces of playground equipment using force and motion sensors	**Design a greenhouse** Build a working greenhouse model and monitor conditions using temperature, light, and relative humidity sensors	**Design a measurement** Choose something to measure and devise a way to do it using any or all of the sensors

Notwithstanding research studies focusing especially on higher grade levels, there is a scarcity of research and data about the efficacy of using probes and computers to teach and learn science in elementary and middle schools. What research does exist has often been done with small numbers of students. For example, a recent study of probeware (also called microcomputer-based labs, or MBLs) that reported significant positive impacts favoring the use of probes was based on a sample of only 65 fourth-grade students (Nicolaidou et al., 2007). One important goal of TEEMSS II was to conduct trials and research with larger numbers of students.

The TEEMSS pilot study was useful in that it found that handhelds and probes can be effective in inquiry learning environments at the middle school level (Metcalf & Tinker, 2003), and that online teacher professional development could effectively prepare teachers to use inquiry-based materials. Nonetheless, the pilot study did not include a comparison group. TEEMS II was funded in part to conduct additional, more rigorous research on the efficacy of using probes and computers to teach science in grades three through eight and report evidence of the effectiveness of technology-based materials.

Another major goal of the TEEMSS project was to explore ways of bridging the "digital divide" by reducing the cost to schools of implementing ICT-based instructional materials. This objective led us to examine to what extent inexpensive handheld computers could be used instead of networked desktop computers. As a result, all TEEMSS classroom units used handhelds, but the probeware and curriculum materials could be used on desktops as well. Our commitment to reducing costs also led to the development of an innovative design for low-cost probeware.

TEEMS Trials

During the period of the grant, TEEMSS units were used in classrooms during academic years 2004-2005, 2005-2006, and 2006-2007. Altogether, 66 teachers, who were located in more than a dozen school districts in three states, used one or more of the units during those years. Data about these teachers and the students they taught provide a rich source of information about both the implementation of the units and student learning outcomes. (Data reported in this paper come from a subset of the 66 teachers, as described below.)

Research Design

This paper focuses on eight TEEMSS units, those for which pre- and post-test data are available for classes using the TEEMSS materials and for comparison classes learning the same topics but without using the TEEMSS materials. (Three other TEEMSS units—the design units, which focus on the *NSES* standard for Technology/Engineering—used embedded performance assessments but did not include pre- and post-tests. For the remaining four units, no data were collected from classes that were learning the science content but without using TEEMSS materials.)

During 2004-2005, some teachers taught the topics of these eight units with the TEEMSS materials, while others taught the same topics without TEEMSS materials. In 2005-2006, all participating teachers used the TEEMSS materials. Table 2 shows the numbers of teachers and students in each of these conditions. Data about thousands of instances of students using a TEEMSS unit were collected.

A number of comparisons were made between students studying the same science topics either without using the TEEMSS materials or with the TEEMSS materials. This paper focuses on the most rigorous comparison, namely cases in which groups of teachers taught science topics without TEEMSS materials in school year 2004-2005 and then the same teachers taught those topics again using TEEMSS materials in school year 2005-2006. (These teachers are included under the 2004-2005 non-TEEMSS column in Table 2 and then again in the 2005-2006 TEEMSS column.) By focusing on groups consisting of the same teachers teaching in two different ways, one important potential threat to the validity of the findings can be eliminated because there were no differences between teachers in the experimental (TEEMSS) and the comparison (non-TEEMSS) classes.

Table 2

Teachers and Students Participating in the TEEMSS Research

		Number of teachers (number of students) used in data analysis		
		2004-2005		2005-2006
	TEEMSS unit	TEEMSS	non-TEEMSS	TEEMS
Grades 3-4	Sound	2 (38)	10 (154)	15 (245)
	Electricity	0 (0)	12 (185)	12 (173)
	Human and Electronic Sensing	7 (126)	1 (35)	13 (224)
Grades 5-6	Water and Air Temperature	5 (253)	4 (149)	6 (228)
	Levers and Machines	0 (0)	6 (120)	8 (248)
	Monitoring a Living Plant	0 (0)	6 (193)	7 (268)
Grades 7-8	Pressure	1 (30)	2 (42)	4 (120)
	Understanding Motion	3 (245)	2 (44)	4 (190)
	TOTALS *	18 (662)	43 (922)	91 (2,198)

* Teachers and students often used more than one unit in a year, and if so they are counted more than once in this table. Note that there were 20 different teachers teaching one or more units without TEEMSS in 2004-2005 and with TEEMSS in 2005-2006.

