Effective Use of Intelligent Software in High School Math Classrooms

K. R. KOEDINGER AND J. R. ANDERSON
Psychology Department
Carnegie Mellon University
Pittsburgh, PA 15213
E-Mail: koedinger@psy.cmu.edu

Abstract: In efforts to move our Intelligent Tutoring Systems from successful demonstrations in the laboratory and college classrooms to the less forgiving setting of urban high school classrooms, we have had to face a number of educational issues that have not typically been part of ITS development. We discuss these issues in the context of a preliminary evaluation of ANGLE, a tutor for geometry theorem proving. Because of recent changes in US mathematics curriculum standards, including "decreased attention" to formal proof in geometry (NCTM, 1989), we needed to think carefully about how best to integrate the system into the new curriculum and, in addition, saw less use than we had originally planned. Nevertheless, the 4-5 weeks of classroom use of ANGLE led to large student gains relative to same-teacher classes, but only for the teacher who had participated in the system development. Two educational lessons have emerged from this evaluation that go beyond our usual concerns for cognitive fidelity and laboratory evaluation. First, we need to be responsive to curriculum changes and issues of integrating tutors into the kind of curricula that educators advocate for the future. Second, we need to better understand the role of the teacher in the classroom use of ITSs and how we can better prepare and support teachers in making full use of the potential of these systems.

Introduction

Most early attempts at applying artificial intelligence to education involved the construction of intelligent tutoring systems (ITSs) as proofs of concept of the promise of the technology (e.g., Sleeman & Brown, 1982). Certainly the difficulty involved in building realistic and usable systems (Schank, 1991) has warranted this focus on the technology, however, it has become well recognized that the field must also provide empirical demonstrations that ITSs are effective in improving learning. In fact, there have been a number of successful evaluations of ITSs in the laboratory (e.g., Mark & Greer, 1991), in job training (e.g., Lajoie & Lesgold, 1989), and in classrooms (e.g., Anderson, Boyle, Corbett, & Lewis, 1990) and there is increasing emphasis on evaluation. At the same time there has been a recognizable movement away from ITSs, curiously enough, toward other forms of educational software that may or may not incorporate intelligence (Lajoie & Derry, 1993; Larkin & Chabay, 1992; Schank, 1991) – this trend appears to be a response, in part, to the difficulty of building realistic ITSs coupled with the growing push to create systems that are demonstrably effective.

The intention of this paper is to present further evidence that ITSs can be applied in realistic settings and lead to improved learning. In addition, we make a point that applies as well to other types of instructional software as it does to intelligent tutors, namely, to effectively implement instructional software in realistic school settings, researchers must carefully address issues of instructional context, most importantly, curriculum integration and teacher training.

We report on a second generation ITS for geometric deduction and its use in classrooms of a local urban high school. Geometry is an excellent model domain to explore AI and education issues for a couple of reasons: 1) cognitive modeling in geometry presents interesting yet tractable challenges for AI and 2) geometry is a proven educational challenge in that students have great difficulty with learning geometric proof (Senk, 1983).

Appears in *Artificial Intelligence in Education: Proceedings of the World Conference on AI in Education, 1993*. Charlottesville, VA: AACE.

ANGLE: A New Geometry Learning Environment

Informed by basic research on the visually-based knowledge representation of competent problem solvers (Koedinger & Anderson, 1990), we built an ITS for geometric deduction called ANGLE (Koedinger & Anderson, 1993). ANGLE includes a graphical editor for entering geometry "flow proofs", a cognitive model or "expert component" for checking student proofs, and a tutor for providing feedback and advice on student's work. Figure 1 provides an example of the ANGLE interface in the middle of a problem solution. A proof solution is represented as a graph that shows how the problem goal ("A is the midpoint of BD" – at the top of the screen in Figure 1) is supported by a chain of inferences from the problem givens ("A is the midpoint of EC" and "BC is parallel to ED" – at the bottom of the screen). ANGLE also includes a proof entry interface that developers, as well as students and teachers, can use to enter new proof problems. A proof is entered by drawing the problem diagram in a MacDraw-like window and entering the given and goal statements. Because of the power of the expert system underlying ANGLE, a student can enter a proof problem and get tutored on it in a couple of minutes.

