Using the Think Aloud Method to Observe Students’ Help-seeking Behavior in Math Tutoring Software

Minghui Tai
Math, Science and Learning Technologies Program
Department of Teacher Education and Curriculum Studies
University of Massachusetts Amherst
Amherst, Massachusetts USA
mtai@educ.umass.edu

Beverly P. Woolf, Ivon Arroyo
Computer Science Department
University of Massachusetts Amherst
Amherst, Massachusetts USA
ivon@cs.umass.edu
bev@cs.umass.edu

Abstract—This qualitative study presented high school students’ help-seeking behavior and how they interacted with hints while they solved math problems on an intelligent tutoring system for math. Eight high school students were observed by using the think aloud method while they solving math problems. We describe students’ help-seeking behavior at “the moment” while they sought for hints for help. We discuss students’ help seeking behaviors and discuss possible reasons to explain use and misuse of students’ help-seeking strategies.

Keywords: think aloud method; help-seeking behavior; intelligent tutoring system; math education

I. INTRODUCTION

With continued rapid development of technology, intelligent tutoring systems (ITS) provide interactive learning environments (ILEs) for education. This new educational technology supports teachers’ teaching and enhances students’ learning, allowing learners to gain knowledge through new media [4]. At the same time, learning becomes boundless, since students can access learning through the computer at home after class. As well, they are able to get more chances and time to learn in their own pace.

One of the characteristics of ILEs is their on-demand help feature. Hints, hyperlinks and glossaries can be provided in computer learning environments, and be sought for whenever necessary to support learning. While Hints are created according to specific content of the target-learning subject, hyperlinks can provide more possible helpful resources on internet, and glossaries list terms and definitions in a specific subject.

There is one apparent difference between human tutor and an intelligent computer tutor. Human tutors can facilitate and decide when, how, and what level of assistance to provide to students depending on cues of the human-to-human interaction. However, in human-to-computer interaction learning environments, learners may take an active role to seek help. They control when, how, and to what extent they need help during learning. This ability and action of help seeking relies on learners’ “metacognition” [9]. Hence, learners’ help-seeking behavior [9] plays an important role in ILEs, and affects learning.

II. MODELS OF HELP-SEEKING BEHAVIOR

A. Models of Help-seeking Behavior in other studies

In Aleven, Stahl, Schworm, Fisher, and Wallace’s review [3], they pointed out that help-seeking has not really been studied in the context of computer learning environments. The model of help seeking was presented originally by Nelson-LeGall [7] in the classroom, the social context. Aleven et al. [3] applied Nelson-LeGall’s model and elaborated it as a framework of help-seeking in computer learning environments. The model has five steps, which are: 1) become aware of need for help; 2) decide to seek help; 3) identify potential helper(s); 4) use strategies to elicit help; and 5) evaluate help-seeking episode.

Aleven, McLaren, Roll and Koedinger [1] designed a model of help-seeking behavior with The Geometry Cognitive Tutor. In their model, the ideal help-seeking behavior of students should spend time to think about steps of solving a problem. If they feel it is familiar, they should try to solve the problem. If they don’t seem familiar with it, they should ask a hint from the tutor. Students should spend time to read a hint and decide if they need more hints to help them solving the problem. They should check Glossary for help if they don’t have sense of what to do. If it is helpful from the Glossary, they should try to solve the problem. If it does not help, they should ask hints for help. After students tried steps to answer, the tutor will tell students the answer is correct or not. If the answer is not accurate, they should try to solve the problem again or ask hints for help.

We agree this model developed by authors in [1] was more fined-grained than models in the past. However, Mercier and Frederiksen [5] argue that students’ actions such as understanding problems, figuring out what to do, reading hints from the tutor, and solving a problem are related to students’ cognitive function. Therefore, they presented their model is rooted in cognitive theory instead of metacognitive theory.

Mercier and Frederiksen [5] presented a cognitive model of help seeking that links the help seeking process to a problem solving episode. Four steps are followed in this refined cognitive help-seeking model:
• Help goal setting. In the repair procedure of problem solving, help seeking starts with creating a new goal. Students decide to seek help in order to proceed the task. Therefore, it can be seen as a sub-goal in problem solving. Three sub-components are included in this first step of the model: a) recognize an impasse, b) diagnose the impasse, and c) establish a specific need for help.

• Find appropriate help. Students need to choose help according to what they need.

• Comprehend the help. Students need to understand the new knowledge from the help and update their mental model in order to progress the performance of the task.

• Evaluate help. Monitoring comprehension and monitoring proper material are important skills that students must develop in order to achieve learning.