Measures of Student Learning

Items on the TEEMSS unit tests were primarily drawn from 12 existing standardized tests, including NAEP and TIMSS, as well as regional and state tests with similar item construction. SRI International, a subcontractor for this project, prepared binders that collected the more than 1,500 items initially identified as potentially relevant and organized these items by TEEMSS curricular unit. Concord Consortium staff reviewed the binders in order to determine item appropriateness for and alignment with TEEMSS curriculum. Approximately 380 items were identified as sufficiently aligned, and these were selected for use during the piloting stage of test development. Because several units still did not have adequate numbers of potential assessment items, SRI and Concord collaborated on developing unit-specific questions developed for TEEMSS2 to supplement those already collected.

The tests were piloted in November 2004. Small numbers of students were asked to "think aloud" while answering test questions in order to identify student comprehension of questions. Finally, 60-100 students from Massachusetts completed each test. Validity testing was used for multiple purposes: to select questions that were appropriate for the target grade level, to evaluate inter-rater reliability for scoring, and to compare student performance on matched pre/post variations of questions. The items that performed the best across this range of priorities were included on the final tests. (Also, a few items were included on the final forms that had not been field-tested or reviewed by the outside science expert for scientific accuracy.)

The differences between the pretest and posttest forms used in the study were minimal, limited primarily to the order of the answer choices and the presentation of slightly different surface features (e.g., changing values of temperature readings for the prompts on a multiple choice test).

SRI International was responsible for scoring student work. Scoring conformed to standard practices; namely, each rater scored a single item for all student samples before being trained, raters were then trained to score using anchor and discussion papers, raters were allowed to score actual student work only after scoring qualification samples with 80% reliability, and a minimum of 20% of student-constructed (open-ended) responses were scored by two raters whose scores were checked by a third person who resolved any discrepancies in the scores. Before the student data were analyzed, inter-rater reliability was verified for constructed response items scored by two people. The results indicated that average agreement was 74% across all units on the pretest and 76% across all units on the posttest, somewhat lower than the 80% usually desired.

Scoring was done blind. That is, the scorers did not know whether the papers they scored were pre- or post-tests, or the names of teachers or students.

Findings

Both TEEMSS and non-TEEMSS students showed statistically significant gains on all eight of the unit tests (from pre- to post-). In other words students learned science content whether taught through traditional means or using TEEMSS.

For four of the units students who used TEEMSS materials showed gains that were statistically significantly higher than students who did not use the TEEMSS materials. These data are shown in Table 3 and Figure 1. The four units were sound and electricity (grades 3-4), temperature (grades 5-6), and motion (grades 7-8). The effect sizes for the four units with statistically significant differences are 0.58 (sound), 0.94 (electricity), 1.54 (temperature), and 0.49 (motion).

For the other four units for which these comparison data are available there were no statistically significant differences between students who did and did not use TEEMSS materials. Those four unit were sensors (grades 3-4), levers & machines, and monitoring a living plant (grades 5-6), and pressure (grades 7-8).

