Artificial Intelligence in the Classroom: The Impact of a Computer-Based Tutor on Teachers and Students

Janet Ward Schofield, Debra Evans-Rhodes, and Brad R. Huber

This study assessed the impact of utilization of an artificially intelligent geometry proof tutor on classroom social processes. Both teachers' and students' behaviors changed. Teachers devoted more time to their slower students, treated students in a more collegial fashion, and increased their emphasis on effort in grading students. Students showed a marked increase in task-related effort and involvement. This change appeared to be due to an increase both in the students' enjoyment of the class and in the level of peer competition. *Keywords:* artificial intelligence, classroom social behavior, teachers' roles, competition, grading, teacher authority, embarrassment.

The past decade has seen an incredibly rapid proliferation of microcomputers in both elementary and secondary schools. For example, between 1981 and 1984 the number of schools with microcomputers more than tripled (Quality Education Data, 1984). By 1985 almost all secondary schools and five-sixths of all elementary schools in the United States had at least some computers for use in instruction (Becker, 1986), and the trend toward the continuing acquisition of computers has continued. Current estimates are that more than two billion dollars have been spent to provide schools in the United States with computer technology in a period when school systems are under heavy pressure to spend their limited resources on numerous other things, including increased salaries for teachers (Buckley, 1988).

Although the remarkable rapidity with which microcomputers are being placed in schools is obvious, the impact of this change on teachers and students is not. In fact, our knowledge of the way in which this change influences classroom structure and functioning is extremely limited (Sheingold, Kane, & Endreweit, 1983; Sheingold, Martin, & Endreweit, 1987) and some studies suggest that the impact of educational software can be quite different from that which its developers intended (Hativa, Swisa, & Lesgold, 1989). Thus we decided to conduct an intensive qualitative study examining a wide variety of computer usage in a

Social Science Computer Review 8:1, Spring 1990. CCC 0894-4393/90/$1.50.

single school. Such an approach allows exploration of the extent to which different kinds of computer usage have similar or different effects as well as analysis of what the impact of any particular usage may be.

The goal of the part of the study reported here was to explore the impact of one unusual but potentially very important usage of microcomputers—their utilization as intelligent tutors—on classroom structure and functioning. Ideally, intelligent computer-based tutors can follow what a student is trying to do, diagnose the difficulties the student is experiencing, and present instruction relevant to those difficulties, providing individually tailored learning experiences which proceed at a pace determined by the student's capabilities (Anderson, 1984). Thus, perhaps more than any other presently envisioned use of microcomputers, their use as intelligent tutors holds the promise of improving schooling as we know it today. The cost of the development of such software is high, and some of it currently requires expensive hardware to operate. However, there is reason to believe that within the relatively near future the cost of artificially intelligent tutors for educational purposes will no longer be prohibitive (Lesgold & Lesgold, 1984). Thus, the study of intelligent tutoring is the study of a potentially revolutionary educational innovation which is close to being a practical reality from a technical and a fiscal perspective.

It is important to recognize, however, that effective usage of artificially intelligent tutors may well produce or require substantial changes in both teachers' and students' behavior. For example, effective use of intelligent tutors is likely to require greater role change on the teachers' part than the use of traditional drill and practice applications in which computers are often used as sophisticated electronic workbooks and thus fit much more readily into established classroom roles and routines. Unfortunately, at this time, we know little about what these changes might be. Given the rapidity with which computers are becoming commonplace in American schools, such knowledge seems important for two reasons. First, it may be helpful to those attempting to prepare teachers to use computers in their classrooms in a maximally effective way. Second, it may identify unintended side effects of computer usage so that educators can decide if and how to use computers with more complete knowledge of the full ramifications of their decisions.

Research Site and Methods

Data gathering took place during a two-year period (1985–1987) in a large urban high school which serves approximately 1,300 students from varied socioeconomic backgrounds. Approximately 55% of the students were black, 40% were white, and 5% were from other, primarily Asian, backgrounds. The two major methods of data gathering were intensive qualitative classroom observations and repeated interviews with students and teachers.

A team of three trained researchers observed all geometry classes taught by the two teachers utilizing a sophisticated state-of-the-art artificially intelligent geometry proof tutor, called the *GPTUTOR*. This tutor has been described in detail elsewhere (Anderson, Boyle, & Yost, 1985, 1986; Wertheimer, 1988), so we will just discuss it briefly here. The *GPTUTOR* software consists of three parts. The first is an expert system which contains the knowledge necessary for constructing geometry proofs. The second is the tutor which includes information used to teach the students, such as tutoring strategies and messages about student errors. The third part of this software is the interface which allows students to communicate with the computer using either a keyboard or a mouse. The givens of the proof are presented to the student at the bottom of the screen and the statement to be proved appears at the top, along with a diagram of the problem. The student's job is to create a "proof graph" which shows how the givens can yield the statement to be proved (Wertheimer, in press). Various help and review functions are available, either at the student's own request or when enough mistakes have been made that the tutor unilaterally intervenes with help. Due to the underlying philosophy and goals of its developers, the *GPTUTOR* was purposely constructed to interact in a business-like way with the students. Thus it lacks the game-like or "humanizing" elements of many pieces of educational software, although it does indicate success on a proof with both a distinctive sound and a message on the screen.

