PAT Goes To College: Evaluating a Cognitive Tutor for Developmental Mathematics

Kenneth R. Koedinger (ken.koedinger@cs.cmu.edu)

Erika L. F. Sueker (sueker+@andrew.cmu.edu)

Human Computer Interaction Institute, School of Computer Science, Carnegie Mellon University, USA

Abstract: This report describes an adaptation of the Practical Algebra Tutor, "PAT," for college-level developmental mathematics, and initial evaluations of PAT at two college sites. PAT is a software learning environment that presents students with real-world problem situations, modern mathematical representational tools to analyze these situations, and constant background support from a "cognitive tutor"-- an intelligent computer tutor based on the ACT theory of cognition. Previous classroom evaluations showed that high-school students using PAT outperformed students in regular classes by 15% on standardized tests, and by 100% on assessments of authentic problem-solving. In two colleges, PAT students performed more than 50% better than students in regular classes on a performance-based assessment. This assessment requires students to use mathematical representations to analyze a real-world problem situation. It captures the reform objectives of the PAT approach, which are consistent with new national standards for mathematics.

The Educational Problem

Developmental mathematics courses, defined as college-level courses prior to calculus, are the fastest-growing post-secondary mathematics courses in the nation. From 1965 to 1985, total enrollment in college mathematics courses increased 60%, while enrollment in college courses in high-school algebra and geometry increased by 250% [Madison & Hart 1990]. Developmental mathematics courses present unique challenges for instruction, but instructors currently have limited resources with which to meet these challenges. Based on input from developmental mathematics instructors, administrators, and publishers, we have identified a profile of the developmental mathematics student population that has served to focus our efforts at instructional innovation. Developmental mathematics students tend to have one or more of the following characteristics: (1) high variability in age and prior mathematical background; (2) a prior history of failure and frustration in mathematics; (3) a fragmented, proceduralized understanding of mathematics; (4) a lack of appreciation for the direct importance of mathematics to their lives; and/or (5) poor study skills, particularly, little awareness of how crucial it is to spend time working through problems if one is to succeed in mathematics. Thus we need an instructional solution that can adapt to the individual background of students, that offers a success experience, that interrelates various mathematics skills during problem-solving, that makes mathematics meaningful, and that can compensate for poor study skills. Also, despite their poor mathematics skills, many students enrolled in developmental courses declare college majors that depend heavily on mathematics (e.g., economics, biology, physics). Thus we need an instructional solution that prepares students for higher-level mathematics courses and that focuses on the application of mathematics to a variety of fields.

In this article we report on the use of a technological tool that addresses these problems. A computer-based environment for learning algebra has been developed through a collaborative effort between cognitive psychologists, computer scientists, and educators at the Pittsburgh Advanced Cognitive Tutoring (PACT) center of Carnegie Mellon University. Here we present the initial results of an ongoing experiment that integrates this computer-based environment into developmental algebra courses at two college sites.

A Solution: Intelligent Computer Tutors

Cognitive Tutors

Cognitive tutors are a type of intelligent computer tutoring system well suited to the unique challenges of teaching developmental mathematics [Anderson, Corbett, Koedinger & Pelletier 1995]. They have at their core a fine-grained computer model of the knowledge, skills, and strategies involved in competent student problem solving in the subject-matter domain. That is, within the tutor is a simulation, developed through psychological research, that can solve mathematics problems in the ways we expect students to solve the problems, as well as simulate the kinds of mistakes that students make.

Cognitive tutoring technology is grounded in the ACT theory of cognition, perhaps the most comprehensive, empirically tested and influential of all contemporary theories of human learning [Anderson 1993]. Two key premises of the theory that are particularly relevant to developmental mathematics are that (1) learning occurs in the context of a problem-solving activity, and (2) learning events occur at a grain size well captured by ACT's production rule notation.

Like much computer-based instruction, cognitive tutors allow for self-paced use. Unlike other approaches, the cognitive model within the tutor facilitates three additional forms of customized instruction: (1) Individualized feedback and advice can be supplied as a student solves a problem; through a technique called "model tracing" [see Anderson, Boyle, Corbett & Lewis 1990], the tutor uses a student's behavior to identify the chosen solution path and prompts the student appropriately; (2) The number and types of problems are individually selected to maximize learning opportunities; using "knowledge tracing" [see Corbett & Anderson 1992], the tutor follows the student's knowledge acquisition process and selects problems based on a student's past successes and failures in applying each individual skill and strategy in the model; and (3) it is easy to customize the tutors for particular courses; this is important both because there is great variation between developmental mathematics courses and because college teachers want to put their own signature on the courses they teach.