The models of help-seeking presented above are the major literature of this field. However, the latter one was not designed for any specific ITS, ILE, or subject domain. We believe that different software with different designs of help functions may lead students have slightly different steps when they seek help in the interactive computer learning environment. Therefore, the model of help-seeking could be altered depending on a specific tutor. This software-dependent model of help-seeking can also be used by the designer or engineer to find possible flaws in the tutor’s help function in the tutor. We applied the model by Aleven et al. [1] in this study to observe students’ help-seeking behavior in Wayang Outpost, an intelligent math tutoring software.

B. A Model of Ideal Help-seeking Behavior for Wayang

This tutor-dependent model of help-seeking behavior was based on Aleven [2]; however, it was revised to fit the Wayang math tutoring software (see Figure 1).

Students with ideal help-seeking behavior, according to the model would behave as follow: 1) they would spend time to read and understand the problem; 2) they would recognize if they know how to solve the problem; 3) they should ask for help from Hint button when they don’t know; 4) they should spend time to understand hints; 4) if a hint is helpful, students should try to solve the problem; 5) if a hint is not helpful, and students are still not clear about how to solve the problem, they should ask more hints for solving the problem; 6) during the time this ideal student solves the problem, he/she would use the formula sheet and calculator if it is needed.

III. THE USER STUDY

We ran a study using think-aloud methods to observe students’ help-seeking behavior during the use of Wayang Outpost, in order to understand proper and improper uses of the software and the help features. We investigated:

a) What are the participants’ help-seeking behaviors when solving math problems in Wayang Outpost?

b) What happens during “the moment” when participants interact with help features?

A. The testbed tutoring system: WAYANG OUTPOST

We examined high school students’ help-seeking behavior within the interactive computer-learning environment, Wayang Outpost. Wayang Outpost is an online ITS for middle and high school students to learn mathematics and help them prepare standardized math test. In this system, students practice math problems in the “Learning Hut”. Each problem provides detailed multimedia hints about how to solve the problem. Students can just click the “Help” button to receive hints. One click of Help button presents one hint. Students can continue to click on the Help button to get more hints until they know how to solve the problem, or see all hints to get full animated processes that show a solution to the problem.

B. Participants

The participants in this study were eight students in a suburban high school in western Massachusetts. They were randomly selected from around 1,000 students in ten math classes, which these ten classes participated in Wayang research studies. Table 1 presents the information of eight participants. The average of their age was sixteen years old, three girls and five boys. One male and one female student had IEPs (are struggling students with Individualized Education Plans).

C. Protocol data and other data collection

The participants were asked to think aloud as they tried to solve math problems. Before starting, we showed participants a video demonstrating how to do a think-aloud while solving a
problem. We explained that we were interested in getting to know the thoughts that come into participants’ minds as they carry out the task of solving math problems. After the participants understood how to do it, the participants got time to practice doing think aloud before protocol recording. We only prompted participants when it was necessary by just saying: “Keep talking” [8] or “Tell me what you are thinking” when participants were silent or stopped talking at all.

The participants were audiotaped while they were thought aloud. They were not given the same amount of time to finish this think aloud task because each participant took different amount of time to solve problems, and they spent time very differently. Also, they were not asked to solve the same amount of problems in Learning Hut. The participants might see easy problems and solve problems correctly without using Help button during the Think aloud task. In order to collect productive protocol data, each participant recorded their think-aloud task under different time and different numbers of problem conditions based on the variety of their behavior. The think aloud task time period was between fifteen minutes to less than thirty-seven minutes (see TABLE 1.)

In order to have more details of the participants’ actions other than audio protocol data, the data from the computer logs in the Wayang Outpost system was gathered to help interpret student behavior along with the protocol data. The computer logs showed participants’ detailed actions on computer when they solved problems, such as Hints seen in each problem, incorrect attempts per problem, seconds to solve and seconds to first attempt for each problem.

D. Interpreting the protocols and coding

The analysis of content in raw protocols was in terms of the model of help-seeking behavior. Every step or process in the model of help-seeking behavior was labeled as a coding scheme. Steps or processes that were not covered by the model of help-seeking behavior were categorized in order to analyze the unexpected help-seeking behavior.

The coded protocol of each math problem presented a unique theme of a student’s help-seeking behavior. Themes from all problems for each participant could be generated as general patterns of his or her help-seeking behavior. Individual differences of help-seeking behavior were summarized.

IV. RESULTS

Coded Think-aloud protocols of each math problem from all participants have been categorizing to nine themes. TABLE II shows the analysis of participants’ actual help-seeking behavior and the number of problems and percentage that each participant saw.

