Barry P. Goettl  Henry M. Halff  
Carol L. Redfield  Valerie J. Shute (Eds.)

Intelligent Tutoring Systems

4th International Conference, ITS '98  
San Antonio, Texas, USA, August 16-19, 1998  
Proceedings
Successful Use of an Expert System to Teach Diagnostic Reasoning for Antibody Identification

Philip J. Smith, Jodi Heinz Obradovich, Stephanie A. Guerlain, Sally Rudmann, Patricia Strohm, Jack W. Smith, John Svirbely, Larry Sachs

Cognitive Systems Engineering Laboratory, The Ohio State University
210 Baker Systems. 1971 Neil Avenue, Columbus OH 43210-1271
philw@osu.edu

Abstract. A previously reported study indicated that, when used by an instructor as a tool to assist with tutoring in a class laboratory setting, use of the Transfusion Medicine Tutor (TMT) resulted in improvements in antibody identification performance of 87-93% (p<.001). Based on input from teachers requesting that TMT be designed for use without the presence of an instructor, a new study on the use of TMT without instructor assistance found that performance improved by 64-66% (p<.001). Finally, based on the results of these two studies, TMT was mailed to 7 sites for beta-testing. In exchange for a free copy of the kit, the instructors (and their students) were asked to fill out questionnaires. Results of these questionnaires are summarized.

1 Introduction

This research project, involving the design, development, and empirical evaluation of a computer-based tutoring system, was highly interdisciplinary in scope, requiring consideration of aspects of artificial intelligence (specifically, expert systems), education, psychology, and human factors engineering, as well as the domain of study (i.e., allo-antibody identification).

More specifically, the focus of this project has not been on the invention of new features for an underlying expert system, although two especially interesting forms of expertise have been noted in the development of the system:

1. The use of initial data to form an abstract model of the solution in order to decompose the task into simpler subproblems ("It looks like I have a two antibody problem with one antibody accounting for these reactions and a second accounting for the others");
2. The use of self-directed models of potential errors ("In a situation like this, it's likely I erroneously ruled out a weakly reacting antibody") to direct attention when an impasse has been reached [1].

Thus, from the perspective of advancing the design of the expert systems underlying tutoring systems, the project is primarily an application that helps to demonstrate that a rule-based system can provide an efficient means of encoding domain expertise, and can do so with sufficient completeness to cope with the varieties of interactions encountered when used in actual instructional settings.
What the project does focus on is the integration of interface design concepts (providing an interface that not only is unobtrusive in collecting data about the student’s thought processes, but that in fact makes performance by the user easier because of the embedded perceptual and memory aids) with instructional design strategies, such as the use of part-task training and scaffolding, and the presentation of instructional messages that generalize based on the specific error made by a student in order to teach him/her how to deal with the broader class of relevant knowledge. In short, the major research question has been: What is required to successfully integrate expert systems technology into an effective instructional and interface environment that teachers and students want to use, and that results in significant learning?

Below, the domain of study is briefly described in terms of its abstract characteristics. Then TMT is described and the underlying design principles are listed. Finally, the results of empirical studies of its use are presented.

2 Antibody Identification as a Testbed

Antibody identification is a laboratory task where medical technologists must run a series of tests to determine the antibodies in a patient’s blood. It has the classical characteristics of an abduction task, including masking and problems with noisy data.

2.1 The Antibody Identification Procedure

In going from raw data to a diagnostic conclusion, blood bankers must call upon a large body of factual knowledge, apply strategies that have either been taught or derived from past experience, and make hypotheses and predictions to help them through the problem-solving process [1].

2.2 Expert Strategies

Studies by Smith et al. [1] have shown that, like experts studied in other medical domains, expert blood bankers try to sort out which antibodies are causing the reactions by recognizing typical reaction patterns and making early hypotheses upon which to base further analyses. In order to minimize the chance for an incomplete or incorrect diagnosis and to protect against human error and the fallibility of the heuristic methods, the expert blood banker also tries to collect independent, converging evidence to both confirm the presence of hypothesized antibodies and to rule out all others. Thus, there is a high level of skill involved in knowing how to combine various problem-solving strategies such that the overall protocol is likely to succeed.
3 The Design of TMT

Based on studies of the expert strategies and erroneous/inefficient strategies found to be used in this domain, as well as studies of current teaching methods, TMT was designed as a coaching system [2-11]. TMT uses a rule-based system to monitor the student’s actions for evidence of errors and provides feedback if errors are detected. To help ensure use of this strategy, TMT monitors for both errors of commission and errors of omission. The types of knowledge encoded into the system include detecting:

1) Errors of commission (due to slips or mistakes):
   - Errors in ruling out antibodies.
2) Errors of omission (due to slips or mistakes):
   - Failure to rule out an antibody for which there was evidence to do so.
   - Failure to rule out all clinically significant antibodies besides the antibodies included in the answer set.
   - Failure to confirm that the patient did not have an auto-immune disorder (i.e., antibodies directed against the antigens present on their own red blood cells).
   - Failure to confirm that the patient was capable of forming the antibodies in the answer set (i.e., that the patient’s blood was negative for the corresponding antigens, a requirement for forming antibodies in the first place if the possibility of an auto-immune disorder has been ruled out).
3) Errors due to masking:
   - Failure to detect and consider potentially masked antibodies.
4) Errors due to noisy data:
   - Failure to detect situations where the quality of the data was questionable.
5) Data unlikely given answer (low probability of data given hypothesis):
   - Failure to account for all reactions.
   - Inconsistency between the answers given and the types of reactions usually exhibited by those antibodies (e.g., that a warm temperature antibody was accounting for reactions in cold temperatures)
6) Unlikely answers according to prior probabilities (regardless of the available evidence)
   - Antibody combinations that are extremely unlikely due to the way the human immune system works.

3.1 Design Strategy

As reported in a previous paper [12], based on the literature, several design principles were used to guide the design of TMT:

Principle 1. If the goal is to offer students an opportunity to actively apply relevant knowledge and develop important problem-solving skills, provide a problem-solving environment that allows them to integrate this declarative knowledge into the procedural knowledge that they need to develop.
Successful Use of an Expert System to Teach Diagnostic Reasoning

Principle 2. Use expert systems technology to efficiently provide students with immediate, context sensitive feedback or critiques as they perform the problem-solving tasks.

Principle 3. Design a user interface that allows the expert system to unobtrusively collect data on the student's reasoning during the problem-solving tasks, thus allowing the computer to give immediate, context-sensitive feedback.

Principle 4. Design the system to support rather than replace the teacher.

Principle 5. If appropriate, decompose a complex task into subtasks, and teach the subtasks first.

Principle 6. Use a mixture of proactive and reactive teaching methods to teach and reinforce the student's knowledge.

Details on the application of these principles to the design of TMT can be found in [12].

4 Previously Reported Findings: TMT with Instructor Assistance

In a previous paper [12], an empirical evaluation of TMT was reported in which it was used in a classroom lab setting with an instructor present to supplement the tutoring provided by the computer. That study is reviewed here, before presenting new work. In that study, thirty students in the medical technology program at a major U.S. university were tested on TMT. These students were college juniors and had completed the didactic portion of their immunohematology coursework and an associated student lab, but had not yet begun their clinical rotation. The study was conducted at a university where the staff had not been involved in the development of the system.

Since a major goal of this instructional system is to teach antibody identification, a major test is whether the students learn effectively from it. This previous study collected data on two groups of students using the Transfusion Medicine Tutor. The Treatment Group received a version of TMT with all of the intelligent functions turned on, use of a Checklist, the Reference Manual, and access to instructor assistance; while the Control Group used a version of the system with the immediate intelligent feedback turned off and with no other support except for the end-of-case summaries.

A within-subjects analysis (using McNemar's chi square test) of the students' performance in the Treatment Group showed a significant ($p < 0.001$) improvement in performance (an improvement of 87%) from the pre-test case to a matched post-test case (Post-test Case 1). Students in the Control Group showed a 20% reduction in errors that was not significant from the pre-test to post-test Case 1 ($p > 0.05$).

Between-subject analysis using Fisher's Exact Test results showed that there was no significant difference in the misidentification rates on the Pre-test case for the Control and Treatment Groups ($p = 0.50$). However, analysis showed a significant difference in performance on each of the post-test cases between the two groups ($p < 0.005$). On Post-test Case 1, subjects in the Treatment Group had a misidentification rate of 13% (2 out of 15 incorrectly identified the antibodies present) while subjects in the Control Group had a misidentification rate of 73% (11 out of 15 students incorrectly identified the antibodies present) ($p=0.0013$). On Post-test Case 2, students in the Treatment Group had a 7% misidentification rate (1 out of 15 students
incorrectly identified the antibodies present), while students in the Control Group had a 73% misidentification rate (11 out of 15 students incorrectly identified the antibodies present) (p=0.0002). Thus, something about the Treatment Group (the use of intelligent tutoring, the checklist, the reference manual and/or instructor assistance) produced a sizable and statistically significant improvement in performance (Additional details are available in [12]).

5 New Findings: TMT without Instructor Assistance

In this new study, a within subjects design was used to study the effectiveness of TMT as a stand alone tutoring system.