Observations were made before, during, and after the part of the year in which the tutors were used in a total of seven different classes. Three "control" geometry classes taught in a traditional manner by the two teachers using the computer tutors were also observed. In addition, several comparison classes taught by two other geometry teachers were observed. The classes using the computer tutors ranged markedly in size. Some were small enough so that each student had his or her own computer to work on. Alternatively, students worked in pairs on the ten available machines. The computer tutors were observed in use for roughly 100 hours. Similarly, over 100 hours of observation were conducted in the control and comparison classes.

Observations were conducted using the "full field note" method of data collection (Olson, 1976) which involves taking extensive handwritten notes during the events being observed. Shortly thereafter these notes were taped and transcribed. Field notes were made as factual and as concretely descriptive as possible to avoid unwarranted inferences. However, one major issue with the use of such notes as a database is what Smith and Geoffrey (1968) have termed the "two-realities problem"—the fact that the notes which serve as the basis for the researcher's analysis cannot possibly include everything that has actually transpired. Hence, a source of potential bias is the fact that the reality captured in the field notes is only a part of the fuller reality that the researcher is trying to understand. Further, there is the possibility that selective recording of certain types of events means that the reality captured in the notes is not only partial but biased in some systematic way.

Although the two-realities problem is impossible to surmount completely in qualitative observation, there are some steps that can be taken to minimize its effects. For example, we found it useful to have two researchers observe the same classroom. Discussion of differences between the two observers' notes helped to make the observers aware of their individual biases and preconceptions. Another technique we found useful in reducing the effect of the two-realities problem was to actively seek out data that undercut our developing assessment of a situation. In addition, numerous other techniques discussed in works on qualitative research in educational settings (Bogdan & Biklen, 1982; Bogdan & Taylor, 1975; Goetz & LeCompte, 1984; Kirk & Miller, 1986; Schofield, in press; Schofield & Anderson, 1987) were utilized in a systematic attempt to provide as unbiased, reliable, and representative a view of what actually occurred in those classrooms as possible given the qualitative and exploratory nature of the study.

The second major data-gathering technique employed in this study was repeated interviewing of students and teachers. All of the students in the classes using the *GPTUTOR* were invited to participate in "pre" and "post" use interviews. Over 90% of these students and 82% of a group of control students randomly selected from other geometry classes actually participated in the interviews. These 45 minute structured open-ended interviews were taped and transcribed. Both formal and informal interviews with the two teachers using the computer tutors, Mr. Adams and Mr. Brice,[1] and the teachers of the comparison classes were conducted throughout the course of the research. To get yet an additional perspective on the computer tutors, interviews were also conducted with numerous individuals closely connected with the development of the tutor and its field test in the school setting, including the project leaders, programmers, and the individual whose job was to oversee the field testing itself. In constructing and conducting these interviews, strong efforts were made to procure valid and unbiased data. For example, questions were posed in a balanced manner so that leading questions were avoided, students were assured that their teachers would not have access to their interview transcripts and the like.

Although observation and interviews were the primary data-gathering techniques utilized, other techniques were employed when appropriate. For example, archival material such as letters sent to parents about the computer-tutor, internal school memoranda and announcements, and copies of the student newspaper were collected and analyzed.

Data Analysis

Briefly describing data analysis procedures in qualitative research is extremely difficult since the process is so complex and iterative. To summarize, observational notes were coded as described in sources like Miles and Huberman (1984) and Strauss (1987). This process involves carefully reviewing field notes as they are collected, creating coding categories of various types, developing and refining coding systems, writing working memos, and then searching for ways to refute or enrich the ideas emerging from the preceding activities. Interviews were analyzed using traditional content analysis procedures.

Three general principles guided both the data-gathering and the data analysis phases of the research. First, a concerted effort was made to be as rigorous and systematic as possible. For example, sampling techniques were employed where appropriate; trained coders coded the open-ended interviews using reliable systems developed for this research; and field notes were carefully indexed so that all notes relevant to a given topic could be examined. Second, we took the importance of triangulating the data quite seriously (Webb, Campbell, Schwartz, & Sechrest, 1966). That is, great care was taken to gather many different types of information bearing on the same issue, to minimize the potential problems with each data source, and to be sensitive to biases which could not be completely eliminated in analyzing and interpreting the data. The third general principle which we took seriously was that data analysis should be an on-going and iterative process. As the field notes and other data accumulated, they were indexed, read, and reread. Informal working memos were written, and data relevant to ideas emerging from the early stages of analysis were actively sought in planned and systematic ways.

Findings

Utilization of the *GPTUTOR* appeared to result in a number of important changes in both teachers' and students' behavior. Three changes appeared most marked in the teachers' behavior. These included a change in the relative amount of attention given to different kinds of students, an increase in the extent to which the teacher functioned as a collaborator with the students rather than as a more distant authority figure, and a change in grading practices. There were also a number of important changes in students' behavior, most notably increases in task-directed effort and peer competition.

