Multi-Year Large-Scale Field Studies of the Fundamental Skills Training Project's Intelligent Tutoring Systems.

Thomas N. Meyer [1], Kurt Steuck [2], Todd M. Miller [1], Monika Kretschmer [1]

[1]Mei Technology
A Division of MATCOM
8930 Fourwinds Dr #450
San Antonio, TX 78249 USA

[2]Air Force Research Laboratory
2509 Kennedy Circle (Bldg 125)
Brooks AFB, TX 78235-5118

Project Background

Since 1990, the Air Force has been engaged in a long-term research effort, the Fundamental Skills Training (FST) project, designed to bring state-of-the-art intelligent tutoring technology to bear on literacy skills problems in mathematics, writing, and science.

The FST project was designed to help school children attain basic literacy skills. To accomplish this goal, three intelligent tutoring systems (ITS), the Word Problem Solving Tutor (WPS), MAESTRO, the Writing Process Tutor, and the science tutor, Instruction in Scientific Inquiry Skills (ISIS) were developed. Throughout the ten years of the project, these systems have been evaluated in 15 year long field studies involving 40 schools in 10 states. In addition, these studies have involved 40-50 teachers and as many as nearly 3,000 students each year to evaluate the effectiveness of the software in enhancing critical thinking skills.

The goal of this paper is to present to researchers, teachers, administrators, and managers a quick summary of the methods and results of the research conducted on all three tutors from September 1991 through May 1999.

The Suite of FST Intelligent Tutoring Systems

The first tutor developed by the FST team was the Word Problem Solving tutor (WPS). This tutor teaches individuals general problem solving strategies as well as how to solve specific word problems. Specifically, WPS helps students learn to analyze and solve word problems by applying a pedagogy derived from contemporary cognitive science, including principles of active problem solving (Anderson, 1994), elaboration theory (Riegeluth, 1987, 1992), categorization by prototype (Rosch, 1988a, 1988b), mastery learning (Bloom, 1984), and worked examples (Ward & Sweller, 1990). The tutor consists of 24 independent modules of instruction that correspond to critical curriculum elements in ninth grade algebra. WPS is designed to enhance, not replace, traditional classroom instruction. During the course of traditional instruction, the class periodically goes to a computer classroom for students to work individually and receive automated feedback and guidance from the tutor. Because WPS does the majority of teaching there, the teacher is free to work individually with students who are having a particularly difficult time learning.

Paper Presented at the Annual Meeting of the American
Educational Research Association, April 2000, New Orleans.

The second FST tutor, MAESTRO: The Writing Process Tutor, was created to facilitate the development of basic writing process skills. *MAESTRO* is a student-oriented writing process tutor that facilitates learning about the writing process through a series of instructional presentations and real-life writing assignments performed in composing workspaces that, when properly used, help the student acquire the writing processes of an expert writer. *MAESTRO* allows the student to control the workspaces through a simple menu system that facilitates the development of cognitive processes often associated with expert-level writing. During use of the *MAESTRO* composing workspaces, the student's writing process is monitored, and the student is coached by an intelligent advise system (Rowley, 1995; Rowley, Miller, & Carlson, 1997; Rowley, K. & Crevoisier, 1997).

The last ITS created in the FST project was the Instruction in Scientific Inquiry Skills (ISIS) tutor. ISIS is a simulation-based cognitive tutoring system constructed to teach high school students scientific inquiry skills and substantive knowledge in ecology. ISIS has been designed to allow students to explore an ecology domain space and practice skills. Students using ISIS are taught under an instructional system that provides an exploratory environment, authentic data, appropriate and timely feedback, and a series of progressively difficult Skill Instructional Modules (SIMs). The primary goal of the science tutor, consistent with major science initiatives and standards, is to increase the level of scientific functioning (Steuck & Miller, 1997; Meyer, Miller, Steuck, & Kretschmer, 1999; Meyer, Steuck, Miller, Pesthy, & Redmon, 1999; Meyer, Steuck, Miller, Pesthy, & Kretschmer, 1999).

Research

In the studies the focus of the research has been on three sets of questions including the instructional effectiveness of the tutors (i.e. "Do students gain more using the tutors vs. traditional educational approaches?"), the effectiveness of individual tutor components, and implementation issues in the use of educational technology. This paper primarily focuses on the instructional effectiveness of the tutors.

Subjects

In general, the subjects were seventh, ninth, and tenth grade students enrolled in sections of introductory biology, life science, pre-algebra and algebra, or English classes at forty junior and senior high schools in states across the nation. Students were demographically diverse (e.g., inner city youth, suburban youth, varying levels of SES, gender, etc.) and participated in this research as part of their normal instruction in biology or life science classes. On average, classes spent 15-20 hours in the computer lab throughout the school year.

Methods

Since random assignment was not feasible, (i.e., the schools having designated students to classes and teachers) the studies employed a quasi-experimental design.

In each of the studies, to determine instructional effectiveness, students devoted a partial class period to pretesting during the first six weeks of the academic year. After taking the pretest, the experimental students attended a computer lab approximately one day every two weeks during their normal class time for an average of 15-20 contact hours during the academic year. Control group subjects returned to their normal classes where instructors covered the same material. At the end of the academic year, students were given a posttest measure to quantify instructional gains.

Results

1. Word Problem Solving (WPS) Tutor Pilot Study. Sep 91 - May 92

Description: A pilot test was conducted to investigate the instructional effectiveness of the first version of WPS and to examine issues surrounding a long-term operational implementation of WPS in a high school setting. The treatment group consisted of 400 students in 16 sections of prealgebra with 8 teachers for approximately 20 hours of instruction. A demographically similar high school served as the non-treatment control group.

Outcome: The tutored group significantly outperformed the control group on structured word problem solving exercises. In addition, their knowledge of the steps involved in solving word problems was significantly better than the control group. Several issues concerning implementation were raised, such as teacher-student-computer interaction and technical support requirements. Data from this study were applied to subsequent tutor design and implementation.

2. Word Problem Solving Tutor Large-scale Implementation Study. Oct 92 - May 93

Description: A large-scale evaluation of WPS was conducted at 7 schools in 3 states across the nation. The treatment group used the WPS software plus classroom instruction, the placebo group used the same set of computerized word problems but with no theory-driven instructional approach plus classroom instruction, and the control group used traditional classroom instruction only. The study involved 632 students and 20 teachers (Wheeler & Regian, 1999).

Outcome: Students who used WPS outperformed students who used the placebo tutor and students who received traditional classroom only. The treatment group improved by 29%, while the placebo group improved by 18% and the control group improved by 19%. Further implementation issues dealing with site communication were raised and solved. In addition, differences in improvement between students' concrete reasoning skills and abstract reasoning skills were examined. In all three groups, there were larger increases in concrete reasoning performance than in abstract reasoning performance (see Table 1). The differences in change scores among the groups on both concrete and abstract reasoning scores were significant.

Groups	Percent change on concrete reasoning scores	Percent change on abstract reasoning scores
WPS	31%	20%
Placebo	19%	15%
Control	22%	11%

Table 1. Percent Changes in Concrete Reasoning Skills and Abstract Reasoning Skills.

