What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Science May 2012

Technology Enhanced Elementary and Middle School Science (TEEMSS)

Program Description[1]

Technology Enhanced Elementary and Middle School Science (*TEEMSS*)[2] is a physical science curriculum for grades 3–8 that utilizes computers, sensors, and interactive models to support investigations of real-world phenomena. Through 15 inquiry-based instructional units, students interact with computers, gather and analyze data, and formulate ideas for further exploration. This information is managed by software in a handheld computer and transmitted to other students and to the teacher. All classroom units use handheld computers to avoid the expense of networked desktop computers. The program includes a web-based teacher-reporting tool that allows teachers to review student portfolios and gather student responses for assessment and class discussion.

Report Contents

Research[3]

One study of *TEEMSS* that falls within the scope of the Science review protocol meets What Works Clearinghouse (WWC) evidence standards with reservations. This study includes 181 students in grades 3–4 in elementary schools in three states.[4] Based on this study, the WWC considers the extent of evidence for *TEEMSS* on elementary school students to be small for the general science achievement domain, the only domain identified by the review protocol.

Effectiveness

TEEMSS was found to have potentially positive effects on general science achievement for elementary school students in grades 3–4.

Table 1. Summary of findings[5]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
General science achievement	Potentially positive effects	+24	+24	1	181	Small

Program Information

Background

Technology Enhanced Elementary and Middle School Science was developed by the Concord Consortium and funded by the National Science Foundation. Address: 25 Love Lane, Concord, MA 01742. Email: info@concord.org. Web: http://www.concord.org. Telephone: (978) 405-3200, (617) 926-0329. Fax: (978) 405-2076.

Program details

The *TEEMSS* curriculum emphasizes the use of technology to support inquiry-based scientific learning. The curriculum incorporates the use of computers or probeware, such as sound graphers, thermometers, and sensors.

TEEMSS includes 15 inquiry-based instructional science units, with five units developed for each of the grade levels 3–4, 5–6, and 7–8. Each set of five units shown below targets the five National Science Education Standards: Inquiry, Physical Science, Life Science, Earth and Space Science, and Technology and Design.

Grades 3–4 units:

- Sound
- Electricity
- Sensing
- Weather
- Design a playground

Grades 5–6 units:

- Water and air temperature
- Levers and machines
- Monitoring a living plant
- Sun, Earth, seasons
- Design a greenhouse

Grades 7–8 units:

- Air pressure
- Motion
- Adaptation
- Water cycle
- Design a measurement

Every unit contains two one-week investigations, each with a discovery question, several trials, analysis, and ideas for further investigations. Each investigation is structured to give guidance to the students as they progress through it. The students interact with the computer for information, data gathering and analysis, and response purposes. This information is managed by the software in the handheld computer and can be transmitted to other students and to the teacher. Thus, the student is led through concept and content development, instruction in the use of probes, and data gathering and graph representations of key concepts.

All *TEEMSS* classroom units use handheld computers to avoid the cost of using networked desktop computers, but the probeware and curriculum materials can be used on desktops as well. Teachers' guides are included in the program, featuring discussion guides, background material on the content, ideas for assessments, information on the technology, and suggested timelines.

Cost

The *TEEMSS* curriculum, including the 15 units and software, is available to download and use free of charge at http://teemss.concord.org/. Cost information for the teacher's guides and other project materials, such as mobile devices (handhelds) and probeware, is available from the developer.

Research Summary

Three studies reviewed by the WWC investigated the effects of *TEEMSS* on elementary school students. One study (Zucker, Tinker, Staudt, Mansfield, & Metcalf, 2008) is a quasi-experimental design that meets WWC evidence standards with reservations. This study is summarized in this report. The remaining two studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 6 for citations for all three studies.)

Table 2. Scope of reviewed research

Grade	3, 4
Delivery method	Whole class
Program type	Curriculum
Studies reviewed	3
Meets WWC standards	0 studies
Meets WWC standards with reservations	1 study

Summary of studies meeting WWC evidence standards without reservations

No studies of *TEEMSS* meet WWC evidence standards without reservations.

