December 2011

Field Observation of Student Behavior in ASSISTments

Breanna Eileen McElroy
Worcester Polytechnic Institute

Mohammad Abdullah Alshuqaiq
Worcester Polytechnic Institute

Field Observation of Student Behavior in ASSISTments

An Interactive Qualifying Project Report:
submitted to the faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science
By

Mohammad A. Alshuqaiq
Breanna E. McElroy

Date: December 14, 2011

Approved:

____

Prof. Ryan S. Baker, Major Advisor

Abstract

We examine the observed behaviors that students made while using a math tutor program.We then study in more detail three of these behaviors, which are Re-entering/Keeping Incorrect Answer, Adding Numbers with Fingers, and Waiting on Teacher. Then, we perform data analysis on our results to see if any behaviors hold significance. We find that demand for the teacher on one side of the room results in students on the other side of the room having to wait for assistance longer. We also find that regardless of using a tutor program, students still work with each other.

Acknowledgments

We would like to thank our advisor, Professor Ryan Baker, for his understanding, patience, and guidance throughout the course of this project.

We would like to thank Mrs. Sue Donas from the Burncoat Middle School for her assistance in setting up studies at the school.

We would like to thank all the students who contributed to our research and allowed us to study their behavior while learning.

Introduction

In recent years, technology has become more and more involved in the teaching environment. One way it is being implemented into classrooms is through tutoring programs on the students' computers, which help aid the students without any help from the teacher (Koedinger, & Corbett, 2006). Having a special "tutor" in each computer can greatly help students learn how to attack problems when they are having difficulty. This type of system helps students learn more than with traditional teaching methods (Koedinger et al, 2006); however, how the student uses the system can affect how much they will learn, as indicated in research by Professor Baker and his colleagues (Aleven, McLaren, Roll, & Koedinger, 2006). In a traditional classroom setting, the entire classroom is the students' learning environment, which is where they receive their instructions, teaching, problem sets, and assistance. However, when using computers in the classroom as a learning tool, the computer system itself is intended to be that learning environment for each individual student; with this, a common problem arises. Students tend to disconnect with this new learning environment or, as it is sometimes called, they exhibit off-task behavior, which occurs when "a student completely disengages from the learning environment and task to engage in an unrelated behavior" (Karweit & Slavin, 1981). In our project, we found a few behaviors that seemed potentially relevant to understanding student learning in intelligent tutors, causing us to research them in more detail. These behaviors consist of: students waiting for assistance from the teacher even though there was both a tutor and a hint button in the program being used, students using their fingers to count numbers, and students reentering the same wrong answer.Our goal is to observe how students interact with the tutor system and possibly find reasons for these interactions.

Burncoat Middle School is our setting to collect our data. The school uses a system run by WPI called ASSISTments (Razzaq et al, 2007). ASSISTments is a math tutoring program that allows teachers to make online quizzes or questions for their students. “The ASSISTment System contains tutoring for over 3,000 problems and is growing everyday as teachers and researchers build content regularly” (Mendicino, Razzaq, &Heffernan, 2009).It is a tutoring program in that it helps students work through types of problems that they are struggling with, usually through the form of scaffolding questions. With scaffolding questions, if a student gets a question wrong, then the program will ask the student two or three related questions which present concepts that the original (incorrectly answered) question was built upon. This helps ensure that the student has a full understanding of the provided material (Heffernan, 2009). The goal of ASSISTment is no student left behind; “The dilemma is that every minute spent testing is a minute taken away from instruction. ASSISTments solves this problem by tutoring students on items they get wrong, thus providing integrated assisting of students while they are being assessed. Teachers can use this detailed assessment data to adjust their classroom instruction and pacing.”(Mendicino et al, 2009.) Studies have shown that student knowledge increases more when using ASSITments for homework rather than doing traditional homework (feedback after the due date) (Mendicino et al, 2009.) In the following sections, we describe the process used for conducting our research and discuss our results.