ANGLE's Cognitive Model of Proof Competence

Rather than assuming that competence in proof is derived purely from internalizing the formal domain rules, the cognitive model underlying ANGLE claims that proof knowledge is induced from and organized around prototypical perceptual *configurations*. The icon menu on the left of the screen in Figure 1 illustrates these configurations. The fifth icon up from the bottom, for example, represents the generic CONGRUENT-TRIANGLES configuration and the box in the middle of the screen labeled $\triangle ABC \cong \triangle ADE$ is a conjectured instance of this configuration (yet to be proven so). Attached to these configurations is the formal knowledge of the domain including both *properties* and *sufficient conditions* of the configuration. The CONGRUENT-TRIANGLES configuration schema has the properties that the 3 corresponding sides of the two triangles are equal and the 3 corresponding angles are equal. The sufficient conditions are subsets of these 6 properties, "shortcuts", that are sufficient to formally prove an object is an instance of a configuration. Certain combinations of 3 properties are sufficient to prove triangles congruent, for example, two sets of corresponding angles, $\angle E \cong \angle C$ and $\angle DAE \cong \angle BAC$, and the corresponding sides between them, $AC \cong EA$, are sufficient to prove $\triangle ABC \cong \triangle ADE$. This sufficient condition is an informal counterpart to what is usually the Angle-Side-Angle theorem or postulate in the traditional curriculum. We call this clustering of informal perceptual knowledge of configurations with the corresponding formal knowledge of properties and sufficient conditions *diagram configuration schemas.*

Part of the strategic competence in proof, according to Koedinger and Anderson (1990), involves the idea that a solution to a nominally *deductive* problem can often be aided by *inductive* reasoning -- that is, by using the problem diagram as a model and inducing from it possible statements that might be provable and thus possibly on the path to the problem goal. Statements that look true in an accurately drawn diagram provide a superset of the possible provable statements -- a set which is much smaller than the set of syntactically possible statements that might otherwise be searched. Further search efficiency is achieved by focusing the induction on those statements that correspond with diagram configuration schemas. As argued in Koedinger and Anderson (1990), this inductive approach focused on key conceptual steps is exactly approach taken by competent human theorem provers. In addition, from an AI perspective, this approach has led to an expert system that can solve problems more effectively and efficiently than the expert component in the previous Geometry Proof Tutor (Anderson et. al., 1990).