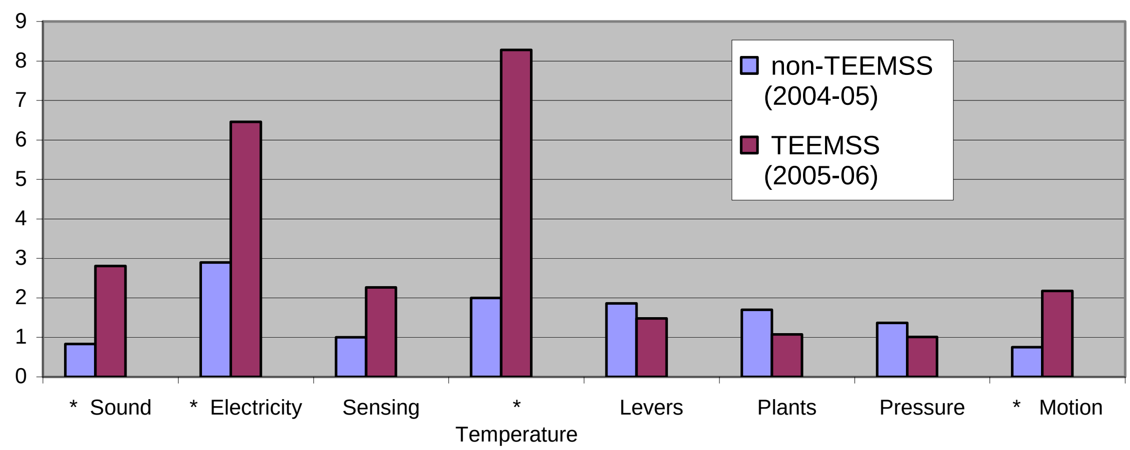
Table 3

A Comparison of Students' Gains in Two Conditions, non-TEEMSS and TEEMSS

	2004-05 non-TEEMSS		2005-06 TEEMSS			
Unit	**n**	**Gain** Mean (SD)	**n**	**Gain** Mean (SD)	**Significance**	**Effect Size**
Sound	84	0.83 (2.35)	97	2.81 (4.24)	t (152.039)=4.282, p<.001	0.58
Electricity	57	2.90 (3.75)	79	6.46 (3.82)	t (134)=5.406, p<.001	0.94
Sensing	19	1.00 (2.11)	15	2.27 (2.79)	No sig. difference	n/a
Temperature	55	2.00 (4.04)	46	8.28 (4.10)	t (99)=7.737, p<.01	1.54
Levers	104	1.86 (2.11)	112	1.48 (2.35)	No sig. difference	n/a
Plants	119	1.70 (3.21)	134	1.08 (2.81)	No sig. difference	n/a
Pressure	41	1.37 (2.27)	120	1.01(1.63)	No sig. difference	n/a
Motion	44	0.75 (2.26)	55	2.18 (3.46)	t (97)=2.369, p=.02	0.49

Figure 1

Gain Scores for the *Same* Teachers in Successive Years, non-TEEMSS vs. TEEMSS



* indicates statistically significant difference

Discussion

Computers are becoming more common in schools than ever before. Entire states are implementing laptop programs, including Maine (which has focused on middle and high schools) and Pennsylvania (which is focusing on high schools). Many other states (including Indiana, Massachusetts, Michigan, New Hampshire, New Mexico, Texas, and Vermont) and hundreds of schools and districts are implementing "one-to-one" computing programs on a smaller scale (for further information see, for example, http://ubiqcomputing.org or http://www.k12one2one.org/).

Furthermore, virtually every school in the United States is now connected to the Internet. In fact, Internet access is available not only in almost all schools but in almost all (94%) *classrooms* (Wells, Lewis, & Greene, 2006). The rapid growth of wireless networks in schools has made it much less expensive to provide ubiquitous Internet access throughout school buildings.

In this technology-rich environment, it is important to show that technology can enhance the teaching and learning of science. Data from the TEEMSS II project demonstrates that the use of computers and probes in elementary and middle school classrooms results in substantial learning gains. For some topics, the use of computers and probes results in larger learning gains than instruction without computers and probes.

Further research and development are called for. It would be good to show learning gains in additional science topics due to technology enhanced instructional units. More development work may be needed to reach that goal. Further research is also needed to see whether technology enhanced instruction can be tied to learning gains on standardized tests, such as those that will soon be required in science under provisions of the No Child Left Behind Act of 2001.

References

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

Adams, D. D., & Shrum, J. W. (1990). The effects of microcomputer-based laboratory exercises on the acquisition of line graph construction and interpretation skills by high school biology students. *Journal of Research in Science Teaching*, 27(8), 777-787.

Beichner, R. J. (1990). The effect of simultaneous motion presentation and graph generation in a kinematics lab. *Journal of Research in Science Teaching*, 27(8), 803-815.

Brassell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, 24(4), 385-395.

Friedler, Y., Nachmias, R., & Linn, M. C. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal of Research in Science Teaching*, 27(2), 173-191.

Hudson, S. B., McMahon, K. C., & Overstreet, C. M. (2002). *The 2000 national survey of science and mathematics education: Compendium of tables.* Chapel Hill, NC: Horizon Research.

Krajcik, J. S., & Layman, J. (1993). *Microcomputer-based laboratories in the science classroom. research that matters to the science teacher, no. 31*. National Association of Research on Science Teaching (NARST). (Available online at http://www.narst.org/publications/research/microcomputer.htm)

Kreikemeier, P. A., Gallagher, L., Penuel, W. R., Fujii, R., Wheaton, V., & Bakia, M. (2006). *Technology Enhanced Elementary and Middle School Science II (TEEMSS II): Research Report 1*. Menlo Park, CA: SRI International.

Laws, P. (1997). Millikan Lecture 1996: Promoting active learning based on physics education research in introductory courses. *American Journal of Physics*, 65(1).

Linn, M. C., Layman, J. W., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. C*ontemporary Educational Psychology*, 12(3), 244-253.

Metcalf, S. J., & Tinker, R. (2003). *TEEMSS: Technology Enhanced Elementary and Middle School Science*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, 24(4), 369-383.

Nicolaou, C., Nicolaidou, I., Zacharia, Z., & Constantinou, C. (2007). Fourth graders ability to interpret graphical representations through the use of microcomputer-based labs implemented within an inquiry-based activity sequence. *Journal of Computers in Mathematics and Science Teaching*, 26(1), 75-99.

National Research Council. (1995). *National Science Education Standards*. Washington, DC: National Academy of Sciences.

Thornton, R. K. (1997). Using Large-Scale Classroom Research to Study Student Conceptual Learning in Mechanics and to Develop New Approaches to Learning. In *NATO ASI Series*. New York: Springer-Verlag.

Wells, J., Lewis, L., & Greene, B. (2006). *Internet access in U.S. public schools and classrooms: 1994-2005 (Highlights)* (FRSS No. 2007-020). Washington, DC: National Center for Education Statistics.