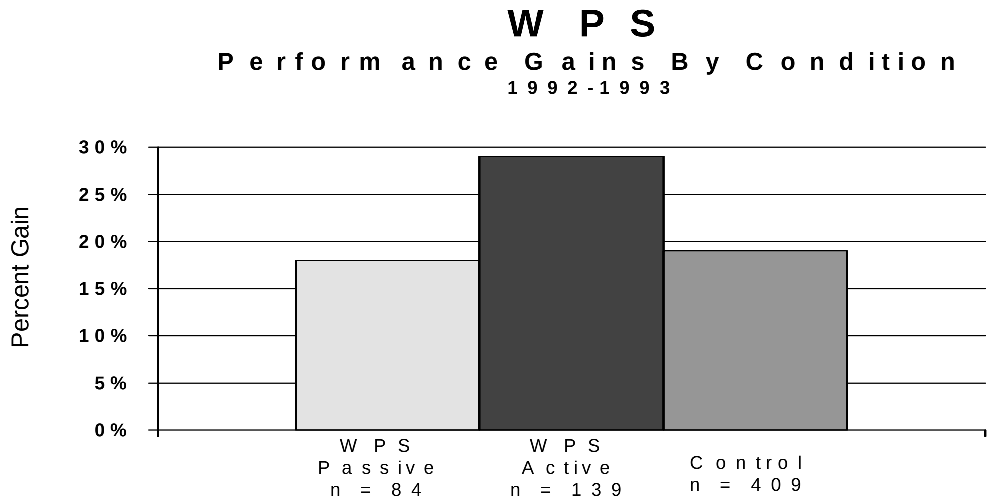

Figure 1. 1992- 1993WPS performance gains by condition

3. Reading and Writing in a Supportive Environment (R-WISE) Pilot Study. Jan 93- May 93

Description: The pilot test was conducted to investigate the effectiveness of the first version of R-WISE and issues surrounding a long-term operational implementation of it in a high school setting. The treatment group consisted of 430 students in 26 sections of ninth grade English, 9 teachers, approximately 9 hours of instruction. Another demographically similar high school group (423 students, 22 sections, 6 teachers) in San Antonio served as the non-treatment control (Rowley, Miller, & Carlson, 1997).

Outcome: The treatment group had higher scores on 5 measures of writing performance even after only 9 contact hours with R-WISE. Implementation issues, such as scheduling classes into the computer laboratories, were raised and resolved.

	Cntrl	Trmnt
Pretest Score	34%	30%
Posttest Gain	-46%	44%

Table 2. R-WISE Holistic Pretest Scores and Posttest Results.

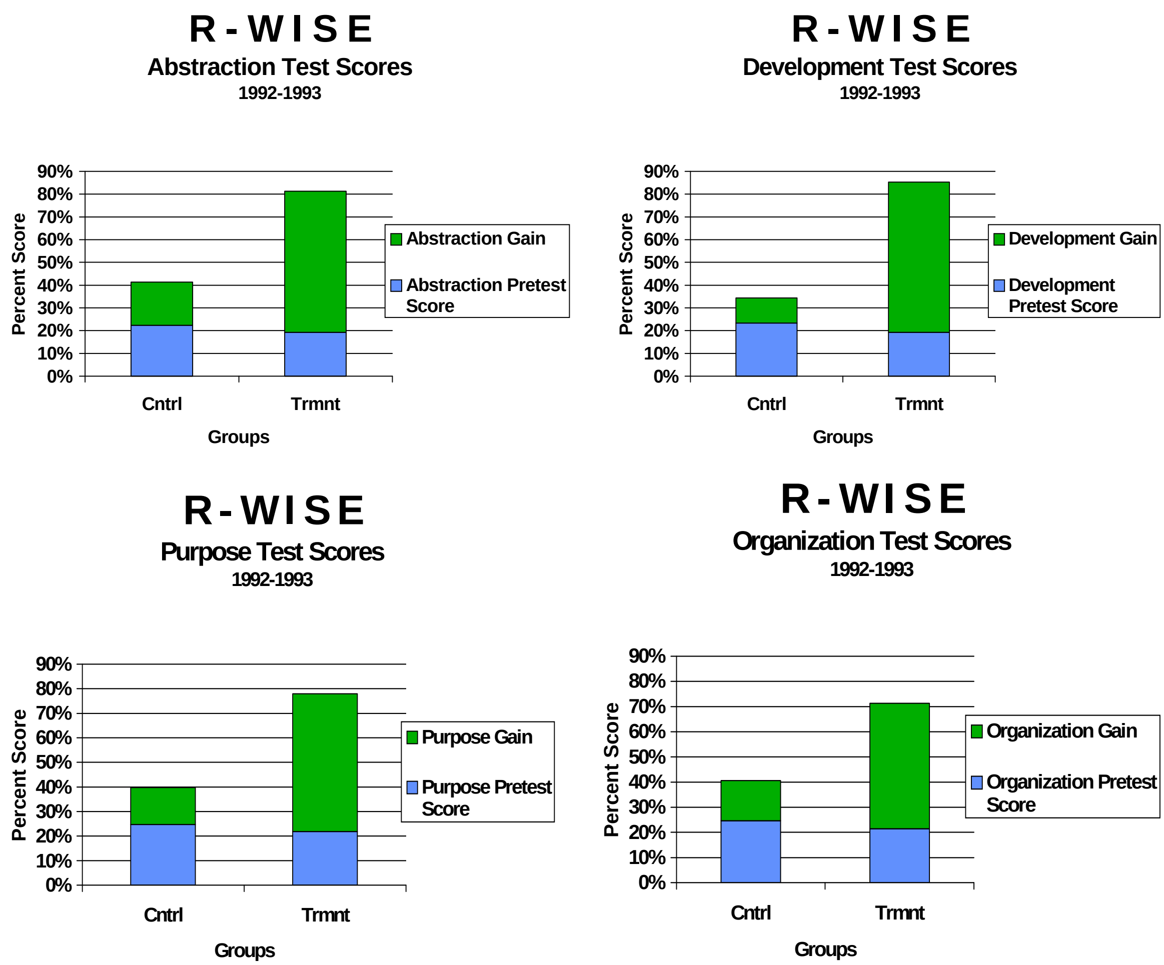

Figure 2. 1992- 1993 R-WISE test scores.

4. A Test of Two WPS Instructional Approaches. Sep 93 - May 94

Description: A study of two instructional approaches (guided and unguided) was conducted at two schools in Dayton, OH involving 194 students and 5 teachers.

Outcome: A significant aptitude-treatment interaction was found. Students with lower levels of achievement as determined by their class level performed better in the guided mode of instruction than those in the unguided mode. Students with moderate to high achievement levels performed better in the unguided mode than the guided mode.