Many laboratory studies have investigated the effects of cognitive tutors on student learning, and the success of these systems has also been demonstrated in real classrooms [see Anderson et al. 1995]. In addition, significant gains in student attitudes and in positive teacher behaviors, including acting and being seen by students more as a facilitator than as an authority, have been reported in classrooms using a cognitive tutor [Schofield, Evans-Rhodes & Huber 1990]. In this paper we focus on a cognitive tutor called the Practical Algebra Tutor, or "PAT". In a previous evaluation of PAT in high school algebra classes, students using the tutor with a reform curriculum performed 100% better on assessments of problem solving and representation use, and 15% better on standardized tests than students who followed a traditional curriculum without the tutor [Koedinger, Anderson, Hadley & Mark 1995]. In this paper we present the initial results of an experiment integrating PAT into developmental algebra courses at two universities.

Curriculum Reform

The current emphasis in mathematics reform efforts is on the use of multiple representations and computational tools to solve difficult real-world problems that involve symbolic expressions. The PAT curriculum is consistent with reform efforts at the high school [NCTM 1989] and new standards of the American Mathematics Association of Two Year Colleges [AMATYC 1995]. The emphasis in PAT is less on symbol manipulation and more on symbolization as way to unleash the power of mathematics for problem solving. [Fig. 1] shows the PAT screen after completion of a problem. For a more complete description of PAT see [Koedinger et al. 1995]. Here we illustrate three key features of PAT:

(1) While PAT covers the traditional content of beginning algebra (e.g., linear functions, systems of equations), these skills are introduced in the context of authentic, realistic problem-solving tasks. The problem in Figure 1 involves using systems of equations to compare the cost of two car rental companies, Avis vs. Hertz. One goal is to find the circumstances under which one company is more cost effective than the other (Answer: rent from Hertz when traveling less than 750 miles).

(2) PAT provides computerized computation and visualization including a tabling tool, (see upper-right of [Fig. 1]) a graphing tool (lower-left) and an equation solving tool (lower-center). By allowing manipulation

of algebraic functions through the use of multiple representations, the tutor promotes deeper understanding of the meaning of the symbolic expressions [Kozma 1993]. One component of the college algebra curriculum that is critical to future mathematics success is skill in judging when two algebraic functions are equivalent. With the multiple representations that are in our current system a student could, for example, use a table to check when the two functions produce the same output for the same input, see where the graphs of the functions intersect, or use symbolic manipulation to solve the equation of the two functions.

(3) Instruction promotes instance-based investigation, a flexible problem-solving strategy often used by competent problem solvers to support their use of abstract mathematical formalisms as they attempt to make sense of a difficult problem [Koedinger & Anderson 1990] [Koedinger & Anderson 1995]. PAT promotes the use of this strategy by asking students to first investigate concrete instances of unknown values, e.g., "How much would you owe each company if you drove 500 miles?" see row 2 of the Worksheet in [Fig. 1]. After answering the questions, they attempt to formulate the abstract rule, see the "Formula" row of the Worksheet in [Fig. 1]. Instance-based investigation can lead students to solutions, particularly in problems requiring a discrete decision (e.g., from which company does it cost less to rent a truck?) as do many real-world problems.

Assessment of Learning Environments

Monitored Design

[Brown 1992] has stressed that success in the learning sciences requires going beyond laboratory instructional experiments to face the complex issues of engineering instructional solutions in real learning environments. Instructional design should (1) be guided by theoretical principles, (2) address multiple facets of instructional innovation, for instance, teacher training in addition to integrated technology use, (3) be evaluated in terms of the instructional objectives of multiple stake-holders, and (4) be made practically feasible so as to be easily disseminated. A point not enough emphasized in [Brown 1992] is the importance of having comparison conditions against which to evaluate the design. In this study, for instance, we monitored outcomes both in classes using PAT and in comparable classes without PAT. These results will provide a baseline to monitor design improvements of future iterations in studies planned for the 4 semesters and summer sessions between now and the summer of 1997. We call our approach to design experimentation "monitored design" to emphasize the close monitoring of outcomes that guide the iterative design process and provide a database for justifying the future dissemination of PAT.