Summary of studies meeting WWC evidence standards with reservations

Zucker et al. (2008) conducted a quasi-experimental study that examined the effects of *TEEMSS* on students in grades 3–8 attending elementary and middle schools in three states. Of the 15-unit *TEEMSS* curriculum, the study examined eight science units, including three units for grades 3–4 (sound, electricity, and sensing), three units for grades 5–6 (water and air temperature, levers and machines, and monitoring a living plant), and two units for grades 7–8 (air pressure and motion).

The same teachers taught both groups of students. Students in the treatment group were taught with the *TEEMSS* curriculum during the 2005–06 school year. Students in the comparison group were taught during the 2004–05 school year using the teachers' regular teaching methods. The study reported students' outcomes after completion of the teaching of the eight units. The only findings that meet WWC evidence standards with reservations were those for the sound unit test. Findings for the remaining seven outcomes do not meet WWC standards.[6] The WWC based its effectiveness rating on findings from the treatment group of 97 students and comparison group of 84 students from grades 3–4 who received instruction on the topic of sound.[7]

Effectiveness Summary

The WWC review of interventions for Science addresses student outcomes in one domain: general science achievement, which includes three outcome constructs: life science, earth/space science, and physical science. The study that contributes to the effectiveness rating in this report covers one construct: physical science. The findings below present the WWC-calculated estimates of the size and statistical significance of the effects of *TEEMSS* on elementary school students. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 11.

Summary of effectiveness for the general science achievement domain

One study reported findings in the general science achievement domain.

Zucker et al. (2008) reported, and the WWC confirmed, statistically significant positive effects on the *TEEMSS* sound unit test for students in grades 3–4.

Thus, for the general science achievement domain, one study showed statistically significant positive effects. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the general science achievement domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	The review of *TEEMSS* in the general science achievement domain had one study showing statistically significant positive effects and no studies showing statistically significant or substantively important negative effects.
Extent of evidence	**Criteria met**
Small	The review of *TEEMSS* in the general science achievement domain is based on one study that included 181 students.

References

Study that meets WWC evidence standards with reservations

Zucker, A. A., Tinker, R., Staudt, C., Mansfield, A., & Metcalf, S. (2008). Learning science in grades 3–8 using probeware and computers: Findings from the TEEMSS II project. *Journal of Science Education and Technology, 17*(1), 42–48.

Additional source:

Zucker, A. A., Tinker, R., Staudt, C., Mansfield, A., & Metcalf, S. (2007, April). *Increasing science learning in grades 3–8 using computers and probes: Findings from the TEEMSS II project*. Proceedings of the National Association for Research in Science Teaching 2007 Annual Meeting, New Orleans, LA.

Study that does not meet WWC evidence standards

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: Center for Technology in Learning, SRI International. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Study that is ineligible for review using the Science Evidence Review Protocol

Metcalf, S. J., & Tinker, R. F. (2004). Probeware and handhelds in elementary and middle school science. *Journal of Science Education & Technology, 13*(1), 43–49. The study is ineligible for review because it does not use a comparison group design or single-case design.

Appendix A: Research details for Zucker et al. (2008)

Zucker, A. A., Tinker, R., Staudt, C., Mansfield, A., & Metcalf, S. (2008). Learning science in grades 3–8 using probeware and computers: Findings from the TEEMSS II project. *Journal of Science Education and Technology, 17*(1), 42–48.

Table A. Summary of findings

Meets WWC evidence standards with reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
General science achievement	181 students	+24	Yes

Setting

The study took place in more than 100 elementary and middle school classrooms in over a dozen school districts in three states during the 2004–05 and 2005–06 school years.

Study sample

In this quasi-experimental study, the treatment group included students of teachers who used the *TEEMSS* curriculum during the 2005–06 school year. The comparison group included students from the prior school year (2004–05) of the same teachers, who taught the same topics but used their regular teaching methods. For this review, the analysis sample consisted of 181 students in grades 3–4 (97 treatment and 84 comparison) who received instruction on the topic of sound.