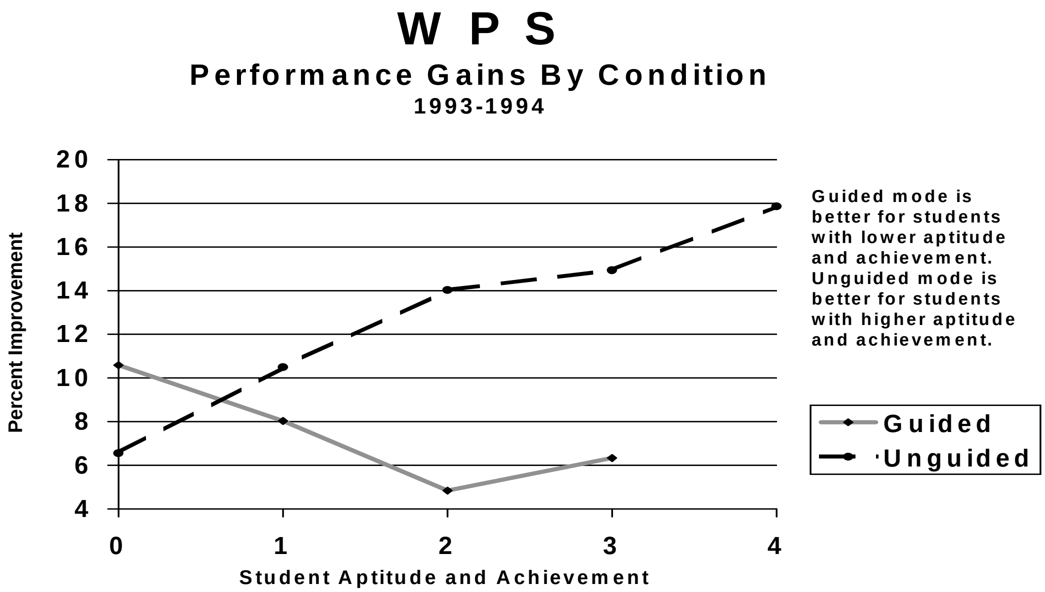

Figure 3. 1993-1994 WPS performance gains by condition

5. Reading and Writing in a Supportive Environment (R-WISE) Large-scale Implementation Study. Sep 93 - May 94

Description: A large-scale operational test of R-WISE was conducted in 8 schools in 5 states across the nation. The treatment group (N=700 students, 17 teachers, 57 sections) used R-WISE for an average of 20 hours during the academic year. The control group consisted of 451 students and 13 teachers (37 sections) from the same high schools. The control group used a word processor for an average of 20 hours during the academic year (Rowley, Miller, & Carlson, 1997).

Outcome: The students using R-WISE significantly outperformed students learning to write using a word processor on all 5 measures of writing performance. The attitudes towards writing of students using R-WISE were significantly more positive than those using a word processor.

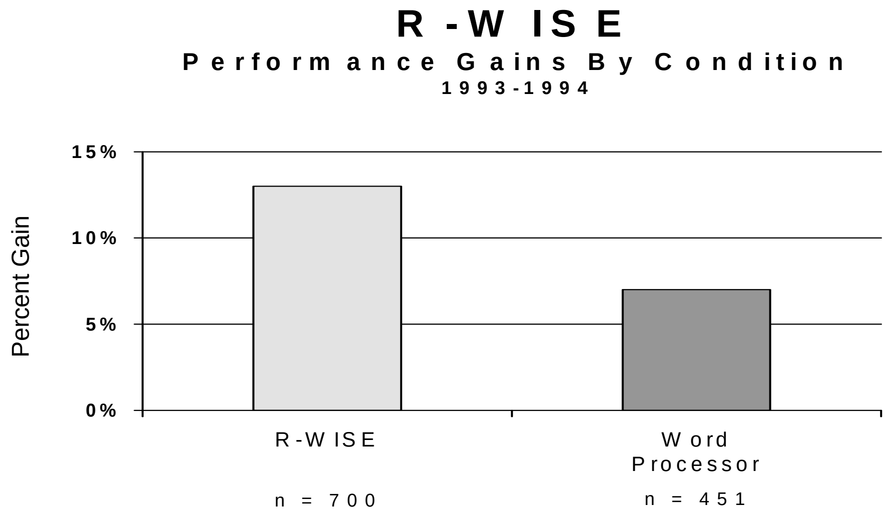

Figure 4. 1993-1994 R-Wise performance gains by condition

6. Extended Operational Test of the Word Problem Solving Tutor. Sep 94 - May 95

Description: A large-scale field test of a revised version of WPS (v4.06) was conducted with 611 subjects at 8 schools in 2 states. In order to test the effects on teacher training on WPS implementation, training was shortened to 1 1/2 days from 2 1/2 days.

Outcome: There was no control group during this year. Students improved their word problem-solving performance by 12%. Shortening teacher training did not appear to negatively affect teacher and student use of WPS.

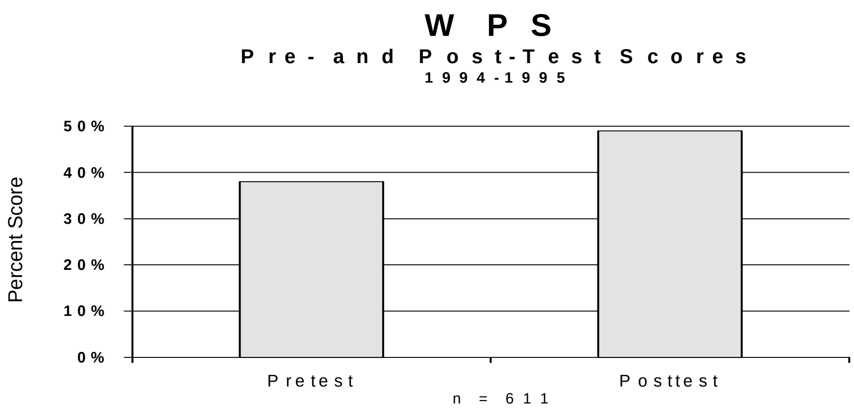

Figure 5. 1994-1995 WPS pre- and post-test scores

7. A Comparison of Two Instructional Approaches for R-WISE. Sep 94 - May 95

Description: Data were collected from students using two versions of R-WISE. One version had extended, elaborated instructional modules which the students were required to complete as they interacted with R-WISE. The other version included shorter, simpler instructional modules. The relationship between teacher style and tutor approach was investigated (Rowley, Miller, & Carlson, 1997).

Outcome: Data from the first semester of the academic year showed positive gains from September through January, though these gains were smaller than previous studies, probably because student contact time with R-WISE was significantly reduced. Due to a network installation "bug" in the updated R-WISE software, teachers used the computer lab less in the second semester. As a result, the data from the spring semester are inconclusive. The data on teacher style interactions indicate that styles of teaching with emphasis on student-centered instruction are more compatible with the tutor produce higher student scores. However, further research is needed before strong conclusions are made concerning the relationship of teacher style and tutoring approach.