Methods

After establishing the target school to conduct the study at, we met with the instructor that we would be working under to discuss our project and to schedule days in which we would conduct our research. In total, we had two observational research days, each with a very different purpose.

The first observational day's work was directed toward gathering qualitative observations (Schofield, 1995), primarily students' interactions with ASSISTments and the environment around them. We conducted this observation in a very simple manner. Each observer started across the room from one another and individually began observingstudents. They watched each student just long enough to decide what behavior the student was engaging in, and then recorded that behavior in a notebook. Once they finished recording the data, they each moved counter-clockwise to observe the next student. In this way, each student was observed approximately the same number of times as every other student. This approach allows observers to pick up on subtle behaviors, as opposed to watching the entire room at once and only noticing behaviors that dramatically stick out, which are a rare happening (Baker, D'Mello, Rodrigo, 2010). We continued this process throughout the entire day, observing a total of five different classes.

After that first field day, all observations were organized into behavioral categories. This was done in order to decide which behaviors would be best to focus our efforts on during our research. Behaviors that were in consideration were:

- Re-entering/Keeping Incorrect Answer (interface action)

- Note Writing
- Talking to Self While Working
- Working Together or "Yoked Collaboration"
- Adding Numbers with Fingers
- Complaining
- Looking at Problem List
- Does Not Want to Be Wrong (would not submit answer)
- Waiting on Teacher

There were four primary requirements of the behavior that were considered during the decision process. We desired behaviors that: were unexpected, had little to no research already completedin the field,happened at least a few times, and had distinguishing physical features that make it easy to observe. We wanted to observe behaviors that might reveal how students are learning from or interacting with ASSISTments. We felt that unexpected behaviors would give us the best starting point for this type of research. Next, we wanted tofocus our research on a behavior that has not been researched thoroughly; this gives us a greater potential to enable a scientifically interesting finding.For instance, "gaming the system" is a very interesting behavior; but substantial research has already been performed in this area (Baker et al, 2004; Cocea et al, 2009). Therefore, we chose not to further research this phenomenon, though we did see it in the classroom. We found several behaviors that were both unusual and new to being studied, however, they occurred so infrequently that it was impractical to conduct an entire study based on them. Lastly, given our short time period to conduct our research, choosing behaviors that are easily identified by an observer rather than cognitive or emotional states was very important to us. This was chosen in order to help the observers avoid confusion or making observational mistakes. With these four requirements taken into careful consideration, we chose to continue our observational research on Re-entering/Keeping Incorrect Answer (interface action), Adding Numbers with Fingers, and Waiting on Teacher.

Re-entering/Keeping Incorrect Answer is a behavior related to how the student directly interacts with the ASSISTments interface. Students would submit an answer to the program, which would in turn be an incorrect answer. Instead of reworking the problem, they would simply resubmit the same answer. This happened in two very specific ways. The first way being, when the student would receive the incorrect answer cue, they would erase what they had previously submitted and retype the same exact thing, thus resubmitting the same answer again. We hypothesize that this maybe a sign of mistrust in the system. It seemed that these student felt assured in their answer and that ASSISTments was wrong, so they simply tried the "right" answer again. The second way this Re-entering happened was when a student would receive the incorrect answer cue, they would continuously hit enter or click submit in a very rapid, successive manner, thus resubmitting the same answer several times over and over again.This action would sometimes be performed by more forceful clicks on the mouse or hitting the enter key harder than usual, also students would sometimes say, "This is stupid." or other similar phrases.This leads us to believe that this type of re-entering may be a sign of frustration. It seemed that these students wanted to progress in the program and were frustrated with the set-back and having to work on the same problem.

Adding Numbers with Fingers is a behavior related to the process used by the student to calculate and solve given math problems presented in ASSISTments. When given a problem to solve, the student would use their fingers to add or subtract rather than simply recognizing the answer without using their fingers or using the online calculator provided for all students' use.

