Using Online Tutors for Learning – What do Students Think?

Amruth N Kumar
Ramapo College of NJ,
Mahwah, NJ 07430 amruth@ramapo.edu

Abstract - Educators and researchers in various disciplines are developing online tutors to help students learn. Such tutors often promote active learning and problem-based learning. They facilitate individualized learning and provide instant feedback. What do students think about their experience using such online tutors? Do they think that the tutors help them learn? Do their opinions reflect their experience using the tutors? Do they feel favorably about using online tutors? In this paper, we will attempt to answer these questions based on our evaluations of two tutors that we have developed for programming topics. We believe that our results will be of interest to developers of online tutors who are interested in the affective learning of their students.

Index Terms - Affective learning, Cognitive learning, Education research, Online tutors, Programming.

INTRODUCTION

Educators and researchers in various disciplines have been developing online problem-solving systems to help students learn. In Computer Science, these systems may be categorized into three types:
1. Some systems solve problems entered by the student, such as JFLAP [6] for Automata Theory.
2. Some systems administer problems generated by the instructor, and test the learner’s answer, such as WebToTeach, now TuringsCraft [1].
3. Some systems generate problems on specific topics – examples include PILOT [3], a problem-generation tool for graph algorithms; Gateway Labs, which generate problems on mathematical foundations of Computer Science [2]; and SAIL [4], which is a LaTeX-based scripting tool for problem generation.

Our tutors differ from earlier systems in the following ways:
1. Our tutors typically deal with problems based on program code, problems for which the answers may not always be quantitative;
2. Our tutors generate the problems randomly. So, no two problems are alike. Therefore, the tutors deter cheating and prevent plagiarism.
3. Our tutors provide detailed feedback to help users learn from their mistakes.
4. Our tutors adapt to the academic needs of the learner, by generating problems on only those topics that the learner has not yet mastered.

To date, we have developed and evaluated tutors on the following topics:
- Expression evaluation in C++ [9] – the learner is asked to solve expressions online; the tutor provides graphic and text feedback to the learner;
- for loops in C++ [16] and selection statement in C++ [17] – the learner is asked to predict the output of the code and in the case of the for loop tutor, the number of times a loop iterates; the tutor optionally explains the correct execution of the code to the learner;
- Encapsulation in C++ classes [8] – the learner is asked to identify syntax and semantic errors in code; the tutor optionally explains the bugs in the code;
- Pointers in C++ [10,11] – the learner is asked to identify syntax, semantic and run-time errors in code; the tutor optionally explains the bugs in the code;
- Parameter passing in programming languages [15] – the learner is asked to indicate the behavior of a program given a parameter passing mechanism (value, result, value-result, reference or name);
- Scope concepts [7,12,13] – the learner is asked to answer questions based on static scope, dynamic scope, and their implementation.

Our evaluation of the tutors indicates that:
1. Using online problem-solving tutors helps Computer Science students learn programming concepts.
2. Tutors that provide feedback are clearly more effective in helping students learn than those that do not.

But, what do students think about using the tutors? Do they think that they learn from the tutors? During controlled evaluations of the tutors, do the control and test groups feel differently about their differing treatments? Do students feel favorably about using online tutors?

In this paper, we will attempt to answer these questions based on our evaluation of two tutors in several sections of our Computer Science courses over multiple semesters. We will first describe the protocol we used to evaluate the tutors. We will then discuss the results of evaluating our tutors on expression evaluation and pointers. We believe that our results...
The Evaluation Protocol

We used controlled tests to evaluate our tutors – either between subjects, wherein a class is randomly divided into control and test groups for a given topic, or within subjects, wherein the entire class acts first as the control group and then as the test group for two different topics. Further, we used the traditional pre-test - practice - post-test - feedback protocol:

- During the pre-test, students answered a set of questions on the topic. Their performance on the test was interpreted as a measure of their preparedness in the topic. The test was either written or online, but students were neither told whether their answers were correct/wrong, nor were they given any feedback during the test. Typically, the test lasted 6-10 minutes.
- During the practice, the test and control groups practiced with different treatments. The treatments we tried were: using our tutor for the test group versus using a printed workbook for the control group; using our tutor with detailed feedback for the test group versus using our tutor with no/minimal feedback for the control group. Typically, the practice session lasted 10-12 minutes. Students did not have access to any other resource during the practice session.
- During the post-test, students answered a new set of questions that were carefully matched with those on the pretest for their level difficulty as well as the order in which they appeared. The post-test was in the same medium as the pretest – either written or online. If the test was written, often, the questions on the pretest and posttest were themselves generated by the tutors, later typeset and printed. The performance of the students on the post-test was interpreted as a measure of how much they had learned (or did not learn) during the practice session.
- Finally, the students filled out a one-page feedback form. The items on the form covered the usability, learnability, and usefulness of the tutor. Students were asked to respond to each item on the form on a Likert scale of 1 (Strongly agree) to 5 (Strongly disagree).