Changes in the Teachers' Behavior

A shift in the amount of teacher attention devoted to different types of students. The introduction of the computer tutors appeared to change the relative amount of attention given to students of different ability levels. More specifically, it increased the amount of time devoted to those having problems. Before the arrival of the tutors, and in the control and comparison classes, teachers often had students work through geometry problems and proofs on the board. Another commonly used teaching method was to work through problems by having students who were seated volunteer answers to the teacher's questions. Not surprisingly, in these situations teachers tended to call disproportionately on the more advanced students, as previous research has suggested is often the case (Bossert, 1979). This saved considerable time, raised the probability of a correct answer, and saved the poorer students the embarrassment of making mistakes in public. The problems posed by waiting for a slower student to supply an answer which other students have already figured out are made clear in the following excerpt from the field notes from a geometry class in which two able students challenge the fairness of Mr. Adams's usual behavior:

> The teacher says, "Ready? Okay, What's the answer? . . . " Tim answers the question correctly. Mr. Adams says to Tim, "You get the extra credit!" . . . Ida says heatedly, "That's unfair! You always call on Tim for extra credit! . . . His hand was up first, but he gets all the credit." Allie chimes in complaining too. Mr. Adams doesn't answer the complaints directly. Instead he assigns another extra credit problem and says, "Ida will choose who answers this time." Allie and Pete finish first. They have their hands up. Mr. Adams says, "Okay. Choose." Ida replies . . . "I want to call on one of those," pointing over to Debbie and Katy, clearly the two slowest students in the class. Both of these girls have their heads bent over their papers, still working. Debbie says to Mr. Adams, "Can I ask a question?" He answers it. She continues to work. . . . Ida says, "Are you ready Katy?" (She is not.) Mr. Adams says, "The bell is going to ring any minute and no one will get the credit . . . Time is going." Ida walks up to the front of the room where Mr. Adams stands. Mr. Adams says, in a loud voice close to a shout, "WOULD YOU PLEASE CALL ON SOMEONE SO I CAN GIVE THE CREDIT OUT!" Ida hesitates and calls on Debbie. Debbie gets 7 angles right, but the 8th (the most difficult which was the real point of the example) is wrong. Pete volunteers the correct answer. Mr. Adams lets him show the class how he got to his answer saying, "This is the one you all missed, so I want you to watch. . . . " As the class files out (but is still within earshot), Mr. Adams says in a very biting tone to Ida, "A wonderful teacher you'd make!" Ida defends herself saying heatedly, "Sometimes it's not having the right answer. It's having a chance. If you give her a chance. . . . " Mr. Adams interrupts saying, "Here are people having difficulty. You . . . focus all the attention on them. Isn't that embarrassing. It puts them in a corner." Ida says, "Okay; Okay. But why don't you even call on them?" Mr. Adams replies, "You need to learn something about people. They get it wrong. They make bad subtraction errors."

When using the computer tutor, in contrast, the slower students often received considerably more attention than the brighter ones. Mr. Adams's comments in an interview suggest that he was well aware of this change:

> Interviewer: Have you noticed that [when using the computer tutors] you're giving [more] attention to certain sorts of students versus other ones? Or does it pretty well even out?
>
> Mr. Adams: No, I would say in general I give much more help to students that are much slower and need the help. . . . I'm giving more help

> to them than I ever was able to in the past. . . . A lot more time . . . [The computer] frees you up for individualized attention . . . knowing that the rest of the class is doing something constructive.

Such attention was not likely to be embarrassing because students working on their computers were often unaware of exactly with whom the teacher was working. In addition, as Mr. Adams pointed out, since the *GPTUTOR* provided a substantial amount of help, the teacher's working with the slower students did not impede the rest of the students as much as it would under a more traditional whole class method of instruction since the remainder of the class could continue to work uninterrupted.

A shift in the teachers' role towards becoming a collaborator. A second shift in the teachers' role behavior was also apparent in the computer tutor classrooms. Specifically, the teachers functioned less as authoritative experts and more as collaborators than they had previously. This shift was beautifully captured by the words of one student who was asked in an interview about whether using the computer tutor had changed his teacher's behavior. He replied, "He doesn't teach us any more. He just helps us."

What is this distinction between teaching and helping? The teaching role as it often appears to be defined in high school consists of rather formally imparting a body of facts to less knowledgeable individuals through lectures, structured class discussions, and problem solving. The teacher's separate and superior status is well symbolized by his or her physical position—typically standing above and in front of students who are expected to be watching or listening carefully. In the control and comparison classrooms, the teacher's authoritative position was made clear by the common practice of calling upon students to answer questions or work problems at the board. In doing this, the teachers exercised control over the class not only by choosing between students who indicated a desire to participate, but also, less commonly, by calling on students who would have preferred not to become the focus of the class's attention.

In contrast, in the computer tutor classes, as in many of the other classes we observed in which computers were used a great deal, the teacher functioned more as a collaborator than was typical under other circumstances. Specifically, rather than addressing the entire class in a relatively formal manner, the teacher tended to work on an individual basis with students. Just as importantly, in the computer-using classes the teachers were much less likely to initiate teacher-student interactions. Rather, they were kept busy responding to student requests for assistance, thus shifting control for initiating teacher-student interactions into the hands of the students.