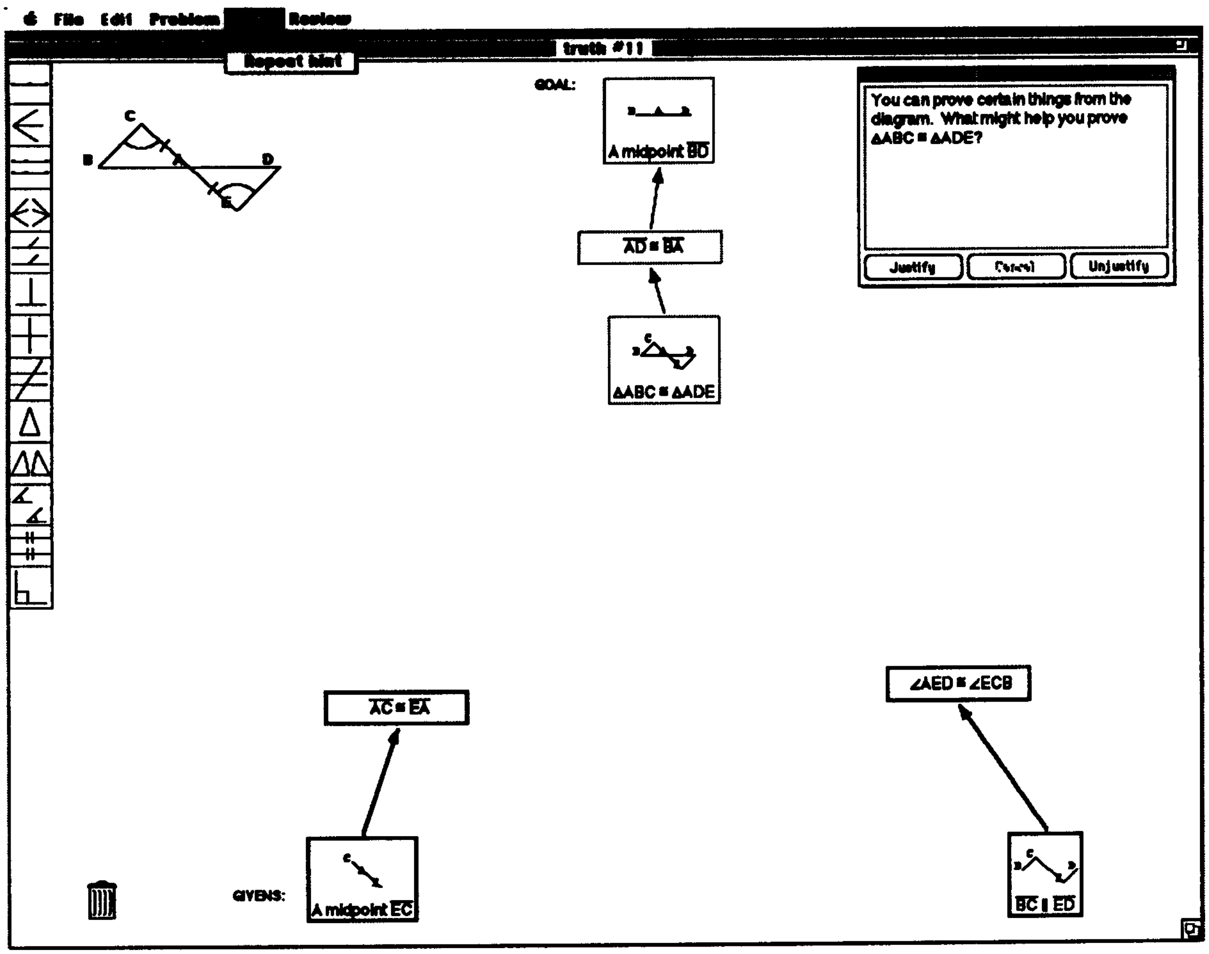


Figure 1. The ANGLE interface in the middle of a proof problem. The student has worked forward one step from each given and worked backward two steps from the goal. At this point a hint was requested from the Info menu and is shown in the message window.

Reification of the Cognitive Model in ANGLE's Interface and Tutor

In addition to the usual pedagogical goals of an ITS as providing timely and individualized advice, the intention of ANGLE was to communicate this particular and efficient cognitive model of proof problem solving. The communication is done implicitly through the structure of the interface as well as explicitly through the tutoring messages[1]. The icon menu on the left in Figure 1 is one form of implicit guidance that indicates the possible configuration schemas the student might consider. The process of entering a statement supports the inductive approach of the cognitive model by only allowing students to enter statements that look true in the diagram. Entering a statement involves using the icon menu to "conjecture" that something in the diagram is an instance of one of the configurations. For example, the configuration $\triangle ABC \cong \triangle ADE$ (in the middle of Figure 1) was entered using the Congruent Triangles icon (10th one down in the icon menu) and indicating where in the diagram this instance appears (in general the diagram may contain multiple instances of any configuration). When a statement is conjectured it appears on the screen in a dotted box. The student can then justify an entered statement by correctly linking it with a statement or set of statements that logically support it. For example, the statement $AC \cong EA$ was justified by indicating if follows from the given A midpoint EC. As shown in Figure 1 the result of a correct justification action is that an arrow is drawn from the premise(s) of the justification, A midpoint EC, to the conclusion $AC \cong EA$. If the premises of the

[1]Koedinger & Anderson (1993) can be consulted for more details on the "model-based design" approach used in the creation of ANGLE.

justification are all given or proven and thus already in a bold box, then the conclusion will now also appear in a bold box. The goal of the proof problem is to get the Goal statement (**A midpoint BD** in Figure 1) into this proven state.