Overview of the Experiment

This experiment is part of a monitored design process of an instructional solution that addresses the challenges of teaching developmental algebra. The tutor was implemented in a total of 11 classes at two universities. Within University P we controlled for teacher, curriculum and computer tool use, in that the same teacher taught the comparison class using the traditional curriculum and the experimental class with the traditional curriculum and the cognitive tutor. Within University L we controlled only for curriculum and computer tool use.

We were interested in two major outcomes of student learning: (1) algebraic manipulation skills, the goal of a traditional curriculum, and (2) use of these skills as well as alternative algebraic representations (e.g., tables and graphs) to analyze and solve real-world problem situations, the goal of the PAT curriculum. To assess these outcomes we used a performance-based assessment that targeted students' problem-solving skills, qualitative reasoning, and ability to communicate effectively about mathematics, as well as traditional algebra exams. We also assessed students' attitudes toward mathematics and computers before and after working with the tutor.

This evaluation differed from the previous evaluation of PAT at the high school level [Koedinger, Anderson, Hadley & Mark 1995] in two ways: (1) in this experiment, PAT was added to a traditional rather than a reform curriculum, and (2) the performance-based assessment was used in comparison classes as well as in the experimental classes.

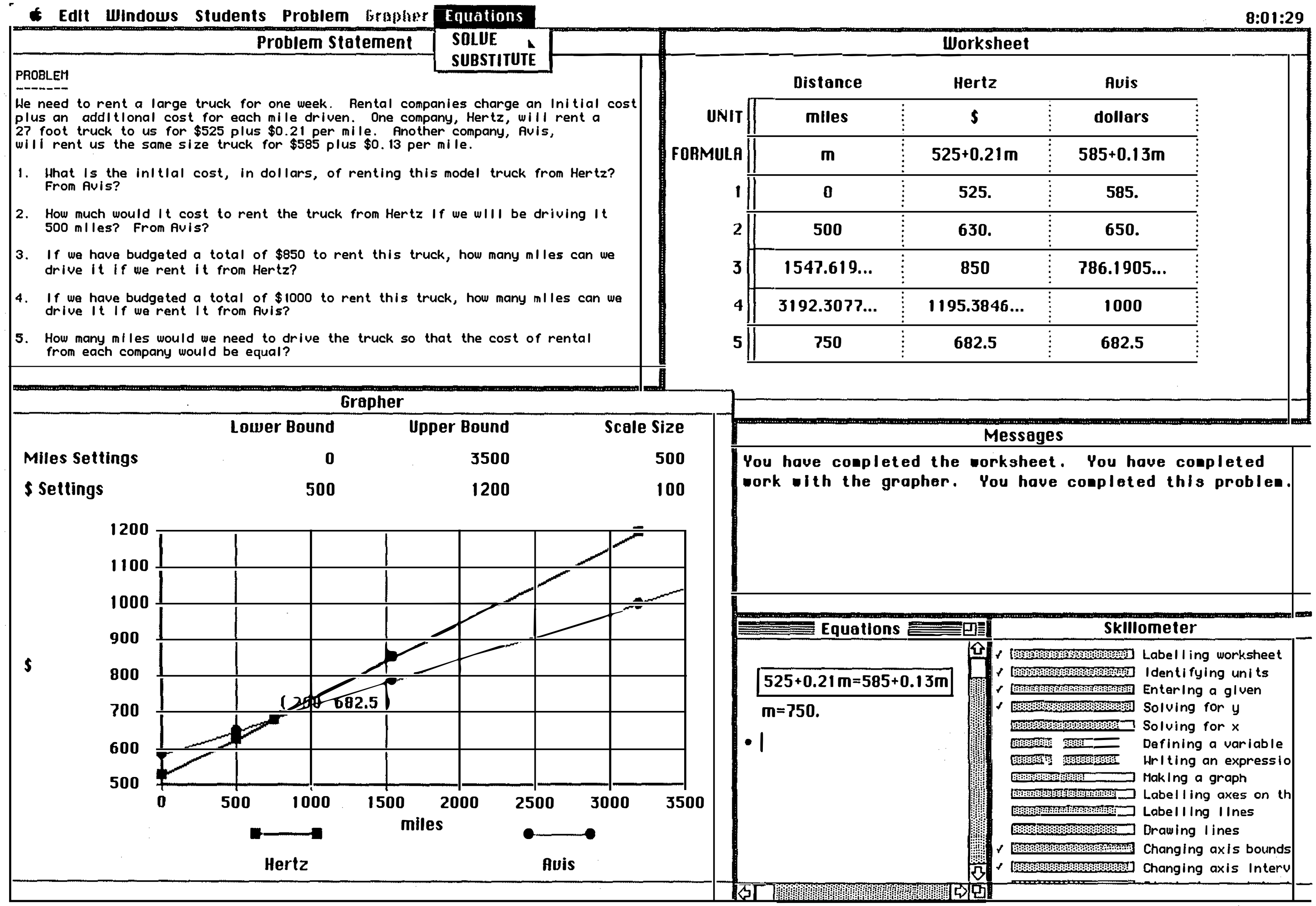

Figure 1. PAT (Practical Algebra Tutor) provides just-in-time intelligent support for the use of multiple algebraic representations in the analysis of real-world problem situations.

Experimental Evaluation

Design

At University P, one comparison class (n = 16) studied a standard beginning college algebra curriculum, and one experimental class (n = 12) studied the same curriculum with the same instructor and also used the PAT cognitive tutoring system. At University L, three comparison classes (n = 88) studied a standard curriculum, and six experimental classes (n = 199) used PAT along with the same curriculum. The nine classes were taught by a total of five different instructors, none of whom taught both an experimental and a comparison class.

Student Population

The two Universities have a large developmental mathematics curriculum. Reflecting the demographics of their respective cities, a large portion of University P's student population is African American. University L's student population is more diverse and includes a large portion of Hispanic students. The students' mean scores on the Mathematics portion of the SAT were 419 (University P) and 395 (University L).

Procedures and Curricula