An increased emphasis on effort in grading. Finally, the utilization of the computer tutors also led to potentially important changes in the teachers' grading practices. Mr. Adams and Mr. Brice, like the other geometry teachers, usually based the grades they gave upon some formal criteria involving an objective level of performance on homework assignments, quizzes, tests, and the like. However, when using the computer tutor they both made a change. Specifically, both decided independently to emphasize effort more than they had previously. Mr. Brice explained it this way:

> Interviewer: Did having the computers change the basis on which you assigned grades at all?
> Mr. Brice: Well, I did give them a grade for the lab (computer) work they did. The control class didn't get a grade for the lab work because there wasn't any lab to work in.
> Interviewer: Was it how much they accomplished, or how hard they worked, or some combination?
> Mr. Brice: Probably how much they stayed on task. . . . Breaks here and there would affect their lab time. It was how much time they spent on task not necessarily how much they learned.

Since one of the major advantages of the computer tutor was that it allowed students to work at their own pace, grading everyone against the same standard of accomplishment no longer seemed consistent with the way the class was structured. It is particularly interesting that Mr. Adams adjusted his grading system to reflect effort even though he was philosophically opposed to this:

> Interviewer: How has the introduction of the computer tutors changed the basis on which you assign grades?
> Mr. Adams: THIS IS A PROBLEM!!! . . . Oh my God, yes, how do I grade them?
> Interviewer: That's what we were wondering. How?
> Mr. Adams: I've had to develop a policy. . . . If they came in and started on the tutor they had a grade of C on the tutor. If they came in and worked everyday and made a legitimate effort, they'd go up to a B. If they came in and made a half-assed effort they'd go down to a D. If they came and didn't give a damn at all they'd go down to an E. If they came in and really knocked their socks off and showed me that they really cared and learned, and learned, then they'd get an A. . . . Effort meant a lot more this time. It had to. See, I'll be honest with you. . . . I just don't buy effort. It just doesn't mean much to me. It doesn't. I mean, you need effort . . . but I'm a geometry teacher. . . . A college is going to assess a student's ability according to that grade. . . . So, I just can't give a B for effort. . . . This is my philosophy. . . . I'm trying to be a bottom line so we don't sell them some bill of goods, like . . . A or B students [who] take the SAT exam and pop out a 450 combined, which happens all the time in the city. . . . So that's why this is a little unusual, because effort is going to count.

Interestingly, the chairman of the math department mentioned to our project staff that he could not evaluate the teachers using the intelligent tutors very well since those classes were run so differently from ordinary ones and different teaching skills were needed. Thus, the utilization of the tutor raised questions about both teacher and student evaluation.

Changes in the Students' Behavior

Increased effort and involvement. There were also significant changes in student behavior which accompanied the introduction of the computer-tutor. One of the most striking changes in the classrooms using the GPTUTOR was the increase in student involvement and effort. This change was evidenced by markedly increased time on task, clear increases in apparent level of concentration, and the like.[2] In fact, when asked in the post-use interviews how using the computer tutor influenced their behavior, students most commonly mentioned an increase in their level of effort. Both of the teachers utilizing the computer-based tutor also spontaneously mentioned this increase in student effort in their interviews. It seems reasonable to suggest that this change may have contributed to the positive impact of the computer-tutor on students' ability to do geometry proofs, which is discussed elsewhere (Wertheimer, in press).

Before the arrival of the computer tutors, and in the control and comparison classes, it generally took the teacher a few minutes to get the class settled down and ready for work. Mr. Adams generally started his classes more promptly and continued them right up to the closing bell more than the other geometry teachers we observed. However, even students in his classes made it difficult to use all of the 45-minute class period productively when they were not using the GPTUTOR as is evident from the following excerpt from our field notes:

> I [the observer] arrive before the bell [which indicates the start of class] rings. Kathy, Tim and Debbie are in their seats already. The two girls are consulting about their homework. . . . The bell rings. Karen saunters in and takes her seat. Ida comes in after Mr. Adams begins to ask for homework saying, "Pass it in if you've got it. . . . " Rachel strolls in while Mr. Adams puts the first problem on the board. He ignores her. Karen says, "Are we almost done with these circle problems? I hate them." Ida is looking at a postcard which Karen has shown her. Mr. Adams says, "Ida, put that thing away and pay attention." Ida looks at the postcard for another 15 seconds or so and then hands it back to Karen who holds it to her nose and inhales happily before putting it away. Katy complains, "It's so hot in here today." Mr. Adams gives the students a problem to work on. Ida and Rachael remove their jackets.

In many of the control and comparison classes it was not unusual for the last five or ten minutes to be devoted to socializing. Sometimes the teacher had simply covered all he wanted to and judged that it did not make sense to start something new in the time remaining. Another common practice which contributed greatly to lost instructional time was giving out a "homework" assignment well before the end of the class period and asking students to work on it in class. Although the rationale for this practice was that the teacher could help students with any difficulties they might encounter, many students chose to chat with friends for ten or fifteen minutes rather than doing this "homework" in class. Thus, in many classes which did not use the computer tutors it was common for a substantial proportion of the students to spend a total of ten to fifteen minutes a period chatting about sports, clothes, teachers and the like. This accounted for a substantial proportion of the 45-minute class periods.