ANGLE's proof graph representation provides a reification of the search process that leads to a proof solution. Students can make forward inferences up from the problem givens towards the goal (e.g., AC ≅ EA follows from **A midpoint EC** -- on the bottom-left in Figure 1) or backward inferences down from the goal towards the givens (e.g., AD ≅ BA can be proven if △ABC ≅ △ADE is -- in the middle of Figure 1). The proof graph records this search process in a way not possible on paper and provides a vehicle for discussion, for example, about garden paths in the search and how they are a normal part of the problem solving process even for expert theorem provers.

Explicit communication of the cognitive model comes through tutoring messages. One form of tutoring message is feedback on logical errors that indicates, for example, when a student has not chosen enough statements to correctly prove a statement. Students can also request strategic hints about what they might do next. Such a hint is illustrated in Figure 1. In order to finish the proof, the student needs to find one more set of corresponding angles in order to show △ABC ≅ △ADE using Angle-Side-Angle. The hint shown is the second in a series of hints that start out vague and get more specific. It guides the student to look in the diagram for a statement that might help prove the current subgoal △ABC ≅ △ADE -- in this case, the student is hinted to notice that the crossing lines CE and BD form "vertical angles" ∠DAE and ∠BAC that are equal.

Many students have used ANGLE in laboratory experiments designed both for the purpose of formative evaluation to help improve the system and to collect data on the nature of human problem solving and learning (e.g., see Koedinger & Anderson, 1993). However, the focus of this paper is on our first formative study of ANGLE in a realistic urban high school setting. Prior to this study, we hired a high school teacher to help us in the further development of ANGLE and, in particular, in dealing with the significant curriculum integration issues involved in effectively using an ITS, or any other piece of software, in the classroom.

Integrating ANGLE into the Existing Curriculum

Any effort at innovation in US public schools, must face the reality of multiple and possibly competing innovation efforts of other stakeholders at multiple levels (federal, state, local school board, principal, department head, individual teachers, other researchers, etc.). One particular concern in this case involved a decision by the Pittsburgh public schools to make a major change in their geometry curriculum. Motivated by a consistent high failure rate in geometry (nearly 50% were getting D's or F's), the city decided to take a risk with a new geometry curriculum by adopting a new textbook called *Discovering Geometry: An Inductive Approach* (Serra, 1989) in the fall of 1991. The school we worked with, in fact, had been piloting this curriculum (with a pre-release version of the text) since 1988. This change was of particular relevance to us since this textbook de-emphasizes the role of proof in the geometry curriculum saving most of the discussion of it for the last three chapters of the book. In addition, such a change is not just a local phenomenon. The standards of the National Council of Teachers of Mathematics (NCTM, 1989) have also called for "decreased attention" to deduction. This is not say that deduction has been or should be eliminated from the geometry curriculum nor that a tutor for proof is no longer needed. It has meant that we have had to take a different approach to tutor development and curriculum integration than we have in the past.

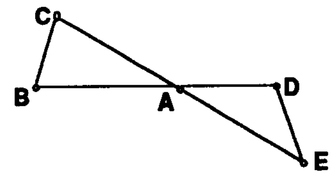
In the context of these changes, we created curriculum materials designed to motivate and justify the study of proof both to students, who often do not understand its purpose (Schoenfeld, 1989; Chazan, 1988), and to our participating teachers, who are committed to fulfilling NCTM standards and to the use of the *Discovering Geometry* textbook. Rather than giving students problems of the form "Prove X" where X is known to be true (as is done in traditional textbooks), we gave students *truth judgment* problems of the form "Is X true?" where X may or may not be true in general. For example, the problem shown in Figure 1 was given to students as a conjecture:

> 11. In all figures like the one shown if you are given that A is the midpoint of segment EC and "BC is parallel to ED, is it *always true* that A is the midpoint of BD?.

Related conjectures using the same starting diagram were given that are not always true:

> 12. In all figures like the one shown if you are given A is the midpoint of segment EC and BC = ED, is it *always true* that A is the midpoint of BD?.