5.1 Subjects

Thirty-six students in Medical Technology Programs at three major U. S. universities were tested on TMT. These students were college juniors and had completed the didactic portion of their immunohematology coursework and an associated student lab. Thus, they were similar to the students in the previous study, except that they were attending different schools. The staff at these universities had not been involved in the development of the system.

5.2 Experimental Design

The students completed five lessons; each of the first four lessons consisted of subtasks involved in solving a complete case. Lessons 1 - 4 each had four to six subtasks to solve. The fifth lesson, Complete Cases, consisted of solving complete patient cases, and included the use of all the subtasks covered in the first four lessons, along with more global strategies for gathering converging evidence to test hypotheses. All of the participants saw the lessons and test cases in the same order, with the exception of the pre-test case and the post-test case, which were randomized with respect to their order of use for each student. The students were allowed to work at their own pace and the entire study took no more than 4 1/2 hours to complete, with most students finishing in approximately 4 hours.

Interface training. In order for the participants to interact as efficiently as possible with the computer, they were introduced to the interface by viewing a 10 minute video while they sat at a computer with TMT running. The video ran through the interface functions available with TMT, instructing them to perform specific actions using TMT. The video did not review any problem-solving strategies. It simply reviewed use of the interface.

Pre-Test. Following the interface training the participants were asked to solve a case, a “pre-test,” that did not contain any intelligent tutoring. This case was used to provide a benchmark, allowing the researcher to determine a student’s overall level of understanding of the task of antibody identification prior to any training using TMT. The purpose of this pre-test case was explained to the students, and they were encouraged to do their best work.
Lessons. Following the pretest case, the students completed five lessons. Each of the first four lessons consisted of subtasks involved in solving a complete case. The fifth lesson allowed the students to work on complete cases. While completing these lessons, the students had access to the following resources:

Checklist. In addition to the immediate, context-sensitive tutoring provided by the computer, the students in the Treatment Group were given a paper checklist that made explicit the high-level goal structure that guided the expert system’s error detection and tutoring.

Reference Manual. The students also received a manual consisting of a concise set of the underlying rules used by TMT when tutoring students. The Reference Manual was meant to be used by the students when they needed a more in-depth explanation of the rules and procedures than was provided by the Checklist.

Post-Test Case. Following Lesson 5 (Complete Cases), a “post-test” case was given to all students in order to assess each student’s overall level of competence on the material just taught. This case was one of two cases matched in characteristics. This post-test case was randomly selected for each subject from one of the two “matched” cases, the other of which was used as the pre-test case. These two cases were matched in that the original testing panel seems to indicate that only one antibody is present, but in actuality, two different antibodies are together accounting for the reactions. For the post-test case, the intelligence was turned off.

Debriefing. A questionnaire was administered to each student to gather additional demographic data (including the student’s age, gender, and previous computer experience), to assess the students’ subjective reactions to TMT and its various functions, and to elicit suggestions for improvement in its design and use. For brevity, however, those results are not reported here.

Data Collection. The computer system logged all of the student’s actions, including final answers. In order to better understand the problem-solving strategies and the errors made by the student (and the possible misconceptions underlying those errors), computer logs for each student were analyzed and coded for certain key behaviors.

5.3 Results

A between-subjects analysis was done on the results for each of the three universities to determine whether there were any significant differences among the three groups which would prevent us from combining the results of the data from the three schools. No significant differences were detected, so the data sets were merged for later analysis. Using McNemar’s Chi Square test for a within-subjects analysis, a significant reduction in errors (from 80% wrong on the pre-test to 13.9% on the post-test) was found from the Pre-Test Case to the matched Post-Test Case (p < 0.001) for these students.

Thus, as measured in this manner, the use of TMT (along with the associated checklist and reference manual) was quite effective even without instructor assistance.

Classes of Errors. In order to better understand the impact of the Treatment condition on learning for the students in Study 2, the computer logs were used to identify error frequencies for four classes of errors (1. Ruling out correct answer due to ruling out incorrectly; 2. Failure to rule out when appropriate; 3. Failure to collect
converging evidence; 4. Failure to check for consistency of data with answer). A within-subjects analysis (using McNemar's chi square test) of the students' end-of-case errors showed a significant ($p < 0.001$) reduction in errors on Errors 2, 3 and 4 from the pre-test case to the matched post-test case. Thus, the tutoring provided to students in this study appeared to be effective in significantly reducing some of the errors that the computer detected immediately after the student marked a final answer for a case.