At University P the course met for 15 weeks, 4 days per week, 50 minutes per day. During 11 class sessions, students in experimental classes met in the computer lab and worked on 3 word problems with the tutor. At University L the course met for 15 weeks, 3 days per week, 75 minutes per day. During 10 class sessions, students in the experimental classes met in the computer lab and completed up to 15 problems with the tutor. The cognitive tutoring software was stored on a server and accessed by students from Apple Macintosh machines. The computer lab was also accessible to students outside of class time. Students were encouraged to use cognitive tutors outside of class, for remedial tutoring, completing unfinished lessons, and as homework when directed by the instructor. More students used the tutor outside of class time at University L than at University P.

The tutor curriculum covered solving linear equations, graphing, slope-intercept graphing, finding unknowns with equations, and solving and graphing systems of equations, all within situated problem-solving scenarios. Classroom instruction covered traditional algebraic symbol manipulation, e.g., simplifying equations, factoring, logarithms, and inequalities. During the last week of class, all students were given approximately 20 minutes to work on the performance-based assessment. Students also took a Final Exam covering traditional algebraic manipulation. Students completed the attitude survey before and after the course.

Outcome Measures

Performance-based Assessment: "The Cellular Phone Problem"

Given pencil and paper (and at University P optional use of a graphics calculator) students are asked to do a mathematical analysis of two cellular phone services, given a basic monthly charge and a rate per minute of usage for each service. The problem was contextualized as a requirement given by a "boss" to a "company employee" (the student). The goal was for the employee to decide which cellular phone service would be appropriate for officers of the company, based on the amount of time each officer uses the cellular phone. The basic elements required for the "mathematical analysis" were clearly listed in the problem statement: (a) defining variables, (b) writing equations, (c) making tables, (d) constructing graphs, (e) finding slopes and intercepts, and (f) finding points of intersection. Additional problem-solving tasks were presented as criteria suggested by the boss, e.g., that the analysis should: (g) specify the amount of usage each service allows for a total cost of $100, and (h) state the range of usage for which each service is cheapest.

The problem context stated that the employee will have to make decisions based on the analysis shortly after the boss would provide the necessary information on each officer's normal phone usage. This implicitly encouraged students to produce a graph of the two functions from which values can be read off quickly.

Performance was evaluated by breaking the Cellular Phone Problem down into eight component tasks that corresponded to the above eight requirements of the problem statement (a-h). A single researcher assigned each student's solution a score of 0, 0.5, or 1 point on each component, based on predetermined criteria. Component scores were analyzed separately and were also averaged together to produce an overall mean score for each student that ranged between 0 and 1.