The change in the student score from pretest to post-test was interpreted as a measure of cognitive learning, if any, that resulted from using the tutor. The responses on the feedback form were used to assess affective learning – how students felt about their learning. In this paper, we will compare cognitive and affective learning from the evaluation of two tutors that we have evaluated over several semesters: expressions and pointers.

Expression Evaluation Tutor

The tutor on expression evaluation is designed for use in Computer Science I/II courses. It covers arithmetic, relational, logical, bit-wise and assignment operators in C++/Java. It can provide feedback to the learner at several levels: 1) no feedback except to specify the correct final answer of evaluating an expression; 2) error flagging wherein an incorrect step in the solution of an expression is color-flagged as soon as the user enters it; 3) graphic visualization using underbraces to indicate the order in which an expression is evaluated; 4) text explanation wherein the order of evaluation of an expression is described in prose, in terms of precedence, associativity, etc.

In Spring 2001, we compared the tutor against textbooks – we tested the hypothesis that practicing with the tutor would be at least as effective as practicing with a printed workbook that provided answers to problems at the end. The improvement in the score per attempted problem from pretest to post-test was 20.1% for workbook users (N=31) and 29.6% for tutor users (N=33), both the improvements being statistically significant. The effect size, calculated as (post-test average – pretest average) / standard deviation on the pretest – was 1.27 for the tutor users versus 0.71 for the workbook users. Table I lists the averages of the student responses on the feedback form. The feedback provided by the tutor and the feedback provided in the workbook were identical – graphic visualization of the steps needed to solve the expression. The 10% or greater difference between the responses of the workbook users and the tutor users on the first three questions may be read as a preference among students for online delivery, which is novel compared to the traditional delivery of problems in print. The difference in the responses for questions 4 and 5 further illustrates that students may more readily use online tutors for problem-solving practice than they might, printed textbooks.

Table I

<table>
<thead>
<tr>
<th>Feedback Item</th>
<th>Workbook Users</th>
<th>Tutor Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It provided useful feedback</td>
<td>2.35</td>
<td>2.16</td>
</tr>
<tr>
<td>2. The provided feedback was clear</td>
<td>2.14</td>
<td>1.88</td>
</tr>
<tr>
<td>3. It helped me learn the material</td>
<td>2.23</td>
<td>1.91</td>
</tr>
<tr>
<td>4. It should be made available to all students</td>
<td>2.39</td>
<td>1.84</td>
</tr>
<tr>
<td>5. I would like to see such instrument for other topics</td>
<td>2.45</td>
<td>2.03</td>
</tr>
</tbody>
</table>

In Fall 2001, we compared two versions of the tutor – one that provided both graphic visualization and text explanation versus one that provided only graphic visualization, i.e., both the groups got to visualize the correct order of evaluation, but, only the test group got additional text explanation of the evaluation. Both the control and test groups used the tutor for practice. We tested the hypothesis that students who got the text explanation would learn better than those who did not. The improvement in the score per attempted problem from pretest to post-test was 51.2% for the test group (N=33) and 39.8% for the control group (N=33), both improvements being statistically significant. The effect size was 1.30 for those who got both graphic visualization and text explanation versus 1.00 for those who got only graphic visualization. Table II lists the average of the student responses on the feedback form. Since both control and test groups used the same tutor, albeit with different levels of feedback, it is reassuring to see that the two
The tutor on C++ pointers is designed for use in Computer Science I/Data Structures course. It covers the concepts of dangling pointers, lost objects, syntax errors and semantic errors. It can provide feedback to the learner at several levels: 1) no feedback except whether the learner’s answer is correct or not; 2) diagnostic feedback which explains only the section of the code that has a bug in it; and 3) explanatory feedback which explains the execution of the entire code from start to finish, including any bugs that may be in the code.

In Spring 2001, we compared the tutor against textbooks – we tested the hypothesis that practicing with the tutor would be at least as effective as practicing with a printed workbook that included answers to problems at the end. The improvement in the score per attempted problem from pretest to post-test was 30.7% for workbook users (N=17) and 33.9% for tutor users (N=16), both the improvements being statistically significant. The effect size was 1.10 for the tutor users versus 0.92 for the workbook users. Table IV lists the average of the student responses on the feedback form. The tutor provided explanatory feedback, whereas the printed workbook merely listed the correct answer. The responses to questions 3-5 may be attributed to the explanatory feedback provided by the tutor – whose level of detail cannot be reasonably duplicated in printed workbooks or textbooks. Students may more readily use online tutors that provide explanatory feedback than they might, printed textbooks. Clearly, the provision of feedback is one of the values added to online delivery of practice problems, and should ideally be a staple of every online tutor.