From personal experience, most students quit this behavior during the middle of elementary school or sooner; however, the students in our study are in 7th-9th grade. Also, a quick search online will show that parents become concerned with this behavior around age six and there has been an experimental study (Albayrak, 2010) attempting to teach 1st graders to quit counting on their fingers. From our research and our personal experience, we believe we are not alone in our assumption that middle school students should no longer count on their fingers. We were interested in why some students never choose to quit this behavior. Another interesting aspect of this behavior was the way in which students carried out counting on their fingers. Very rarely would a student openly count with their hand visible for all to see. Rather, they would count with their fingers in a much more subtle way; it seemed as though they were purposefully disguising the behavior. For instance, a student may keep their hand on their desk and look like they are simply tapping their fingers; however, through observation it was noticed that this “tapping” happened with one finger after another in a sequential manner, often accompanied by lips mouthing numbers with each “tap”. It was very common for this type of counting to happen with their hands under their desk or on their laps. This “hiding” of behavior suggests that the students realize the behavior is not openly accepted for their age, yet they still choose to rely on it as a tool for calculation. We do not explore the answer to this question in our study; however, it is an area for more research in the future.

Waiting on Teacher is a behavior that is partially related to students’ interaction with the software and is partially related to the students’ environment around them. As a student makes mistakes, ASSISTments should determine why they are making the mistake and provide hints in accordance with their mistakes to lead them through the process of understanding and success. If a student does not know where to start with a problem, they are able to click the hint button to get help before submitting an answer. However, students also have the freedom to try the problem and upon an unsuccessful attempt, they can then visit the hint button for help. Also, as previously mentioned, ASSISTments uses scaffolding questions to help students fully understand the given math concepts (Heffernan, 2009).

Initially, points were deducted from students for using the help button. The teacher whose students we studied felt that students were not using the help button as much as expected and attributed it to the point penalty, so she removed this rule in hopes that students would feel more free to use the program's provided help. Though the teacher feels that students reacted positively toward the change, we still saw many students wanting help directly from the teacher instead of utilizing the help built into the program. Often the teacher would ask the student if they had used the hint button yet, which they would reply that they had not. She would then instruct them to use it first; otherwise, she would not provide help. At times, even after her direct instruction to seek help from the program, they would be hesitant and say they do not need the hint button. She would explain that no points would be deducted for its use; sometimes the students would then ask if she could see how many times they used the help. She would say yes and they would often continue to not use the button.

We found that most students used ASSISTments as designed; however, there were times that students both used all of the hints and still could not understand how to do the problem, or they would not understand the hints themselves. These are other instances in which the students would require assistance directly from the teacher.

Also during our observations, we noticed that certain question's hint buttons would not work when clicked on. In the event a student needed help on those particular problems, they were given no help from the program. This naturally resulted in the student needing assistance from the teacher if they were not able to ask a peer. These situations and the situations previously mentioned contain the primary instances that students would spend time waiting on the teacher.

After determining our coding scheme and defining each behavior, our Professor provided us with handheld devices which were updated to use our coding scheme (the use of handheld devices to code classroom behavior is discussed in Baker, Moore, Kalka, Karabinos, Ashe & Yaron, 2011).We first had to go through the process of logging in, then we would enter the school, the class name (in our case, the period observed), number of students, title for our project, and then our actual name in order to differentiate between whose data is whose. After the devices had the proper information submitted and the majority of students were seated, we started our observations which consisted of three different types of rounds, being:

- Training Round
  - Once students were all logged on to the program, we together would start with one student and observe their behavior for no more than 20 seconds.
  - We then would discuss what we saw and decide together on what behavior to categorize our observation as.
  - We then both submitted the categorical data on each of our devices.
  - This process was continued together in a sequential order around the room until we arrived back at the first observed student.
- Inter-rater Reliability Round
  - Again, with the same first observed student, we would observe the student's behavior together for no more than 20 seconds.