Almost immediately after beginning to use the GPTUTOR, many students began working on their proofs well before the starting bell, a situation virtually never observed in the control and comparison classes. In addition they frequently continued working after the bell, also very atypical behavior in other geometry classrooms. In one extreme case, a fistfight nearly broke out between one student staying after his class was over to try to finish a proof and another who arrived early for the next class and wanted to get started on the same machine. An excerpt from the project's field notes illustrates the unusually prompt start of the classes using the computer tutor:

> By the time the bell to start class rings, three-fourths of the students in class today have problems on the screen and are working on them. The others (have all logged on and) appear to be waiting for their problems to appear. . . . I'm struck by the fact that the students have started their work without a word from the substitute teacher who is in charge of class today.

In addition to starting work more promptly and working through the last minutes of the period, students using the GPTUTOR also appeared to be more engrossed in their work than those in the traditional classes. Of course, in observational work it is difficult if not impossible to get a precise record of whether students are concentrating since students skilled in the art of classroom behavior may find many ways to appear to be working when in reality letting their attention wander. Yet all indications were that the level of concentration rose. Students often spontaneously mentioned this change in interviews:

> Interviewer: How did using the computers change the way you behaved in class?
> Diane: Well, we didn't talk as much. On the computer you really concentrated on the screen—didn't have time to talk to the person next to you.

The fact that students made much quicker progress on the proofs than either of the teachers anticipated suggests that the students were indeed focusing their attention on the proofs. This increased focus is apparent from field notes like the following, as well as student and teacher interviews:

> The room is extremely quiet now (as students continue to work at the computers). There are just little beeps (from the machines) every minute or two. Except for one brief interchange between two white boys

there has literally been no student to student talk in the 15 minutes of class so far.

Factors contributing to the increase in effort and involvement. Since several sources of evidence all converged to suggest that the computer tutor enhanced student effort and involvement, an obvious question that arises is "Why?" The data suggest several complementary reasons for the change, including an increase in competition between classmates and an increase in the students' enjoyment of their work.

As discussed earlier and consistent with other research (Hawkins & Sheingold, 1986), a change in grading practices followed introduction of the computer tutors into the classroom. Specifically, both teachers began to count effort more than they had, in spite of the fact that one of them was philosophically opposed to this practice. One possibility is that this accounted for the students' increased effort. Yet the evidence suggests that this was not a major contributing factor. Specifically, only two of the over seventy students interviewed about how using the *GPTUTOR* influenced their grades said they thought their level of effort on the computer contributed to their grades. A few additional students remarked that persistent and unnecessary use of a software feature called "system select," which essentially presented the student with the problem's solution, would be viewed negatively by their teacher, although they did not specifically indicate that it would hurt their grades. Fewer than 10% of the *GPTUTOR* students who were asked why the computer tutor kept track of all their work (which was done primarily for research purposes) believed that this record might be used in grading. Instead, they typically said it would be used to isolate areas in which students needed more help or to see if students were using the system select feature so much that they would not learn anything. Thus, overall, the change in the teachers' grading practices does not seem to be a major factor accounting for the widespread and striking increase in the students' effort and involvement.

Increased competition. A more tenable explanation for the students' increased effort and involvement was the unusually high level of competition which developed between students. Although the extent of this increase in competition varied somewhat from class to class, it was apparent in all the computer tutor classes. In fact, in some computer tutor classes a high proportion of the relatively infrequent student conversations concerned a comparison of how many of the available problems the students had completed.

The teachers, students, and observers all noticed the change in the level of competition. Both Mr. Adams and Mr. Brice remarked on it in interviews:

> Mr. Adams: Just listen to them. Just watch them as they're waiting for a problem to come up. They say, "What problem are you on? "Where are you at?" "Oh, you're only there. Oh, I was there two days ago!" That kind of stuff.

When the students were asked in the post-computer use interviews why they had started getting to class early after the arrival of the computer tutors, roughly 40% spontaneously indicated that it was because of the competition. In addition, when asked directly whether the introduction of the tutors had changed the level of competition in the classroom students overwhelmingly responded in the affirmative. Well over two-thirds of the *GPTUTOR*-using students perceived more competition. Comments like the following were common:

> Interviewer: Did geometry class seem more or less competitive when you were using computer tutors?
> Mike: It was more competitive. We went at each other. . . . It was like we were having a race—who would get to the end of each chapter in the book through the computer. I won . . . because I knew it!

Rather ironically, the very fact that the tutors were designed to let students progress at their own pace created a situation which fostered competition between students. In traditionally taught geometry classes students never have the chance to get far ahead of or behind each other in the way they can in a class using computer tutors. All of the geometry teachers we observed typically began with a brief lecture or a discussion of specific geometry problems, either newly presented ones or problems from the previous night's homework. Some students were able to do these problems faster or more accurately than others, and in that sense the potential for competition existed. However, since the teachers normally kept the class focused on a particular problem or set of problems until they felt most of the class understood what was happening, the stronger students were not able to move through the curriculum faster than their peers. They could differ from their peers in the quality of their work, but they could not get substantially ahead. Similarly, although the slower students sometimes got lost on particular problems, they were soon presented with a different one which they had at least some chance of solving. Thus, the daily situation was not conducive to intense competition since the opportunity for pulling dramatically ahead of their peers was not there.