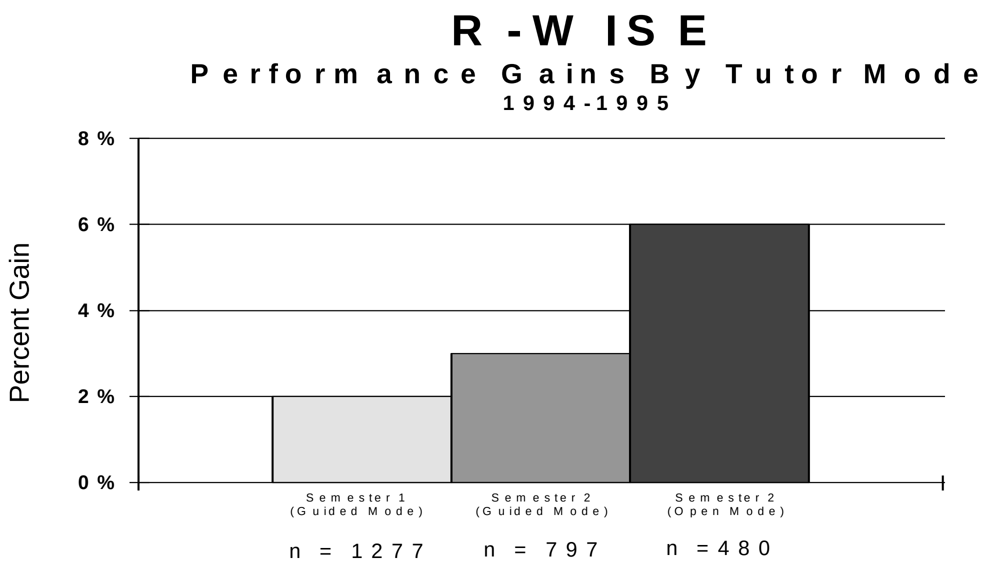

Figure 6. 1994-1995 R-Wise performance gains by tutor mode

8. Instruction in Scientific Inquiry Skills Tutor (ISIS) Pilot Study. Oct 94 - May 95

Description: A pilot test was conducted to investigate the effectiveness of the first version of ISIS and issues surrounding a long-term operational implementation of ISIS in a high school setting. The treatment group consisted of approximately 400 students in 16 sections of ninth grade biology with 4 teachers for approximately 15 hours of instruction. Ten other sections of freshman biology at the same high school served as the control sample.

Outcome: Students using ISIS gained more scientific inquiry skills during the academic year than their control counterparts. Qualitative data from teacher-student-computer interactions indicated several directions for refinements for the subsequent version of ISIS.

9. Small School Implementation of WPS. Jan 96- May 96

Description: WPS 4.06 was implemented in 10 rural sites in New Mexico. Teacher training was shortened to one day in order to test the effects of reduced teacher training sessions on the implementation of WPS. The instructional effectiveness of an updated version of WPS in those schools was tested.

Outcome: Students showed gains comparable to previous years. Students improved their word problem-solving performance by 6%. The project staff met with site coordinators from the rural sites to collect information on the implementation of WPS in their schools. Specifically, the information will be used to assist users in remote sites with only one math teacher using WPS Several site coordinators reported that their teachers sometimes needed to discuss questions with other teachers.

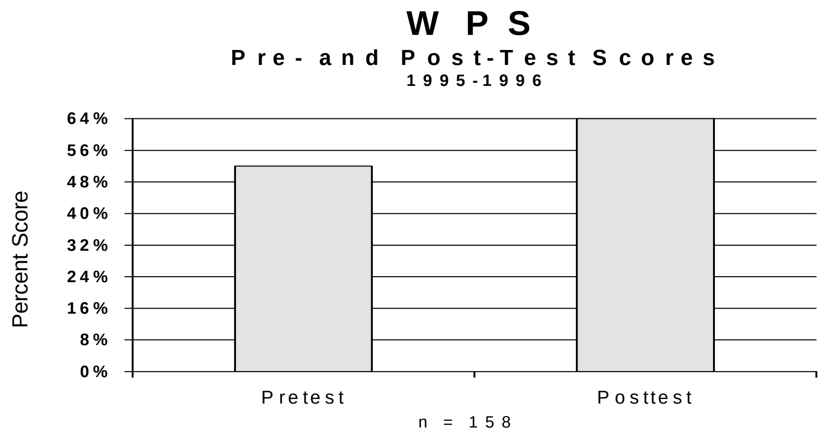

Figure 7. 1995-1996 WPS pre- and post-test scores

10. Addressing Replicability and Measurement Issues in R-WISE studies. Sep 95 - May 96

Description: Data were collected from students using the previous year's version of R-WISE in a replication study. Research design also tested for equivalence of two presumably parallel writing prompts. The study included 356 students of 8 English teachers in the treatment group and 261 students of 5 English teachers in the control group, with 39 English classes participating (Rowley, Miller, & Carlson, 1997).

Outcome: Findings supported the results of previous studies. Results indicated a small but significant effect of the R-WISE condition on learning outcomes. Results also raised questions as to the reliability of the two prompts certified as equivalent. Field input regarding lessons learned suggested directions for the design of a follow-on writing process tutor, *MAESTRO*.

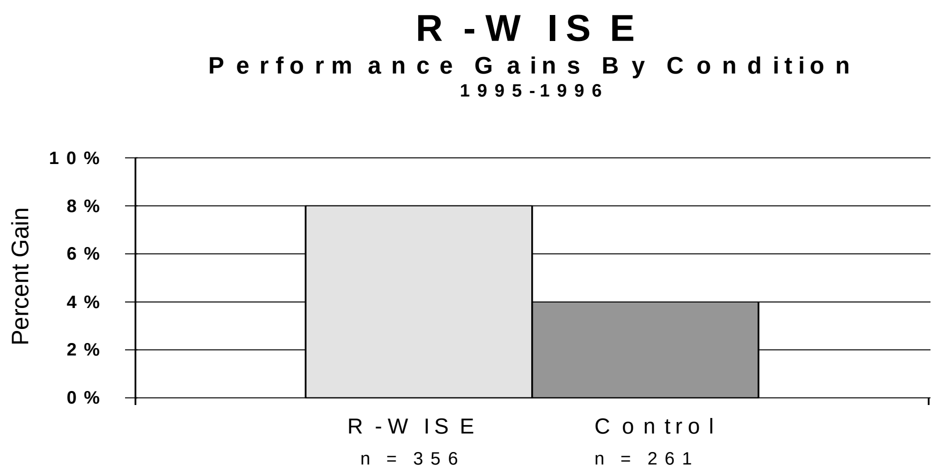

Figure 8. 1995-1996 R-Wise performance gains by condition

11. Instruction in Scientific Inquiry Skills (ISIS) Tutor Large-scale Implementation Study. Sep 95 - May 96

Description: A large-scale test of the effectiveness and implementation of ISIS was conducted in 14 schools in 5 states. There were over 30 teachers using ISIS in 83 sections of ninth-grade biology. There were 46 sections in the same high schools serving as non-treatment control groups. The study also investigated the feasibility of using ISIS in 7th grade honors biology and explored several research questions concerning the use of concept mapping in a knowledge-rich domain (Steuck & Miller, 1997).

Outcome: Observing the overall gains, the treatment group improved by 8%, while the control group improved by only 4%. Specifically, the amount of gain was directly influenced by the skill level obtained by the students. Students using ISIS (having met a threshold of assignments completed) outperformed control group students on scientific inquiry skills as measures of domain knowledge. Teacher and student focus groups were used to improve the design and implementation of ISIS.