Both types of problems were entered into ANGLE (of course for false conjectures, students could not use ANGLE to successfully complete a proof since no proof is possible). Students were taught to use proof as a tool to help judge the truth of conjectures of this kind -- they used ANGLE to try to find a proof and when they did, they concluded the conjecture is always true. They were also instructed in searching for counterexamples as a tool for conclusively determining when a conjecture is not always true. For example, here is counterexample for problem #12 above:



Students could start by either trying a proof or looking for a counterexample. However, failure to find a proof (or counterexample) is not conclusive evidence that the conjecture is false (or true) -- it simply indicates that the student should try switching strategies and look for a counterexample (or proof) instead. In a traditional geometry course where students are asked to prove statements that are always true, it is perhaps no surprise that they typically view proof as a difficult and mundane task that is just part of fulfilling the requirements of school and that has no particular purpose on its own. They fail to see proof, as mathematicians do, as a tool that helps distinguish what is true from what is not, perhaps, precisely because they are rarely or never asked to use it in such a fashion. Truth judgment problems give them this opportunity.

Because of the inductive slant of ANGLE's cognitive model of deduction, we were perhaps in a better position to integrate it with the inductive approach of the *Discovery Geometry* textbook than we might have been. Nevertheless, teachers liked how the truth judgment problems made the connection between the rule-based deduction process tutored by ANGLE and the example-based discovery process emphasized in their textbook. In truth judgment problems both the rule-based and the example-based search process need to be brought to bear to be successful. In addition to aiding the tutor integration problem locally, truth judgment problems fulfill one of the major themes of the new NCTM standards (NCTM, 1989) by making a *connection* between alternative mathematical representations and strategies.

A Controlled Study of Classroom Use of ANGLE

Method

Three teachers participated in the classroom use of ANGLE in the spring of 1992. Given the potential of differences based merely on teacher effectiveness, every teacher had at least one experimental class that used ANGLE and one control class that did not. One of these teachers was the one who had worked with us on the project. He taught 2 experimental "ANGLE" classes and 2 control "no-ANGLE" classes. The other two teachers taught one ANGLE class and one no-ANGLE class each for a total of 4 experimental classes and 4 control classes.

Students worked with ANGLE during the end of the spring semester for about 20-25 class periods (44 minutes each). Students were given pretests at the beginning of the year and we found no significant differences in the quality of students in the 8 classes. At the end of the semester students took a proof posttest containing 8 proof problems designed to be challenging -- the problem in Figure 1 is representative of a medium difficulty problem. They also took a truth judgment posttest.

The proof posttests were scored using the four point Senk (1983) scheme and this score was converted to either 0 (for 0-2) or 1 (for 3-4) to measure how many proofs students got essentially correct. In other words, minor errors in details could be made and the student could still receive a 1 as long as they had all the key steps of the proof correct. An 8 would be a perfect score.

Results

ANGLE classes averaged almost one proof higher on the posttest (4.10 correct to 3.36 correct in the no-ANGLE classes), but this difference was not statistically significant ($F(1, 60) = 1.81$, $p = .18$). A closer look at the data reveals that the effect of ANGLE was specific to the teacher that participated on our project. Students in the ANGLE classes of the project teacher scored much higher on average than students in any of the other classes – see Figure 2. Not only is this difference statistically significant ($p < .02$), it is also quite large – the mean of the project teacher's ANGLE classes, 5.25, is a standard deviation above the mean of the other three groups, 3.21. Students in these classes were essentially getting two more proofs correct on average than students in the other classes. For the non-project teachers' classes, there is no reliable difference between students in the ANGLE and no-ANGLE classes (2.9 vs. 3.4, $p > .49$).

It is important to note that this result is not an effect of poor scores in the control classes. In fact, students scores appear to compare favorably with Senk's (1983) sample of more than 500 students from both urban and suburban high schools. On a comparison problem, students in the control classes scored significantly better (88% correct) than students in the Senk sample (72%; $X^2(1, 539) = 4.1, p < .05$).