In contrast, when students used the computer tutors some students were able to progress much faster than others. The faster students were not held back by their teacher's desire or need to teach to the class as a whole. Neither could the slower students skip over the problems which gave them difficulty in one area and hope to do better in another, for the software was organized into a series of problems which were to be solved in a specified order. The fact that the problems were numbered and that students were seated close enough to each other that they could talk without shouting also encouraged the development of competition because it made it easy for the students to communicate simply and clearly about exactly how far they had progressed. The kind of interchange illustrated in the following field notes from Mr. Brice's class was common:

> Dan says to Val, who is at the computer next to him, "What number are you on?" Val replies "Fifty-two." Dan says "Shit, we're on 41." Milton . . . hears this and laughs.

Increased enjoyment. Students may have been motivated to work harder when using the computer tutor simply because they enjoyed using the computers. The large majority of students indicated in interviews that using the GPTUTOR was more fun than learning conventionally and some specifically linked this increased enjoyment to an increase in motivation. Prior theory and research (Malone & Lepper, 1987) suggest the kind of linkage between enjoyment and motivation implied by Paul's comments below:

> Interviewer: What do you think are the major advantages of using a computer to help you learn geometry?
> Paul: If it's fun, it makes you want to learn something! It's fun!

The students' increased enjoyment of geometry when using the computer tutors had numerous sources. First, many students expressed pleasure at their relative independence from direct adult control when working on the computer. Furthermore, there was a strong link in many students' minds between computers and playing games which predisposed them favorably to working on computers and led many of them to work in a productive but playful manner (e.g., seeing who could take the most steps to complete a proof successfully) in spite of the competition.

> Mr. Brice looks next at Ben and Marcus at the computer next to Andy. He says, "Whoa, guys. What are you doing there?" Mark says, "We're making a road map," referring to the very complicated proof graph on their screen which has resulted in a large array of lines. Mr. Brice . . . asks them what they are trying to prove and comments, "Well, you've taken a roundabout way." Marcus says, "Well, we like to get a lot a steps in there!"

One feature of the software which may have contributed to this sense of playing while working was a "success sound" which occurred when students completed a proof. Considerably more important appeared to be the sense of personal challenge that students felt which is also a characteristic feature, as is competition, of many games. Numerous students remarked on this sense of challenge in their interviews.

In addition, rather ironically, students responded positively to the computers because they felt free to express frustration and anger to them in a way they could not with a teacher without violating strong norms. Specifically, it was fairly common for students to tell the computer to shut up, to call it names like "stupid fool" or even to swear at it. In fact, in one class one student swore at the computer so much that her teacher joked about wiring the computer so that it would shock her every time she swore at it. Sometimes, though not often, students would hit or shake the computer or slap the mouse down very hard in spite of the instructions the teachers had given at the beginning of the computer use period about the proper treatment of the machines. This kind of verbal and physical venting of frustration was virtually non-existent in interactions directed from students to teachers at this school. (Not surprisingly, it was explicitly forbidden in the *Student Handbook*.) However, many students clearly felt free to speak much more freely when dealing with the machine-based tutors, and teachers generally let at least the verbal abuse of the computer pass as a kind of amusing expression of emotion. The difference between what is acceptable when students are dealing with the computer and with an adult in a position of authority is made clear in the following field notes:

> Tara says, "This computer should be shot." Mr. Adams squats down and begins explaining the problem to her. Bob says, "Shit" to his computer. Ms. Donavon (who, as part of the staff involved with the field testing of the computer, assists students with bugs in the program or hardware failures) overhears him and thinks he said, "Sh," possibly to her since she's standing nearby talking to one of his classmates. She says, looking at Bob, "What was that?" The blood rushes to his face and he doesn't answer. She says to him, "Why are you turning every shade of red?" Then the light dawns and she says, "It wasn't "sh." It was . . . " She then gives a small laugh and Bob looks down at his terminal and types away industriously.

Another factor contributing to students' involvement with work on the computer and their feeling that it was fun was most likely a sense of being more comfortable in the computer tutor classrooms because of a decreased fear of embarrassment. Students generally agreed that working on the computer tutors was less likely to be embarrassing than doing geometry the more traditional way. The following is typical of the students' responses to questions about this issue:

> Interviewer: Was it more or less embarrassing to make a geometry mistake when you were using the computers than it was before?
> Alice: I think it was less because you were working all by yourself. If you were answering a question (in a traditionally structured class) and you answered it wrong, the whole class would know. When you were working with the computer, nobody really knows.

In traditionally structured geometry classes, students are often called upon to perform before others, as the teacher has them do board work, answer questions at their seats, and the like. This can be embarrassing for the students who are behind or just plain lost since their difficulties are often very public. Even seat work can be embarrassing because teachers commonly comment on it in a tone clearly audible to the rest of the class, as is apparent in the following field notes taken by an observer seated in the back of the classroom:

> Students begin working quietly at their seats. Mr. Adams walks around. . . . He says, "Good," to Sally when he sees her work. He says, "Wait a minute," to Irene and "No, No, No!" to Linda after looking at

her work. He then goes back over and looks at Sally's work again and says, "Wait a minute, Wait a minute, Sally. What's 180 plus 4?"