---

**TABLE I. THE INFORMATION OF EIGHT PARTICIPANTS**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Gender</th>
<th>IF</th>
<th>Time seen during Think Aloud (min)</th>
<th>Problems seen during Think Aloud</th>
<th>Problems solved during Think Aloud</th>
<th>Problems seen during protocol</th>
<th>Problems solved during protocol</th>
<th>Learning % gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>F</td>
<td>7</td>
<td>19:47</td>
<td>11</td>
<td>15</td>
<td>73.9%</td>
<td>15</td>
<td>73.9% 00.0%</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>M</td>
<td>8</td>
<td>21:15</td>
<td>16</td>
<td>15</td>
<td>66.7%</td>
<td>15</td>
<td>66.7% 00.0%</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>M</td>
<td>8</td>
<td>19:02</td>
<td>16</td>
<td>15</td>
<td>73.9%</td>
<td>15</td>
<td>66.7% 00.0%</td>
</tr>
<tr>
<td>4</td>
<td>N/A</td>
<td>M</td>
<td>5</td>
<td>23:20</td>
<td>16</td>
<td>15</td>
<td>43.9%</td>
<td>15</td>
<td>53.9% 25.9%</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>F</td>
<td>7</td>
<td>36:39</td>
<td>16</td>
<td>15</td>
<td>86.7%</td>
<td>15</td>
<td>86.7% 00.0%</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>M</td>
<td>8</td>
<td>17:55</td>
<td>15</td>
<td>15</td>
<td>80.0%</td>
<td>15</td>
<td>80.0% 00.0%</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>M</td>
<td>8</td>
<td>33:13</td>
<td>15</td>
<td>15</td>
<td>73.9%</td>
<td>15</td>
<td>73.9% 00.0%</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>F</td>
<td>7</td>
<td>31:54</td>
<td>15</td>
<td>15</td>
<td>60.0%</td>
<td>15</td>
<td>60.0% 20.0%</td>
</tr>
</tbody>
</table>

---

**TABLE II. THEMES OF PARTICIPANTS ACTUAL HELP-SEEKING BEHAVIOR AND NUMBERS OF PROBLEMS**

<table>
<thead>
<tr>
<th>Themes</th>
<th>Participants</th>
<th>Solved problem correctly without using Hints (%)</th>
<th>Using all Hints and solve problem correctly (%)</th>
<th>Using all Hints but failed to solve problem correctly (%)</th>
<th>Can’t solve the problem but do not use hints (%)</th>
<th>Can’t solve the problem and did not use either (%)</th>
<th>Guessing (%)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Solved problem correctly without using Hints (53%)</td>
<td>74</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>74</td>
<td>74%</td>
</tr>
<tr>
<td>b) Solved problem correctly by the help from hints (13%)</td>
<td>94</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94</td>
<td>94%</td>
</tr>
<tr>
<td>c) Using some hints, but still can’t solve problem (2%)</td>
<td>15</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>15%</td>
</tr>
<tr>
<td>d) Using all hints, but still fails to solve problem (4%)</td>
<td>15</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>15%</td>
</tr>
<tr>
<td>e) Using all hints, but no focus on problem solving (6%)</td>
<td>15</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>15%</td>
</tr>
<tr>
<td>f) Can’t solve the problem, but did not use hints (3%)</td>
<td>15</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>15%</td>
</tr>
<tr>
<td>g) Guessing (18%)</td>
<td>25</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>25%</td>
</tr>
</tbody>
</table>

---

A. What are participants’ help-seeking behaviors when solving math problems while using Wayang Outpost?

1. **Solved problem correctly without using Hints (53%)**
   - 74 out of 140 problems fell into this first category based on the coded think aloud protocol. Participants solved math problems without using Hint button. This means they tried their steps when solving a problem and were able to solve it and got the answer correct.

2. **Solved problem correctly by the help from hints (13%)**
   - This theme included participants that were able to solve the problem correctly after asking help from hints. Participants in this theme demonstrated the ideal help-seeking behavior that we expected. Excluding 74 problems that participants could solve correct by themselves, 18 out of the 66 problems were categorized in this theme.

3. **Using some hints, but still can’t solve problem (2%)**
   - The third theme, not being able to solve the problem after receiving hints included 3 problems. Participants did not see all hints, but just saw couple of hints. They stopped asking for further hints, and finally failed to solve problems or got the correct answer by luck.