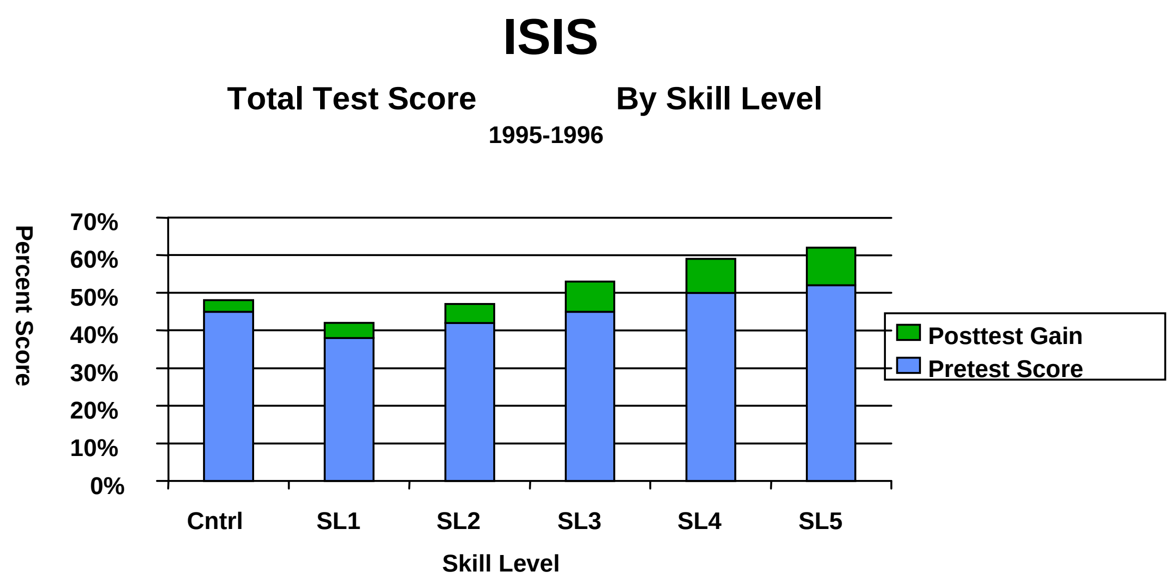

Figure 9. 1995-1996 ISIS pre-test scores and posttest gains by skill level

The scores of the design experiment subscale demonstrate the fact that students who used ISIS and obtained higher skill levels outperformed control group students on scientific inquiry skills as measures of domain knowledge.

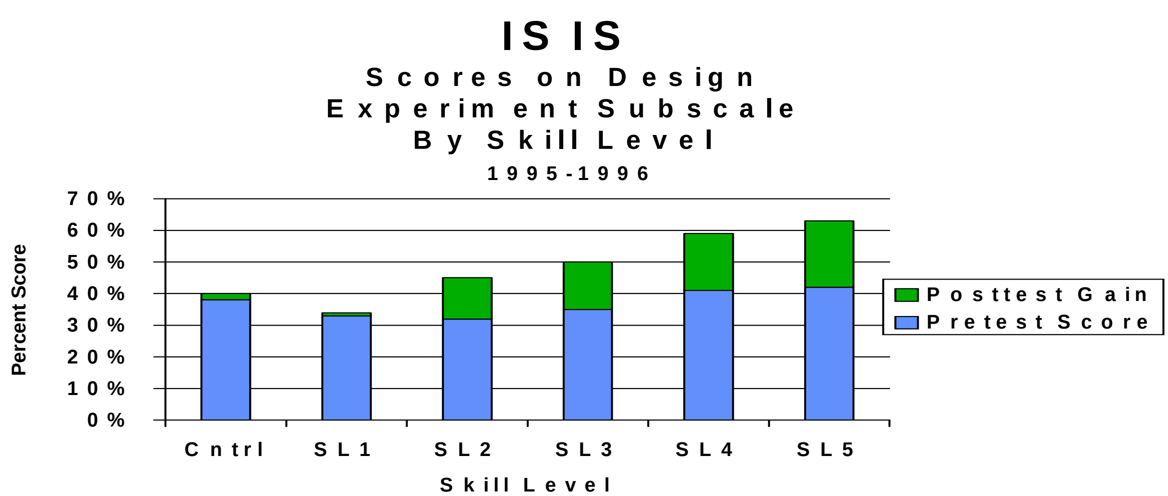

Figure 10. 1995-1996 ISIS design experiment test scores by skill level

12. A Study of the Effectiveness of the Commercial Version of WPS. Sep 96 - May 97

Description: The commercial version of WPS was fielded in 2 middle schools in San Antonio and 3 high schools in New Mexico in order to examine the effectiveness of WPS with younger student populations and to replicate previous results.

Outcome: Students showed gains comparable to previous years. Overall, students (n=244) improved their word problem-solving performance by 9%. Specifically, upward gains for the middle schools (n=63) were 14% while the high schools (n=181) showed an 11% upward gain.

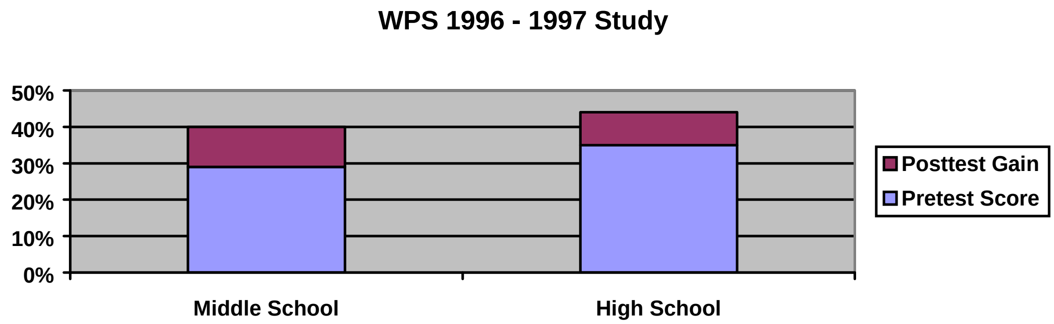

Figure 11. 1996-1997 WPS pretest scores and posttest gains

13. MAESTRO The Writing Process Tutor Initial Functionality Study. Sep 96 - May 97

Description: MAESTRO was designed as a follow-on tutor to R-WISE, based on results of the R-WISE studies, teacher input, cognitive research into the writing process, and the cognitive apprenticeship instructional strategy. The functionality and efficacy of the initial version of MAESTRO was tested in a large-scale pilot study. The study groups include classes of 54 teachers at 23 schools.

Outcome: Similar to results with the other tutors in the FST suite, student gains were directly influenced by the amount of time spent on the Writing Process Tutor. Specifically, students spending at least 11 hours using Maestro improved by 11% while the control group improved by only 3% (Steuck, Rowley, & Kretschmer, 1999).

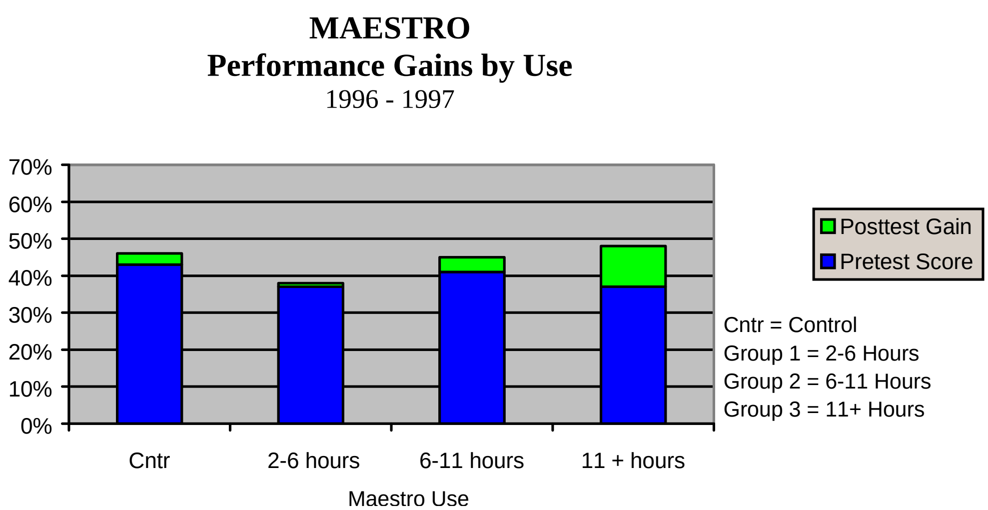

Figure 12. 1996-1997 MAESTRO performance gains by use

14. ISIS: A Replication Study. Sep 96 - May 97

Description: The purpose of this study is to replicate previous findings with the newest version of the software in 19 schools with over 30 teachers utilizing ISIS (Meyer, Steuck, Miller, Pesthy, & Kretschmer, 1999).