Mistakes on the computer tutor were more likely to be private, since the computers were placed so that students could not see each other's screens. In addition, although the teachers did circulate and make comments on students' work at the computer, the fact that students were facing different directions as well as working on different problems made it less likely that others were monitoring these comments. Furthermore, the fact that the tutor had a number of help functions which let the student ask the computer for a review of previous material or hints on how to do the problem meant that students who were particularly sensitive about appearing to need assistance had a readily available non-human source of help. It is interesting to note, and consistent with our argument, that one of the few features of the computer which drew frequent criticism from the students was the use of a beeping sound to indicate a mistake. However, there are two reasons that this beeping may not have been overly inhibiting. Specifically, beeping could be triggered by things other than student errors, and even when a student made an error other students were often unaware of whose computer had beeped. This freed students to continue working on the computers even when they were not sure they were correct, rather than stopping out of fear of making a mistake which would result in a beep.

Summary and Conclusions

In summary, both teachers' and students' behaviors appeared to be influenced in important ways by the utilization of the computer tutors. Teachers began to devote more time to the slower students. They also began to act in a somewhat more collegial fashion and increased their emphasis on effort in grading students. Students showed a marked increase in task-related effort and involvement. This change appeared to be created by a wide variety of factors including a major increase in the amount of peer competition and greater enjoyment of their geometry classes.

Since artificially intelligent tutors are still the exception rather than the rule in present day classrooms, many readers may wonder about the implications of this study for understanding the likely impact of more common kinds of educational software, such as educational games or drill and practice packages. Because educational software is so varied, and much of it differs in many ways from the software studied here, it is hard to draw any certain conclusions on this topic. However, we believe this study does have some implications for those interested in more traditional kinds of CAI programs.

On the issue of whether CAI programs are likely to produce the specific kinds of changes documented in the classroom using the *GPTUTOR*, we would argue that the likelihood of similar changes is related to the similarity of the software along certain crucial dimensions. For example, since the increase in student motivation and effort appeared to be related to, among other things, a lessened sense of embarrassment at mistakes, one might hypothesize that CAI programs that encourage the sense that one's mistakes are private might be more conducive to enhancing student effort than those which do not.

One very practical implication of this study for developers and users of educational software, be it traditional CAI or more advanced ICAI software, is the importance of looking beyond the obvious questions of whether and how such software will improve students' learning. Specifically, much more attention needs to be given to assessing its actual impact on a broad array of teacher and student behaviors. For example, the developers of the *GPTUTOR* had no idea that the self-pacing feature would influence teachers' grading practices. Neither, to our knowledge, has there been much consideration of this issue by most school systems adopting CAI for classroom use. Yet if the self-pacing characteristic of many popular CAI programs influences teachers' grading practices, as it did those of the teachers using the artificially intelligent *GPTUTOR*, this is an issue which merits real attention.

It is becoming clear that students may use both traditional CAI software and the kind of software studied here in rather different ways than its creators intended. For example, no one anticipated the playful elaboration of unnecessarily complicated solutions to proofs which some students engaged in when using the *GPTUTOR*. Although this kind of deviation from expected usage was constructive, not all deviations were. For example, in a different part of the broader study on which this paper is based, students were observed using a computer-based educational game which was supposed to encourage them to learn through using a variety of reference tools. They played this game for months before actually using even one of these tools, since they were too excited and involved in the game to want to bother to interrupt their play to consult these sources (Schofield, 1989). A study by Hativa, Swisa, and Lesgold (1989) also demonstrates the slippage between developers' expectations and actual practice in CAI software. Specifically, this study demonstrated how a widely used piece of CAI software designed to encourage students to learn mathematics in an individualized and non-competitive manner actually appeared to encourage competition. Thus, the study reported here is part of an emerging literature which suggests the importance of classroom-based field tests of educational software designed so that they are sensitive to the unexpected.

In conclusion, we would like to emphasize that we do not contend that all the changes documented in this study will necessarily follow the introduction of computers, or even the *GPTUTOR* software itself, into any and all classroom environments. Indeed, recent papers have pointed out the futility of thinking of classroom computer usage as if it were a conceptually satisfying independent variable (Lepper & Gurtner, 1989; Schofield & Verban, 1988). The effect of computer usage will undoubtedly depend on a plethora of factors including the kind of software used (e.g., drill and practice, simulations, tutoring, etc.), the kind of students using the software, the social and physical context of the computer usage, and the ways in which it changes students' learning experiences. We have focused on the changes which occurred in the classrooms studied to indicate that important and often unplanned changes in student and teacher behavior are likely to occur when new technology is introduced and to suggest that greater attention needs to be devoted to understanding precisely what these changes are.

Notes

Janet Ward Schofield is a professor of psychology and a senior scientist at the Learning Research and Development Center at the University of Pittsburgh. She and David Verban have co-authored a paper on teachers' reactions to instructional computers entitled "Computer Usage in the Teaching of Mathematics: Issues Which Need Answers" which was published in 1988 in Dr. Grouws and T. Cooney (Eds.), *Effective Mathematics Teaching*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 169–193. She is now working on a book entitled *Computers in the Classroom* which explores the impact of a wide variety of different kinds of computer usage on classroom structure and functioning. Debra Evans-Rhodes is a graduate student in social psychology at the University of Pittsburgh. She has been associated for several years with the project which resulted in this paper. Brad R. Huber is an assistant professor of anthropology at the College of Charleston. He is currently planning a study on computer usage in Costa Rican schools.