4. **Using all hints, but still fails to solve problem (4%)**
   - Six problems were categorized in this theme, and 3 participants (participants 1, 4, and 5) showed they behaved this way. This is the worst case of participants’ help seeking behavior. They had seen all hints in a problem, and they also tried steps to solve problems, but still failed to solve it, even though participants saw the correct answer from the last hint.

5. **Using all hints, but no focus on problem solving (6%)**
   - This theme, in which participants use hints but no problem solving steps, is different from the previous theme. Participants tried to understand what each hint told them, but attempt to solve problems or they might just click through all hints quickly. Participants got the correct answer because the last hint contained the answer, but no effort was exerted. There were 8 problems categorized in this theme based on the behavior of four participants (2, 4, 6 and 7).

6. **Can’t solve the problem, but did not use hints (3%)**
   - Five problems were categorized in this theme including two participants (participant 5 and participant 8) rejected asking for hints and were not able to solve the problem correctly.

7. **Guessing (18%)**
   - Twenty-five problems were categorized as guessing. Four participants (1, 3, 4 and 8) guessed when they tried to solve
problems until they got the correct one. The protocols suggested reasons for guessing was that, after clicking on help a couple of times, those hints did not really help them or gave enough information for solving problems. They did not choose to see more hints and began to guess. In another case, participants did not even want to ask help. They just tried wrong steps to make up numbers, which would sound like they were generating an answer, or they randomly guessed.

h) Memory (<1%).

Only one problem categorized in this theme, that is, participant 4 saw the same problem twice. In his second time seeing the problem, he used the memory strategy.

i) Other (<1%)

One problem categorized in this theme was from participant 1. This was the case that the participant was able to solve the problem correct at the beginning. The correct answer that the computer system gave was “48-9x.” However, she could not find the answer from options that matched her own answer, which was “20.” She then tried to do it again and asked help from hints as well. She was confused by hints, and failed to solve it. She finally chose the correct answer from the last hint showed her.

B. What happened at “the moment” when participants interacted with hints?

a) See some hints and solve the problem right by chance

From participants 5 and 8’s protocols, we saw that both of them had asked for hints and yet they both used the wrong steps to solve problems without asking more hints and then got the correct answer by luck.

b) Using hints for help, but still fail to solve problem

The log on Wayang system showed that participant 7 saw 3 hints on problem and made one incorrect attempt. The total hints for this problem was four. After he had seen three hints, he said, “ok… I have no idea… umm… not sure… I am still not sure…” He did not check last hint and commented himself that “math is not my best subject.”

c) The last hint showed the correct answer, and hints were not helpful

Participant 4’s protocols showed that he did not really understand hints he had seen: “Oh… I am not sure if that’s gonna help me or not.” His protocols indicated that he was not satisfied because hints did not really help him or give him enough information to solve problems but showed him the correct answer too soon: “Choose A????! That was a hell of a hint.” In another case he was forced to choose the correct answer by the last hint button: “Ok~Ok~ I will choose answer E, calm down!!! It’s still going. It’s really persistent.” Protocols from participant 4 indicated that he did not understand how to solve a problem; he chose the correct answer because the last hint told him to do so.

d) Don’t like problems, so no solving steps but use all hints

Participant 2 always commented that he did not like problems at the very beginning. He said, “Oh… Geez… I don’t like these problems.” This negative mood made him not even try to carry out any steps to solve problems. Therefore, he asked for hints and tried to understand what hints had showed him. He then waited for the correct answer from the last hint, but did not try any steps to solve the problems, even though those hints made sense. He said, “Ok… Great… it tells me to choose answer D, I’m going with that one (sarcastic)!!

e) Click through hints

Participant 4 was a case, which showed the behavior of “click through hints.” He did not really “think aloud.” And did not try steps to solve problems either. He just clicked through hints and got the correct answer from the last hint.

f) Check Hints but went guessing

Many participants (3, 4, and 8) saw hints and yet began to guess. The following protocol from participant 8 presents the example that she only saw 1 hint and began guessing.

1: Well… I am gonna try hint...
2: because last time I just guessed,
3: and I [already] knew I was trying to find the shaded area…
4: I am gonna try 6 plus 8, which is 14… which I didn’t really want to get…
5: I think it’s B or C…
6: I am just gonna go with a guess of C,
7: and I am… B…
8: and B came out to be right…
9: New problem…

Line 1 showed she was going to ask for a hint. Line 3 was the comment after she saw the first hint. The first hint seemed did not give her enough information about how to solve this problem; therefore, she stopped asking help from hints and tried incorrect steps at line 4. She then started guessing from line 5 to line 7. Line 8 showed she finally got the correct answer by guessing.