Outcome: Observing the overall gains, the treatment group improved by 11%, while the control group improved by only 6%. Specifically, the amount of gain was directly influenced by the skill level obtained by the students. Once again, students using ISIS (having met a threshold of assignments completed) outperformed control group students on scientific inquiry skills as measures of domain knowledge. Teacher and student focus groups were again used to improve the design and implementation of the next iteration of the tutor.

ISIS
Total Test Score By Skill Level
1996 - 1997

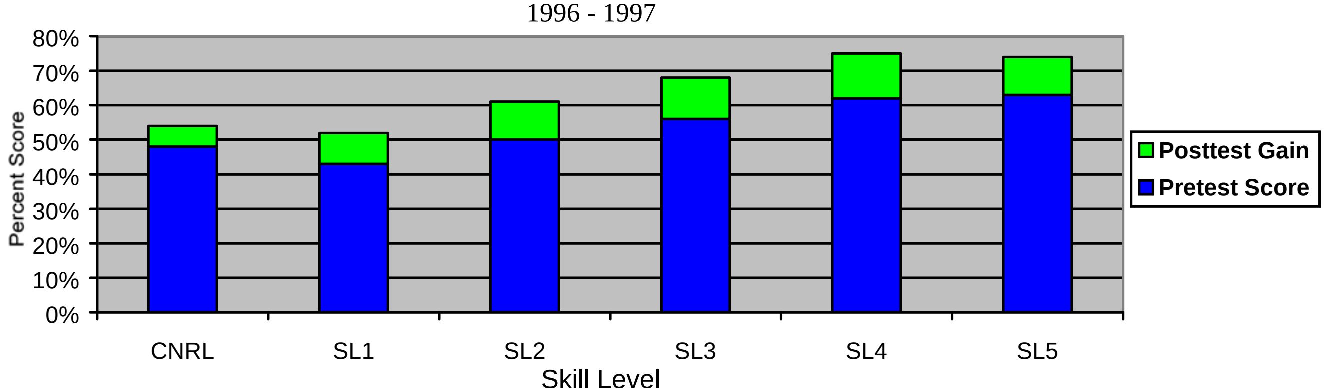

Figure 13. 1996-1997 ISIS pre-test scores and posttest gains by skill level

As seen in the past study, the design subscale showed that students who used ISIS and obtained higher skill levels outperformed control group students on scientific inquiry skills as measures of domain knowledge.

ISIS
Scores on Design Experiment Subscale
By Skill Level
1996-1997

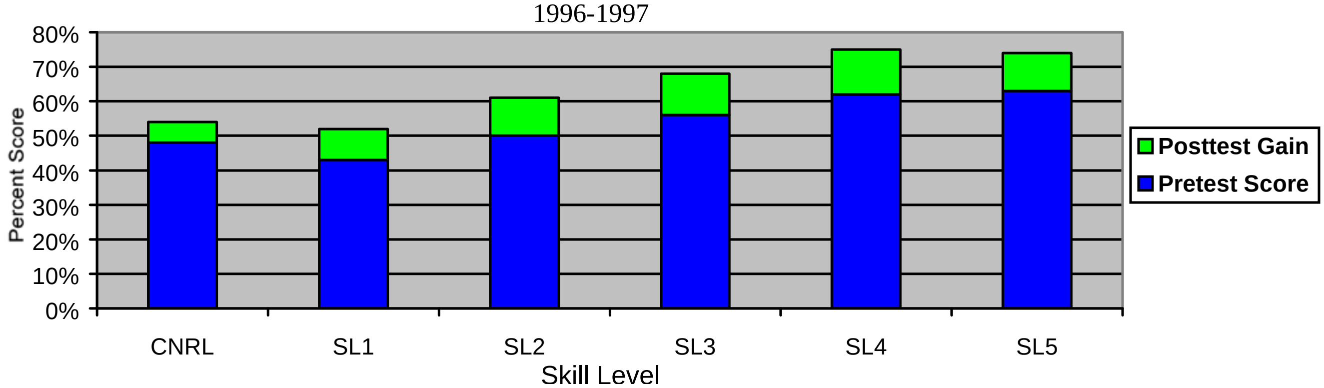

Figure 14. 1996-1997 ISIS design experiment pre-test scores and posttest gains by skill level

15. ISIS: A Replication Study. Sep 97 - May 98

Description: The purpose of this study is to replicate previous findings with the newest version of the software in 4 schools with 9 teachers utilizing ISIS (Meyer , Miller, Steuck, & Kretschmer,1999).

Outcome: The results again depicted a linear trend with students who completed more assignments showing larger gains on both a Skills and Content Test.

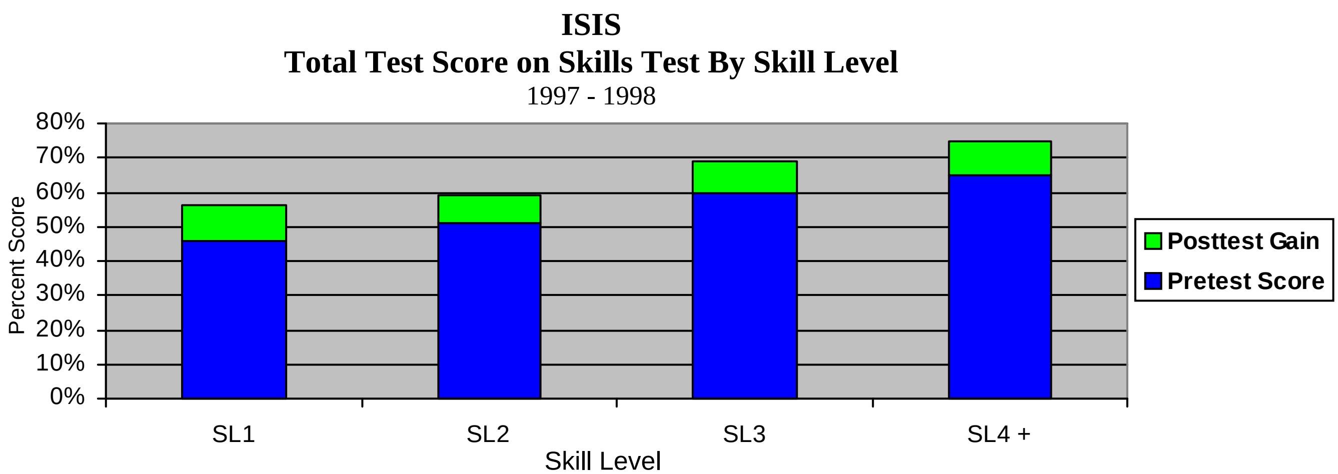

Figure 15. 1997-1998 ISIS skills test pre-test scores and posttest gains by skill level