The interaction between effects of teacher and intelligent tutor replicates a finding from the classroom study of the previous Geometry Proof Tutor where two teachers were involved. An independent evaluator of the social effects of this study (Schofield, Evans-Rhodes, & Huber, 1990) reported that the non-project teacher essentially treated the tutor as a replacement and took the time to do things like grade papers. The project teacher, on the other hand, saw the tutor as providing more opportunities to give individualized help (Wertheimer, 1990). While students were engaged in interacting with the tutor, he was free to roam around the classroom giving extra help to poorer students who needed it or challenging better students to do more than they might otherwise. In other words, the individualized help provided by the ITS gave the teacher a greater chance to provide individualized help himself.

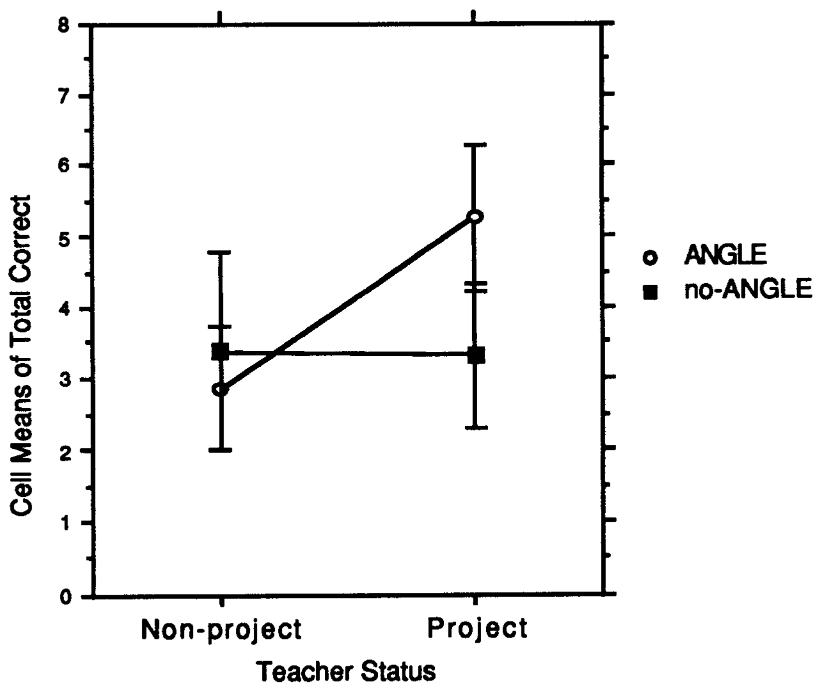


Figure 2. The students in the project teacher's ANGLE classes scored about one standard deviation higher (5.25 proofs on average) than students in other classes (3.21 proofs).

As indicated by the comparison of the control classes with Senk sample, all 3 teachers in the current study were quite effective in teaching students this difficult topic. We did not formally collect the kind of social process data that Schofield, et. al. (1990) did in the previous study. However, from informal observations of these teachers, it did appear that because of his knowledge of the problem solving and tutoring approach of ANGLE, the project teacher engaged in more content-related interactions with students while they worked with the tutor than the other teachers did. One of the non-project teachers, in particular, had little prior experience in computer use and was observed to focus more of his interactions with students on technical problems and advice about the interface. The project teacher moved students quickly past any technical or interface problems to get them back to thinking about geometry.

In addition to being more familiar with ANGLE, the project teacher was more familiar with the theme and content of the Truth Judgment curriculum materials and how they were to support interaction with ANGLE. As evidence of this his classes performed significantly better on the truth judgment test than students in other classes ($F(2, 51) = 6.64, p < .01$). This may be why his ANGLE classes did so much better on proof -- having given students a more meaningful context in which to understand the use of proof (i.e., through the truth judgment instruction), they were better able to take advantage of the proof instruction provided by ANGLE.

It was clear from observing students interacting with ANGLE, that is was a motivating experience. Many students would come to class and begin busily working with ANGLE before the bell rang to start class – this type of independent and active engagement in the subject matter is not typically observed in normal classrooms. A number of students came after school and in one case, a couple of students made up a story to skip another class in order to come into the computer lab to work with ANGLE.

However, there were some clear limitations with the system. One problem we observed is that a few students occasionally took a rather mindless trial-and-error approach to working with ANGLE. This would involve more-or-less randomly entering statements and making justification links until they arrived at a correct step and/or they made enough errors so that the tutor would begin to provide successively more detailed hints as to what the next step should be. The use of this mindless guessing strategy is not particular to ITS use in any way -- it is not uncommon in student attempts to answer teacher questions in normal classes. Nevertheless we are pursuing modifications of ANGLE that might discourage this approach.