Final Exam

The final covered traditional symbolic manipulation skills, including solving equations and inequalities, evaluating logarithmic expressions, simplifying equations, and graphing the equation of a line. Problems were presented as mathematical expressions without context (no word problems), and no verbal responses were required. A typical problem was:

Factor this expression: $x^3y - 3xy + x^2 - 3$

Attitude Survey

The survey was adapted from two existing surveys [Kloosterman & Stage 1992][Nickell & Pinto 1986] and contained 50 items presented in a random order. Items comprised 6 subscales that assessed students': (1) lack of computer anxiety, (2) belief that effort contributes to successful mathematical problem-solving, (3) belief that skill in mathematics is useful in life, (4) belief that there are word problems that cannot be solved by following a set step-by-step procedure, (5) confidence in their own ability to solve difficult mathematics problems, and (6) belief that word problems are an important component of mathematics. The first subscale, lack of computer anxiety, was made up of 20 items, the 5 mathematics subscales contained 6 items each. All responses were made on a five point scale ranging from "strongly agree" to "strongly disagree."

Results

Problem-solving Ability

On the performance-based assessment, students in the experimental groups scored higher than students in the comparison groups overall see [Tab. 1], last row. On two of the eight component scores, Making a Table and Stating the Ranges of Usage, there were statistically significant differences between the two groups at both universities, with the experimental group outperforming the comparison group. On the remaining six component scores, Defining Variables, Solving for y = $100, Writing Equations, Finding Slopes and Intercepts, Constructing a Graph, and Finding Points of Intersection there were significant differences between the experimental and comparison classes at University L, but not University P. This is likely due to a lack of statistical power at University P, given the smaller number of students participating at that site.

Traditional Algebraic Skills

The scores of students in the experimental and comparison classes on the departmental final exams did not differ significantly at either university (*F*s < 1). This shows that the reduced lecture time in the experimental classes (because some lecture periods were spent in the computer lab instead) did not detract from the experimental students' acquisition of the algebraic manipulation skills targeted by the exams.

	University L (n = 287)		University P (n = 28)	
Component Score	Experimental (n = 88)	Comparison (n = 199)	Experimental (n = 12)	Comparison (n = 16)
(a) Defining variables	.83 ***	.54	1.00	.84
(b) Writing equations	.72 ***	.51	.75	.69
(c) Making a table	.52 ***	.22	.67 **	.16
(d) Constructing a graph	.38 ***	.11	.42	.34
(e) Finding slopes and intercepts	.12 ***	.03	.02	.02
(f) Finding points of intersection	.16 ***	.02	.12	.06
(g) Solving for y = $100	.49 ***	.28	.54	.31
(h) Stating ranges of usage	.17 ***	.07	.17 *	.00
Mean Overall	.42 ***	.22	.46*	.30

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 1. Average percent correct on eight component scores of the Cellular Phone application problem

Attitudes Toward Computers and Mathematics

In general, students at both universities reported a positive attitude toward mathematics and computers, one that changed little prior to and following instruction. After instruction, attitudes of students in the experimental group at University P tended to differ from those in the comparison group on two subscales: Students who used the tutor had slightly less computer anxiety than comparison students (M(exp.) = 2.4, M(comp.) = 2.6, one-tailed $t = 1.99$, $p < .05$) and they were more likely to believe themselves capable of solving time consuming mathematics problems (M(exp.) = 2.3, M(comp.) = 2.6, one-tailed $t = 1.58$, $p = .06$).

Discussion

Overall, results showed that students who used the PAT tutor solved the Cellular Phone Problem, a real-world application of algebra, more accurately than comparison classes. Using the tutor did not negatively affect students' ability to perform traditional symbolic manipulation, as demonstrated by their final exams, or their attitudes toward mathematics or computers. However, there is much room left for improvement, especially on higher-level tasks such as judging the range of airtime for which each cellular phone service is the cheapest.

There was some evidence that students using the tutor had a better understanding of problem goals in that they were better able to state the range of values for which each cellular phone service was the cheapest. More students in the comparison classes seemed to believe that the problem goal had been met when they had solved the equations and found that one service allowed a greater amount of phone time for a total cost of $100. However, the functions describing the total cost of each service intersect, creating two ranges of values. Basing a decision about the "best" cellular phone service on a single point does not take this fact into account.