Correspondence should be sent to Dr. Janet Ward Schofield, 816 LRDC Building, University of Pittsburgh, Pittsburgh, PA 15260. The research reported here was funded by the author's Contract #N14-85-K-0064 with the Office of Naval Research. However, all opinions expressed herein are solely those of the authors and no endorsement of the conclusions by ONR is implied or intended. The authors would like to express appreciation to the students and faculty at Whitmore High School for their participation in this research. Thanks also go to John Anderson and Franklin Boyle for their generous cooperation with this independent study of the impact of their computer-based tutor on classroom social processes. Finally, our sincere thanks to Dave Verban for his capable assistance in the data collection stages of this project.

1. For the sake of maintaining the confidentiality of individuals participating in this research, all names used are pseudonyms.

2. Although this phenomena was most obvious in the computer tutor classes, positive motivational consequences associated with computer usage were visible in a great many other settings in the school. (Over 200 hours of observation were also devoted to these other settings which included business classes in word processing and accounting, computer science classes, and classes in the visual arts.)

References

Anderson, J. R. (1984). *Proposal to the Carnegie Corporation to support demonstration and development of a geometry tutor.* Unpublished proposal, Carnegie-Mellon University, Pittsburgh, PA.

Anderson, J. R., Boyle, C. F., & Yost, G. (1985). The geometry tutor. *Proceedings of the 9th International Joint Conference on Artificial Intelligence* (pp. 1–7), Los Angeles.

Anderson, J. R., Boyle, C. F., & Yost, G. (1986). The geometry tutor. *Journal of Mathematical Behavior*, 5, 5–19.

Becker, H. J. (1986). *Instructional uses of school computers.* Reports from the 1985 National Survey, Issue 1, Center for Social Organization of Schools, The Johns Hopkins University, Baltimore, MD.

Becker, H. S., & Greer, B. (1960). Participant observations: Analysis of qualitative data. In R. N. Adams & J. J. Preiss (Eds.), *Human organization research* (pp. 267–289). Homewood, IL: The Dorsey Press.

Bogdan, R. C., & Biklen, S. K. (1982). *Qualitative research for education: An introduction to theory and methods.* New York: Allyn & Bacon.

Bogdan, R. C., & Taylor, S. J. (1975). *Introduction to qualitative research methods: A phenomenological approach to the social sciences.* New York: Wiley.

Bossert, S. T. (1979). *Tasks and social relationships in the classroom.* New York: Cambridge University Press.

Buckley, W. M. (1988). Computers failing as teaching aids: Heralded revolution falls short due to lack of machines, training. *Wall Street Journal,* p. 17.

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and qualitative design in educational research.* New York: Academic Press.

Hativa, N., Swisa, S., & Lesgold, A. (1989, March). Competition in traditional CAI: Motivational, sociological and instructional-design issues. In Andrea di Sessa (Chair), *Computers and classroom social processes.* Symposium conducted at the meeting of the American Educational Research Association, San Francisco.

Hawkins, J., & Sheingold, K. (1986). The beginning of a story: Computers and the organization of learning in classrooms. In J. A. Culbertson & L. L. Cunningham (Eds.), *Microcomputers and education.* Chicago: The University of Chicago Press.

Kirk, J., & Miller, M. K. (1986). *Reliability and validity in qualitative research.* Beverly Hills, CA: Sage.

Lepper, M. R., & Gurtner, J. L. (1989). Children and computers: Approaching the twenty-first century. *American Psychologist, 44,* 170–178.

Lesgold, A. M., & Lesgold, S. B. (1984). *Classroom computers and state curriculum policy.* Unpublished manuscript, University of Pittsburgh.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivation for learning. In R. E. Snow & M. C. Farr (Eds.), *Aptitude, learning, and instruction: III. Conative and affective process analyses.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods.* Beverly Hills, CA: Sage Publications.

Olson, S. (1976). *Ideas and data: Process and practice of social research.* Homewood, IL: The Dorsey Press.

Quality Education Data (1984, January). *Microcomputer data.* Unpublished raw data presented to the Naval Materials Council, Dallas, TX.

Schofield, J. W. (in press). Generalizability in qualitative research. In E. Eisner & A. Peskin (Eds.) (in press). *Qualitative inquiry in education.* New York: Teachers' College Press.

Schofield, J. W., & Anderson, K. A. (1987). Combining quantitative and qualitative components of research on ethnic identity and intergroup relations. In J. S. Phinney & M. J. Rotheram (Eds.), *Children's ethnic socialization: Pluralism and development* (pp. 252–273). Newbury Park, CA: Sage.

Schofield, J. W., & Verban, D. (1988). Computer usage in the teaching of mathematics: Issues which need answers. In D. Grouws & T. Cooney (Eds.), *Effective mathematics teaching* (pp. 169–193). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schofield, J. W. (1989, March). *Computers and classroom social processes.* Paper presented at the conference on Intelligent Computer-Assisted Instruction, Orlando.

Sheingold, K., Kane, J., & Endreweit, M. (1983). Microcomputer use in schools: Developing a research agenda. *Harvard Educational Review, 53,* 412–432.

Sheingold, K., Martin, M. W., & Endreweit, M. E. (1987). Preparing urban teachers for the technological future. In R. D. Pea & K. Sheingold (Eds.), *Mirrors of minds: Patterns of experience in educational computing* (pp. 67–85). Norwood, NJ: Ablex.

Strauss, A. (1987). *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press.

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