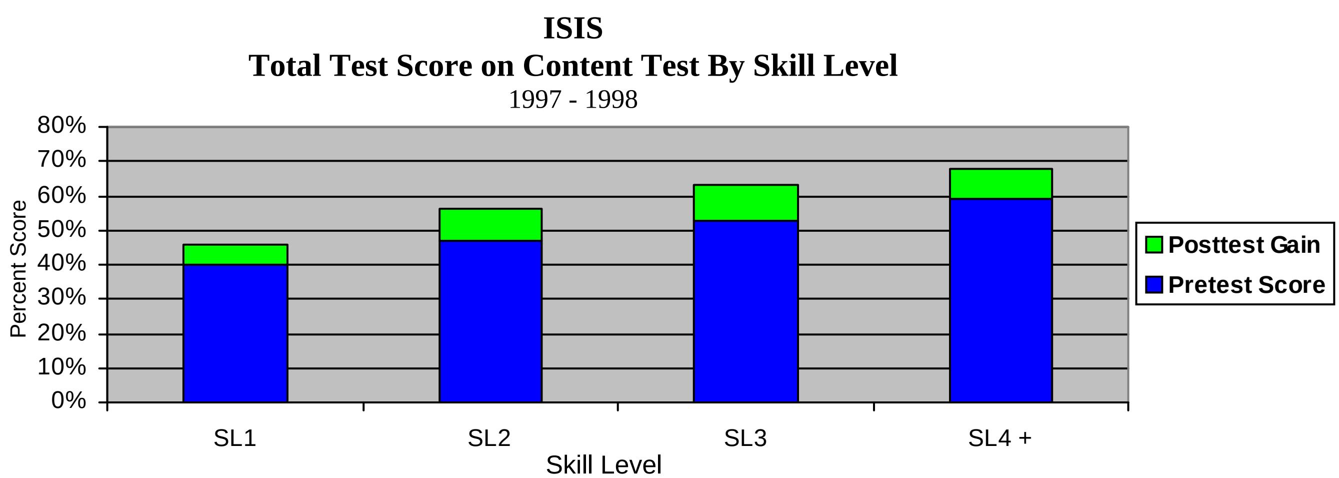

Figure 16. 1997-1998 ISIS content test pre-test scores and posttest gains by skill level

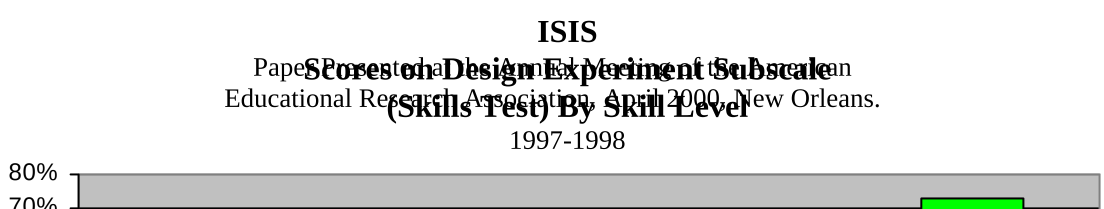

Figure 17. 1997-1998 ISIS skills test design experiment pre-test scores and posttest gains by skill level

Consistent with the earlier ISIS studies, the design subscales show that as students obtain higher skill levels their knowledge is increasing. It appears that there exists a threshold for students to develop a coherent understanding of scientific inquiry. Students may have to progress to a certain level of experience with scientific inquiry activities in order to get “the big picture.” This is supported by the design of ISIS in that students reaching the third level of skill development use the simulated biomes to conduct their experiments. It is at this point that they carry out a wide range of scientific activities: planning and conducting their experiments.

On-Going Studies and Analyses

Data for MAESTRO and WPS for the 1997 - 1998 school year has been collected and is currently being analyzed, as well as data for all three tutors for the1998 -1999 academic year.

Discussion

The results over the different studies for the different tutors show that the FST tutors have been instructionally effective in real-life educational settings. The studies have not only compared the FST tutors to traditional teacher-based, large-class methods, but also other forms of technology (e.g., using a word processor). The studies also show that the effects are greatest for those students who spend more time on the tutors or complete more assignments.

In general, while the overall results show that the FST tutors have been instructionally more effective than the traditional, teacher-based, large-class methods, the effects appear only for those students who have had a moderate amount of success with the tutors. In general the results of these studies demonstrate that the instructional gains are fairly stable over the years with the research designs employed and the varying samples. The results demonstrate that the students, who had more success using the tutors in terms of number of research assignments or modules completed showed higher gains from the pre- to post-test. One possible interpretation is that academically better students learned more, implying that only bright students learn from computer-based forms of instruction. However, in-depth analysis with ISIS showed that inquiry skill level achieved predicted performance even when pre-test scores were accounted for in the analysis (Steuck & Miller, 1997; Meyer, Miller, Steuck, & Kretschmer, 1999). Furthermore, the strength of the covariate regressions suggest that while initial aptitude is the best predictor of final test performance, it is not necessarily the best predictor for gains across the academic year. An alternative possibility is that there is a threshold for students to develop a coherent understanding of scientific inquiry (Schauble, Glaser, Duschl, Schulze, & John, 1995). Students may have to progress to a certain level of experience with each tutor to obtain the "the big picture."

Similar results were also evident in the Maestro data when analysis were run to determine performance gain by use. Students who spent more time working on the writing tutor have significantly better gain scores. While time on task and other environmental issues may have impacted the results, the overall gains are evident for the students who invested into their assignments.

Owston (1997) asked three questions about the implementation of Internet technologies in the classroom that can be raised for any educational technology not just internet-based approaches. The questions address the instructional efficacy of the technology, making learning more accessible, and containing the cost of education.

The outcomes of this research demonstrates that ITSs promote learning relative to traditional, non-technology-based approaches. The gains are moderate, and they are present for those students who complete a modest number of assignments or modules. Providing educational opportunities for students to learn scientific inquiry skills was a point raised in the interviews with teachers using the tutors in their classrooms. For example, with ISIS the teachers pointed out that their students learned science-related skills faster and more systematically than in previous years. Moreover, many believed that ISIS gave students experience with scientific thinking that they could not provide in the traditional, cookbook lab settings. While anecdotal evidence must be taken with caution, the information from the teachers lends credence to the notion that technology-based learning environments, such as ISIS, provide learning opportunities that teachers cannot normally provide.

The cost of implementing the tutors is difficult to assess. Financially, the cost is not great because the software was on loan in return for participation in the research study, and it is not yet a commercial product with an associated price. The hardware used in this study was either owned by the school district or on loan, again, in return for participation in this study. The course materials were initially developed and maintained by an interdisciplinary team funded under the research project. To continue using the tutors, though, would not be very costly. The primary cost would be to repair and upgrade the outdated hardware and to provide a technician in the computer lab. Curricular "costs" are actually more of a burden than financial costs. Teachers must decide what topics they must cut from their normal classroom curriculum in order to allocate time to using the software. In some cases, the choice is easy because the topic is covered in the tutoring system. However, eliminating 15-20 hours of in-class time over an academic year requires teachers to seriously reconsider their curricular choices. These choices are not easy, given the pervasive emphasis on state testing requirements.