Intervention group

The curriculum received by the treatment group was *Technology Enhanced Elementary and Middle School Science* (*TEEMSS*) for grades 3–8. The curriculum included 15 units that were customized to grade levels. Of the 15-unit *TEEMSS* curriculum, the study examined eight science units, including three units for grades 3–4 (sound, electricity, and sensing), three units for grades 5–6 (water and air temperature, levers and machines, and monitoring a living plant), and two units for grades 7–8 (air pressure and motion). Among the eight unit outcomes, the only findings that met WWC evidence standards with reservations were those for the sound unit test for grades 3–4. The unit contained two one-week investigations of sound and vibrations with the sound grapher, a software program that is used with a microphone to record the pattern of sound vibrations, and included a discovery question, several trials, analysis, and ideas for further investigations.

Comparison group

The comparison group included students who were taught the same science unit topics using current teaching practices. The authors indicated to the WWC that there was no single comparison curriculum, and the comparison group curricula addressed science education standards. Authors did not state if the comparison curricula were inquiry-based or used technology.

Outcomes and measurement

For the pretest and posttest, students completed the sound unit test. The pretest was given to students before the teacher taught the unit, and the posttest was given upon the completion of the teaching of the unit. The posttest differed slightly from the pretest in the order of the response options and the values of the prompts (e.g., temperature) in the questions. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers had access to an online training course that provided information about the *TEEMSS* curriculum and technology. The study did not discuss additional support or training for teachers.

Appendix B: Outcome measures for each domain

Comprehension	
Physical science construct	
Sound unit test	The sound unit test was developed by the researchers and designed to align with the *TEEMSS* curriculum and science education standards.[8] The eight unit tests used in the study included items from 12 standardized tests, including the National Assessment of Educational Progress (NAEP) and Trends in International Mathematics and Science Study (TIMSS), as well as regional and state tests. The sound unit test consists of nine items (three multiple-choice and six constructed-response items). The total number of points possible for this test was 21. Interrater reliability was 74% across the pretests and 76% on the posttest (as cited in Kreikemeier et al., 2006; Zucker et al., 2008; author response to WWC, 2011).

Appendix C: Findings included in the rating for the general science achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Zucker et al., 2008[a]								
Sound unit test	Grades 3–4	181	14.78 (3.38)	12.81 (2.56)	1.97	0.65	+24	0.03
Domain average for general science achievement (Zucker et al., 2008)							**+24**	**Statistically significant**

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The statistical significance of the study's domain average was determined by the WWC; a study is characterized as having a statistically significant positive effect when univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

[a] For Zucker et al. (2008), no corrections for clustering or multiple comparisons were needed. The reported *p*-value was computed by the WWC. The WWC calculated the program-group mean using a difference-in-differences approach (see the WWC Procedures and Standards Handbook, Appendix B) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. The sample sizes, group means, and posttest standard deviations presented in this table were based on information provided by the study author.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: the program's website (http://teemss.concord.org/project/, downloaded October 2011) and Zucker et al. (2008). The WWC requests that developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in October 2011, and we incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by June 2011.

[2] *Technology Enhanced Elementary and Middle School Science (TEEMSS)* encompasses two sequential projects funded by the National Science Foundation: a *TEEMSS* pilot project and *TEEMSS II*.

[3] The studies in this report were reviewed using WWC Evidence Standards, version 2.1, as described in the Science review protocol, version 2.0. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[4] The study analysis sample for Zucker et al. (2008) included 1,181 students in grades 3–8 from elementary and middle schools. However, only findings for the sound unit test, which was administered to 181 elementary students in grades 3 and 4, meet WWC evidence standards with reservations.

[5] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 11. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[6] Findings for the electricity and pressure outcomes were not included because the difference between the intervention and comparison groups on the baseline measure exceeded the WWC's allowable difference. Findings for the remaining five outcomes (human and electronic sensing, water and air temperature, levers and machines, monitoring a living plant, and understanding motion) were not included because, according to WWC standards, the difference found between treatment and comparison groups on the baseline measures required statistical adjustment, but these statistical adjustments were not made. Please refer to the WWC Procedures and Standards Handbook, version 2.1, section III for information about evidence standards.

[7] The sample size data and corresponding statistics for the sound unit outcome came from an author response to questions from the WWC.

[8] Based on the author's response, the unit test and the comparison curricula address the same science education standards.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, May). *Science intervention report: Technology Enhanced Elementary and Middle School Science (TEEMSS)*. Retrieved from http://whatworks.ed.gov.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If treatment assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 11.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to treatment and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into treatment and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 11.

Single-case design A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.