The current developmental mathematics curriculum over-emphasizes formal skills to the detriment of useful algebraic skills. The need for reform in algebra courses is made particularly salient by the contrast between the large amount of time spent on helping students master skills of limited applicability, like factoring expressions in the final exam problem illustrated above, and the poor performance of students on applying algebra to problem situations. Less than 7% of the 215 students in the traditional classes in this evaluation were able to make a reasonable conclusion about the Cellular Phone Problem. Unlike the current situation in traditional algebra courses, we believe that useful algebra should lead formal algebra instruction, not lag behind it. Practical problems such as those the tutor presents should serve as the context in which symbol manipulation skills are learned. By integrating PAT into reformed curricula that have this objective, we expect to see even more dramatic learning gains that are particularly relevant to student retention and success in college. Reform efforts should include the creation of new assessments, such as the Cellular Phone Problem, that better reflect the students' future needs for algebraic representations and reasoning processes.

Our future research will continue development and evaluation of PAT, focusing on student learning and attitudes, and successful teacher behavior. We will work with universities to reform developmental mathematics curricula, and strive to help teachers achieve maximal integration of the software into existing curricula.

References

[AMATYC 1995] American Mathematical Association of Two-Year Colleges (1995). Crossroads in mathematics: Standards for introductory college mathematics. Memphis, TN: AMATYC.

[Anderson 1993] Anderson, J.R. (1993). Rules of the Mind. Hillsdale, NJ: Erlbaum.

[Anderson, Boyle, Corbett & Lewis 1990] Anderson, J.R., Boyle, C.F., Corbett, A., & Lewis, M.W. (1990). Cognitive modeling and intelligent tutoring. Artificial Intelligence, 42, 7-49.

[Anderson, Corbett, Koedinger & Pelletier 1995] Anderson, J.R., Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. Journal of the Learning Sciences, 4, 167-207.

[Brown 1992] Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. Journal of the Learning Sciences, 2, 141-178.

[Corbett & Anderson 1992] Corbett, A.T. & Anderson, J.R. (1992). Student modeling and mastery learning in a computer-based programming tutor. In Proceedings of the Second International Conference on Intelligent Tutoring Systems. Montreal, Canada.

[Kloosterman & Stage 1992] Kloosterman, P. and Stage, F.K. (1992). Measuring beliefs about mathematical problem solving. School Science and Mathematics, 92(3), 109-115.

[Koedinger & Anderson 1995] Koedinger, K.R., & Anderson, J.R. (1995). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. Manuscript submitted for publication.

[Koedinger & Anderson 1993] Koedinger, K.R., & Anderson, J.R. (1993). Reifying implicit planning in geometry: Guidelines for model-based intelligent tutoring system design. In Lajoie, S., & Derry, S. (Eds.) Computers as Cognitive Tools. Hillsdale, NJ: Erlbaum.

[Koedinger & Anderson 1990] Koedinger, K.R., & Anderson, J.R. (1990). Abstract planning and perceptual chunks: elements of expertise in geometry. Cognitive Science, 14, 511-550.

[Koedinger, Anderson, Hadley & Mark 1995] Koedinger, K.R., Anderson, J.R., Hadley, W.H., and Mark, M.A. (1995). Intelligent Tutoring Goes To School in the Big City. In Proceedings of the World Conference on Artificial Intelligence in Education, (pp. 421-428). Charlottesville, VA: Association for the Advancement of Computing in Education.

[Kozma 1991] Kozma, R.B. (1991). Learning with media. Review of Educational Research, 61, 179-211.

[Madison & Hart 1990] Madison, B.L. and Hart, T.A. (1990). A challenge of numbers: People in the mathematical sciences. Washington D.C.: National Academy Press.

[NCTM 1989] National Council of Teachers of Mathematics (1989). Curriculum and Evaluation Standards for School Mathematics. Reston, VA: NCTM.

[Nickell & Pinto 1986] Nickell, G.S. and Pinto, J.N. The computer attitude scale. Computers in Human Behavior, 2, 301-306.

[Schofield, Evans-Rhodes & Huber 1990] Schofield, J.W., Evans-Rhodes, D., & Huber, B.R. (1990). Artificial intelligence in the classroom: The impact of a computer-based tutor on teachers and students. Social Science Computer Review, 8, 24-41.

Acknowledgments

Support was provided by the Fund for the Improvement of Post-Secondary Education, Dept. of Education, grant number P116B41269 to the first author. We wish to thank Bonnie Goins for her research assistance, Lora Shapiro and Mark Clark for their dedication as site coordinators, and all the instructors who participated in the experiment.