V. DISCUSSION

In this study, we observed participants’ help-seeking behavior by using think-aloud methods while students solved math problems in Wayang Outpost software. We identified distinct problem solving behavior namely not using hints associated with this ITS (53%), ideal help-seeking behavior that conforms to the model (13%). Help-seeking that did not conform the model (34%), which the primary behavior was guessing (18%), which included guessing while checking hints and guessing without checking hints.

In Aleven’s et al. study [2], they found 7% of help-seeking behavior included a bug called Guess Quickly when Help Use was Appropriate. This means that less skilled students tried to attempt too quickly (within 7 seconds) but still got it wrong after the appropriate help from hints. In our study, though we did not analyze students’ problem solving speeds; we observed those moments cause students went guessing. The first type of guessing behavior was students did not satisfy with hints and did not have clear idea of how to solve the problem. Our results showed that students usually checked no more than 3 hints and then began to guess. This was because they felt they did not receive enough information that could enhance their problem solving skills after they checked the first couple of hints. They finally gave up seeing more hints, and went to guess the answer.
The second type of guessing behavior included students who avoided asking help from hints, and went on to guess the answer. Participant 8 especially exemplified this type of behavior. She preferred guessing until she got the correct answer rather than asking for hints. We are not clear why she did not want to use hints. However, the possible reason was that she was a good guesser. She seemed to enjoy guessing answers correctly instead of learning math.

A student with ideal help-seeking behavior should see hints if they don’t know how to solve a problem. Hints are supposed to help students solve problems along with accurate problem solving steps. However, when looking at the think-aloud protocols, we found that students actually used incorrect solving steps and got the correct answer by coincidence. The protocol suggested that students might not be clear what the problem was actually asking; they made up some numbers and thought they had solved the problem. Besides, the protocol also showed that students might lack or not be clear about certain math knowledge, or they misunderstand certain math rules. Thus, they misused it, and solved the problem correctly with luck. Similar situation happened when computer logs showed students solved problems correctly without using hints. However, the true “moments” behind the logs record was students solved the answer correctly by coincidence, according to the protocol analysis.

Another case we observed was when students saw some hints but failed to solve the problem correctly. They, however, did not see all the hints after their incorrect attempt possibly because they were frustrated and were not confident.

When students saw all hints until the answer hint was shown in a problem, we observed many different situations in protocols. This “bottom out” situation presented in Aleven’s et al. article, (see [1] and [2]) was “Click through hints”. They concluded that “click through hints” was the most frequent errors student made in their studies. Students spent very little time on one hint and then moved to the next hint. In our study, we only had two problems categorized as “click through hints” situation. The possible reason was because of the think-aloud method; since we required students to think out loud while they solved problems, “click through hints” did not happen as frequently in our study. We are not clear if students in our study would have the same improper behavior if think-aloud were not also used. However, we presented another perspective to understand what happened after students had “bottomed out”. The gratifying situation was students bottomed-out with their own problem-solving steps as well, and realized how to solve the problem finally after they saw all hints. Another case was that students did not just “click through” but also tried their steps to solve the problem. However, unfortunately, the protocol showed that students may still not understand how to solve the problem even after they have seen all hints. They got the correct answer because the last hint showed them the answer.

Another situation we observed from the protocol was that students did “bottom out” because they did not like certain problems. They chose to let hints show them how to solve the problem and they tried to understand it instead of trying steps. They knew the last hint would tell them the correct answer. We suggest that the last hint should not be accessible through the “HELP” button and should be reached by another button in the system. In this way, the function of “HELP” button only has the function to show hints and provide help. It can avoid students “clicking through” hints and only waiting for the answer. It can also “motivate” students to try to use their own steps to solve the problem after they had seen all hints.

VI. CONCLUSION

This study showed an analysis of how students interact with help in a tutoring software environment, in a qualitative way. We concluded that Guessing was the main behavior that was not expected in the ideal help-seeking behavior with the Wayang Outpost system. We presented possible reasons that made students guess. We also presented different situations after students “bottom-out.” We provided a possible solution to motivate students and avoid students “clicking through hints.” We also detected a different result between logs in the system and “the moments” presented while students were solving the problem. A quantitative analysis should be conducted in the future to link students’ pre test and post-test performance with their hint use behavior in the Learning Hut. Moreover, questionnaires about students’ math value, self-concept, and frustration level should be analyzed as well to study if this is related to their improper help-seeking behavior.

The future questions in this field we would like to ask is how to motivate proper help-seeking behavior in order to enhance learning with the system.

REFERENCES