Formative Evaluation and Summary of Lessons

Studies of this kind should not be viewed as a final and conclusive statement on the effectiveness of a particular system, like ANGLE, and/or ITSs in general. Instead they should be viewed as a formative evaluation that provides information for making improvements both in the ITS itself and in the instructional context of it's implementation. They also provide lessons that may generalize to other system building and classroom implementation efforts. A major lesson from this study is that there is a very real issue of integrating ITSs with existing curricula – particularly at a time when there are major changes afoot in the curriculum goals and standards for mathematics education in the US (NCTM, 1989). We are now taking a more "client-driven" rather than "technology-driven" approach to the development of new systems. The choice of tutor content in the past was driven largely by what would be a challenging test of the tractability of cognitive theories and AI technology. Now we are treating educators and teachers as clients and getting up-front guidance on what the curriculum should be and how it should best be taught. This approach is being taken in our beginning efforts to build a learning environment for geometry that will integrate mathematical investigation, discovery, and proof as is done in the *Discovering Geometry* textbook (Serra, 1989) that local educators have chosen to best fulfill the NCTM standards. This environment will incorporate 1) sophisticated computer tools for geometric induction (Geometer's Sketchpad, 1989) and deduction (ANGLE's interface) and 2) tutoring facilities to support students in doing and learning geometry more like mathematicians do. Among other things, this system will implement the Truth Judgment curriculum discussed above, that is, by adding to the existing tutoring support for proof, tutoring support for counter-example search and for switching between the two strategies.

A second lesson from this study is that the teacher and the kind of instruction that surrounds the use of an ITS can have serious impact on its effectiveness. Having a system that is more intentionally integrated into the curriculum will help in this regard. In addition, we are seeking guidance from the education community to help us develop materials and workshops that can better prepare teachers to provide the kind of instructional context that makes the full use of the technology's potential.

References

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

Chazan, D. (1988). Ways of knowing: High school students' conceptions of mathematical proof. Unpublished Doctoral dissertation, Harvard Graduate School of Education, Cambridge, MA.

Geometer's Sketchpad (1989). Software by Key Curriculum Press, Inc., Berkeley, CA.

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

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

Lajoie, S. P. & Derry, S. J. (1993). *Computers as Cognitive Tools.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Lajoie, S. P., & Lesgold, A. (1989). Apprenticeship training in the workplace: Computer coached practice environment as a new form of apprenticeship. *Machine-Mediated Learning, 3*, 7-28.

Larkin, J. H., & Chabay, R. W. (1992). *Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Mark, M. A. & Greer, J. E. (1991). The VCR tutor: Evaluating instructional effectiveness. In *Proceedings Thirteenth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates.

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: The Council.

Schank, R. C. (1991). Where's the AI? *Artificial Intelligence Magazine, 12:4*, Winter, 38-49.

Schoenfeld, A. (1989). Explorations of students' mathematical beliefs and behavior, *Journal for Research in Mathematics Education, 20:1*, 338-355.

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, 8:1*, 24-41.

Senk, S. L. (1983). Proof-writing achievement and Van Hiele levels among secondary school geometry students. Doctoral dissertation, Department of Education, University of Chicago.

Serra, M. (1989). *Discovering Geometry: An Inductive Approach.* Berkeley, CA: Key Curriculum Press.

Sleeman, D. H., & Brown, J. S. (1982). *Intelligent Tutoring Systems.* New York, NY: Academic Press.

Wertheimer, R. (1990). The geometry proof tutor: An "intelligent" computer-based tutor in the classroom. *Mathematics Teacher*, April, 308-317.

Acknowledgments

This research was supported by grant 86-17083 from the National Science Foundation and by a James S. McDonnell Foundation Postdoctoral Fellowship to the first author. The computers in the high school were donated by Apple Classrooms of Tomorrow, Apple Computer, Inc. Thanks to the anonymous reviewers for their helpful comments.