6 Beta-Testing

Following completion of the two studies described above, copies of the TMT "kit" (software, video, checklist, user's manual and reference manual) were sent to seven medical technology programs around the country. (Several of these programs were hospital-based and teach only 2-3 students at a time.) A cover letter was included that described the purpose of TMT, and that asked the instructor if he/she would consider using the system on a trial basis. The instructor was told that in exchange for a free copy of the kit, we would appreciate it if he/she would return questionnaires for the instructor and the students after its use.

6.1 Objectives

The goal of this study was to determine whether TMT could and would be used in a setting where the designers were not present to assist or motivate the instructor, and to solicit subjective evaluations of its use in such a setting.

6.2 Procedure

A kit containing the software, video, checklist, user's manual and reference manual was sent to each of seven medical technology programs. These materials were the same as those described in the study above, except that the video was redone so that it not only introduced the interface, but that it did so by running through a complete, correct analysis of a full patient case. Thus, the video provided an introduction to the problem-solving strategies taught by TMT.

The cover letter indicated that TMT could be used in a classroom setting or for stand alone instruction. It also indicated that studies had found use of TMT in both settings to result in significant learning, but that the most effective use appeared to be in a classroom lab setting with the instructor present.

If they chose to use TMT, the instructors and students were asked to fill out short questionnaires following its use.
6.3 Results

Results from all seven programs are summarized below.

**Instructors.** On an ordinal scale with levels of Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree, responses by the 7 instructors to a subset of the rating questions asked are summarized below.

This tutorial provided a useful teaching strategy: 2 Agree; 5 Strongly Agree
The software was easy to use: 1 Disagree; 2 Agree; 4 Strongly Agree
I would like to incorporate this tutorial into my classroom: 4 Agree; 3 Strongly Agree.

Responses to open-ended questions indicated that 6 of the 7 instructors chose to use TMT to provide stand alone instruction, without the instructor present. The 7th used it in a classroom lab setting. Comments about its use included (the range of positive and negative comments are listed):

- "It's appropriate for their level of knowledge"
- "A good review of information"
- "Makes the student think logically"
- "I liked that it thoroughly covered antigen/antibody identification"
- "This is a great program for the student who is not strong"
- "It doesn't have the ability to resume where you left off"
- "I am very pleased with TMT and look forward to using it in the future."

**Students.** Responses by the 23 students to a subset of the rating questions asked are summarized below.

The program was easy to use: 3 Disagree; 3 Neutral; 9 Agree; 8 Strongly Agree
I learned a great deal from the program: 2 Disagree; 2 Neutral; 14 Agree; 5 Strongly Agree
I would recommend this program to other students: 1 Disagree; 3 Neutral; 10 Agree; 9 Strongly Agree

A representative set of comments to open-ended questions includes:

- "I liked that it led you step by step through identification procedures"
- "It introduced many small details about the procedure and helped you remember them"
- "I didn't like] having to correlate the video with the tutorial"
- "The case studies were very helpful for practicing for exams for the class and the boards"
- "I was confused at first on how to run the program, which buttons to push"
- "I didn't like] a separate manual from the computer program"
- "It was thorough and was a great help in my understanding of blood bank"
- "It was easy to click on rule outs for the panels"
- "Good review and useful scenarios"
- "Comments were too negative when errors were made. Positive reinforcement would be nice"

As with the instructors, note that most of these comments focus on and are very positive regarding the content and the instructional strategy with the exception of one ("Comments were too negative when errors were made"). The remaining comments deal with interface design.
7 Conclusion

The results of these studies suggest that a principled and multi-disciplinary approach to design can lead to the successful development of an effective computer-based tutoring system. Broadly speaking, our conclusions are the following:

1. Expert systems technology offers an efficient and successful approach for helping to teach problem-solving skills like antibody identification.
2. The successful use of this technology is critically dependent on:
   a. The design of an interface that unobtrusively obtains data on the student’s thought processes, so that context-sensitive feedback can be provided. (Of particular significance in this regard is the provision of memory and perceptual aids that, in the process of helping the user to more easily accomplish his/her task, also serve to encourage the user to provide the computer with a rich set of data about the user’s intermediate thought processes);
   b. Selection of effective teaching strategies in which to embed the use of the technology (a problem-based curriculum; part-task instruction; proactive as well as reactive teaching);
   c. Careful design of the target material (identification of effective problem-solving strategies and structuring of knowledge) to be taught.

More specifically, the studies indicate that, even without the presence of an instructor to supplement the tutoring provided by TMT, use of the system resulted in significant learning. Given this is an often cited benefit from the use of such technology (and one that instructors in this domain specifically noted as highly desirable), such empirical support for the stand-alone use of such an educational tool is important.

References

Teaching the crafts of reading, writing and mathematics. In L. B. Resnick (Ed.), Knowing, learning, and instruction. Hillsdale, N.J.: Lawrence Erlbaum