### Table IV

<table>
<thead>
<tr>
<th>Feedback Item</th>
<th>Workbook Users</th>
<th>Tutor Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It provided useful feedback</td>
<td>3.06</td>
<td>2.20</td>
</tr>
<tr>
<td>2. It helped me learn the material</td>
<td>2.88</td>
<td>2.31</td>
</tr>
<tr>
<td>3. It should be made available to all students</td>
<td>2.65</td>
<td>1.56</td>
</tr>
<tr>
<td>4. If this tutor is made available, I will use it</td>
<td>2.65</td>
<td>1.93</td>
</tr>
<tr>
<td>5. I would like to see such instrument for other topics</td>
<td>2.59</td>
<td>1.44</td>
</tr>
</tbody>
</table>

In Fall 2001 and Spring 2002, we compared two different versions of the tutor – one that provided graphic visualization, versus one that did not provide any explanation except for the correct final answer. Our hypothesis was that those who got the graphic visualization would learn better than those who got no explanation. The improvement in the score per attempted problem from pretest to post-test was 53.1% for the test group that got graphic visualization (N=24) and 33.4% for the control group that got no explanation (N=24), both improvements being statistically significant. The effect size was 1.35 for those who got graphic visualization versus 0.84 for those who got no explanation. Table III lists the average of the student responses on the feedback form. From the responses to questions 1 and 2, it is clear that students think that online tutors that provide some form of explanation are more effective and useful than those that do not provide any explanation. In question 3, they tend to disagree that providing only the final answer is sufficient for learning from an online tutor. Even though graphic visualization by itself is insufficient (as demonstrated by the results of Fall 2001), students still prefer to have it to not having any feedback, as seen in question 4. Curiously, students do not find a tutor that provides only graphic visualization any more compelling to use than a tutor that provides no feedback at all, as seen in question 5.

### Table III

<table>
<thead>
<tr>
<th>Feedback Item</th>
<th>No explanation</th>
<th>Visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The generated problems were instructive</td>
<td>2.65</td>
<td>2.15</td>
</tr>
<tr>
<td>2. The provided feedback was useful</td>
<td>3.19</td>
<td>2.27</td>
</tr>
<tr>
<td>3. The provided feedback was sufficient</td>
<td>3.38</td>
<td>2.42</td>
</tr>
<tr>
<td>4. It helped me learn new material</td>
<td>3.00</td>
<td>2.44</td>
</tr>
<tr>
<td>5. If this tutor is made available, I will use it</td>
<td>2.46</td>
<td>2.42</td>
</tr>
</tbody>
</table>
the explanatory feedback was too verbose. This prompted students to ignore the feedback rather than read it carefully. Table IV lists the average of the student responses on the feedback form. Students were more positive about the sufficiency of explanatory feedback than no feedback, as seen in the response to question 2. But, from the response to question 3, it is clear that students did not feel that they were learning from the verbose explanatory feedback. They were less inclined to use a tutor that provided verbose feedback than one that did not.

<table>
<thead>
<tr>
<th>Feedback Item</th>
<th>No feedback</th>
<th>Explanatory feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It provided useful feedback</td>
<td>2.56</td>
<td>2.23</td>
</tr>
<tr>
<td>2. The provided feedback was sufficient</td>
<td>2.93</td>
<td>2.23</td>
</tr>
<tr>
<td>3. It helped me learn new material</td>
<td>1.69</td>
<td>2.05</td>
</tr>
<tr>
<td>4. If this tutor is made available, I will use it</td>
<td>1.53</td>
<td>1.90</td>
</tr>
</tbody>
</table>

### DISCUSSION

In the preceding section, we listed only those items from the feedback forms for which the Likert average differed by at least 10% between the control and test groups. From the analysis of these figures, it appears that there is a good correspondence between the cognitive and affective learning of students using online tutors – whether students actually learn versus whether they think they have learned. Some conclusions that we can draw from the above analysis about what students like to see in online tutors include:

1. When presented with a choice of online tutors and printed workbooks/traditional textbooks for learning, **students prefer to use online tutors**. However, it is common knowledge among developers of such tutors that unless students are required to solve problems as part of the requirements of a course, they are not very likely to use either online tutors or printed workbooks.

2. **Students prefer tutors that provide feedback** to those that do not. They think that online tutors that provide some form of explanation are more effective and useful than those that do not provide any explanation. Whereas providing explanation/feedback in the print medium is laborious and impractical, the same is not true for online tutors that are not constrained by space limitations – ideally online tutors should play up this advantage by providing feedback for every problem.

3. **Students prefer text explanation to accompany graphic visualization**. They may more readily use a tutor that provides both graphic visualization and text explanation than one that provides only graphic visualization. However, when presented with a choice between graphic visualization and no explanation, they prefer tutors that at least provide graphic visualization.

4. The amount of feedback provided by the tutor also affects the acceptability and learnability of a tutor - students not only tend to learn less with a verbose tutor, but also feel that such tutors are less helpful for learning. The designers of online tutors must strike the right balance between no feedback and verbose feedback in order to make a tutor both acceptable to student learners and effective at promoting learning. Adapting the amount of feedback to the level of knowledge of the learner would be the ideal solution.

We plan to continue to evaluate our tutors in additional classes and at other institutions, and analyze the collected data to identify correspondences between cognitive and affective learning engendered by the use of online tutors.

### ACKNOWLEDGMENT

Partial support for this work was provided by the National Science Foundation’s Course, Curriculum and Laboratory Improvement Program under grant DUE-0088864.

### REFERENCES