In the ten years of the project, the FST team has found that the implementation of the technology is just as important as the efficacy of the software (Steuck, Rowley, & Kretschmer, 1999). For example, teacher training is one of the more important components required for successful implementation of the tutors. Many teachers do not have experience using computers to deliver student instruction. Instead, they use computers for recording grades and to develop paper-based handouts and exams. Teachers need not only to understand instructional software from a user's point of view, but also how to teach with it. The teacher's role in the computer lab classroom is not to put hands on the keyboard or mouse, but to stand next to the students acting as a partner in the instructional process.

Another key factor in the success of technologies is the support of the administration. A case in point is the critical issue of the number and arrangement of computers in an educational setting. One arrangement is to have 25-30 computers networked in one room. This arrangement provides opportunities for equal access to the technology for all students. Depending on the software, teachers are able to individualize the instruction to the needs of the students. An alternative arrangement that is growing in popularity is to have 4-5 computers in the back of the regular classroom. Proponents state that this arrangement will provide teachers more opportunities to individualize the curriculum. On the other hand, some teachers in the project have expressed concern that the computers in the classroom arrangement will cause classroom management problems. They are concerned that teachers will not be able to provide equal access to the technology. If access is based on student-choice, then students who are less computer literate or hold negative attitudes toward computers may not benefit as much as computer enthusiasts. If access is teacher-driven, administrative overload may occur trying to balance access time and curriculum covered. Another issue is the amount of computer use. In the FST lab-based arrangement, computers are used almost constantly. Skeptics of the classroom-based arrangement fear that the computers in the classroom will not be used as frequently by the students, lowering the total access time to the available technology.

Specifically, the FST team has developed a model of implementation that is comprised of seven factors including:

1) Learning outcomes (i.e. what students learn). This is affected by the quantity of time students interact with the software and the quality of that interaction.
2) Teacher variables. Teachers' workload, skills, and motivation are impacted by external requirements (e.g., state-mandated testing), decisions by administrators, quality of training on the technology, and their own history with the domain and technology.
3) Administrators. They affect the quality and quantity of technology available to teachers and students. Administrators also affect teachers' motivation by giving incentives, stipends, workload reduction, or by placing requirements to use the hardware and software.
4) External requirements, such as system mandated curricula.
5) Technology variables, such as the presence and quality of technical support, hardware, and software.
6) Student variables, such as motivation, access, and history.
7) Project Staff. The staff affects teacher training, administrative decisions, technical support, and the quality of the software.

Steuck, et.al. (1999) further describe that lower than desired learning outcomes may be attributed to one or more problems associated with a component in the model. An example of one of the problems would be lack of administrator support. This, in turn, may lead to decreased teacher motivation in implementing the technology and result in its discontinuation. However, some of the potential "pitfalls" that make it difficult, if not impossible, for the beneficial software's implementation to be successful, may be avoided by working closely with a project staff. However, as Baker states, "if researchers are not prepared to engage to some extent in these social and institutional issues, then I do not think it is advisable for them to seriously envisage putting ITS or any other systems into schools."(Baker, 1999)

Summary

In summary, the outcomes of this research demonstrate that ITSs promote learning relative to traditional, non-technology-based approaches. While the gains are modest, they are present for those students who complete a moderate number of assignments embedded in the tutors.

Furthermore, the implementation information gathered and studied illustrates that while placing ITSs into schools is not impossible, there are several issues that need to be identified in each setting and addressed. It is not an unattainable goal to successfully implement proven technologies into schools. In addition to the achievements of the FST tutors, the research described by Koedinger, Anderson, Hadley, and Mark (1997) with the Pump Algebra Tutor (PAT) also provide insight into the success of implementing technologies into the schools and involving teachers at the "grass roots."

In the near future, large-scale implementation of ITSs into school systems may be obtainable but not without firm commitments from research projects, strong support from administrators, and continuing support from teachers. Until such engagements and responsibilities can be shared, ITSs should only be implemented in small doses to demonstrate the continuing success of this type of technology. These systems may also serve to prepare future educators in the use of more sophisticated learning technologies which will benefit students.

References

Anderson, J.R., & Fincham, J.M. (1994). Acquisition of procedural skills from examples. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1332-1340.

Baker, M. (1999). The roles of models in Artificial Intelligence and Education research: a prospective view. http://cbl.leeds.ac.uk/ijaied/abstracts/Vol_11/baker.html 11, to appear in the International Journal of Artificial Intelligence.

Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 41, 4-17.

Koedinger, K.R., Anderson, J.R., Hadley, W.H. & Mark, M.A. (1997). Intelligent tutoring goes to school in the big city. International Journal of Artificial in Education, 9, 45-87. [http://cbl.leeds.ac.uk/ijaied/]

Meyer , T. N., Miller, T. M., Steuck, K., & Kretschmer, M. (1999). A multi-year large-scale field study of a learner controlled intelligent tutoring system. Proceedings of the 9th International Conference on Artificial Intelligence in Education, Le Mans, FR.

Meyer, T., Steuck, K., Miller, T., Pesthy, C., & Redmon, D. (1999, March). Lessons learned from the trenches: Implementing technology in public schools. . Paper presented at the Society for Information Technology and Teacher Education Conference, San Antonio, TX

Meyer, T., Steuck, K., Miller, T., Pesthy, C., & Kretschmer, M. (1999, March). Teaching scientific inquiry skills with an intelligent tutoring system. Paper presented at the Mathematics/Science Education & Technology Conference, San Antonio, TX

Owston, R. (1997). The World Wide Web: A technology to enhance teaching and learning? Educational Researcher, 26 (2), 27 - 33.

Reigeluth, C. M. (1987). Lesson blueprints based on the Elaboration Theory of Instruction. In C. M. Reigeluth (Ed.), Instructional theories in action (pp. 245-288). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M. (1992). Elaborating the elaboration theory. Educational Technology, Research & Development, 40, 80-86.

Rosch, E. (1988a). Coherences and categorization: A historical view. In F. S. Kessel (Ed.), The development of language and language researchers: Essays in honor of Roger Brown (pp. 373-392). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, E. (1988b). Principles of categorization. In A. Collins & E. E. Smith (Eds.), Readings in cognitive science (pp. 312-322). San Mateo, CA: M. Kaufmann Publishers.

Rowley, Kurt. (1995). Understanding software interoperability in a technology-supported system of education. Cause/Effect, 18(3), pp.20-26.

Rowley, K. & Crevoisier, M. (1997, June). MAESTRO: Guiding students to skillful performance of the writing process. Proceedings of the Educational Multimedia and Hypermedia conference, Calgary, CA.

Rowley, K., Miller, T., & Carlson, P. (1997). The influence of learner control and instructional styles on student writing in a supportive environment. ERIC Document ED404668.

Schauble, L., Glaser, R., Duschl, R.A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. The Journal of the Learning Sciences, 4(2), 131-166.

Steuck, K. & Miller, T. (1997, March). An evaluation of an authentic learning environment for teaching scientific inquiry skills. Paper presented at the Annual meeting of the American Educational Research Association, Chicago, IL.

Steuck, K.W., Rowley, K., & Kretschmer, M. (1999). Partnering to implement computer-based tutoring systems in secondary schools. Journal of Interactive Instruction Development, 12(1), 16-22.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. Cognition and Instruction, 7, 1-39.

Wheeler, J. L., & Regian, J. W. (1999). The use of a cognitive tutoring system in the improvement of the abstract reasoning component of word problem solving. Computers in Human Behavior, 15, 243-254.