    - We would then individually decide the category we felt the student's behavior most reflected and would separately submit that category on our individual devices.
    - This process was continued together in a sequential order around the room until we arrived back at the first observed student.
- Round Three
    - With the same first observed student, one of us would start their observations. The other observer would simply wait approximately 30 seconds and then start their observations on that same first student.
    - This process was continued in a sequential order around the room, with each observer individually studying a separate student. The observers did this until the class was close to being out of time, often getting two or three rounds of data in this manner.
    - After we finished collecting data, at the end of class, we would email the information to our advisor.

This "Three Round" process was used in every class observed. The training round served as a way to keep the observers in agreement with one another in respect to what defines a particular behavior. Also, in the instance that one observer was not noticing certain behavior signals, this round gave the observers a chance to discuss different cues to look for.The inter-rater reliability round is a way to check that each observer was actually in agreement with the definition of each behavior. The data entered by one observer was compared with the data entered for the same student by the other observer. In our case, every observation matchedwhich is unusually good agreement. This confirmed that both observers were in agreement with what the behavior definitions. The third round and all of those that followed are simply rounds to gather as much data as possible, which is why the observers were studying students individually. It is important to note that the very first observation submitted started a timer within the coding scheme; this allowed us to know precisely when each observation was made in relation to the very first observation. On average, each observation, including transition time, lasted approximately 12 seconds. We observed each student for 10 seconds and had an average transition length of 2 seconds.

Results

The participants for our observations on the second field day were 77 students, from four different classes, all taught by the same teacher, and all using the ASSISTments system. There were 45% female and 55% male. As previously stated, we only focus on three behaviors; waiting, reentering, and fingers. In Table.1 we show the total amount of time spent collecting observational data, the total number of observations, and the total percentage of each behavior.

Table.1 Percentage of behaviors based on the overall observations

Total number of students	77 students
Session time	166.5 minutes
Total number of observations	947 observations
% of waiting	10.15%
% of reentering	3.07%
% of using fingers	1.58%
% of other behaviors	85.20%

Based on our data, we found that 50% of all studentsspent time waiting on the teacher at least once, 25% of all students reentered at least once, and 11.1% of all students usedtheir fingers for counting at least once.

To determine if one class has a higher percent of behavior than the other classes we separate each class by each behavior in table.2.

Table.2: the percentage of each behavior at each class

Class	# of Students	# of Rounds	Waiting	Reenter	Fingers	?
One	24	16	4.01%	1.27%	0.63%	34.59%
Two	23	12	1.38%	0.32%	0.53%	10.04%
Three	4	12	3.81%	0.74%	0.32%	23.68%
Four	26	8	0.95%	0.74%	0.00%	16.92%

After organizing the data, we saw what appeared to be clumps of Waiting on Teacher. The most reasonable explanation that a clump of students would be left waiting is that the teacher is helping another student at that particular time; however, it is also possible for a clump of students to randomly be waiting at the same time. Therefore, we conducted Z statistical tests to see if our data held any statistical significant relationships, where a clumping of waiting students was more frequent than could be expected by chance. We conducted the Z tests in four different grouping sizes. These four grouping sizes were organized in two different ways, which were by the observer, or coder, and by the time the observation happened. This results in eight total Z tests calculations for the data set. To better understand how we organized our data, an example of the organization for each calculated Z tests can be seen in following figures:

Data organized by observer, or "coder":

W= waiting
N= Non-waiting

Figure 1: Pattern of 3/3 (WWW)

Student	Observmin	Behavior	Coder	class
0904	1370923	WAITING	BeBe	three
0714	1376635	WAITING	BeBe	three
0607	1382258	WAITING	BeBe	three

Figure 2: Pattern of 2/3 (In this example WNW)

Student	Observmin	Behavior	Coder	class
0426	517343	WAITING	BeBe	one
01010	524509	?	BeBe	one
1114	549570	WAITING	BeBe	one

Figure 3: Pattern of 3/4 (In this example WWWN)

Student	Observmin	Behavior	Coder	class
0714	1376635	WAITING	BeBe	three
0607	1382258	WAITING	BeBe	three
1205	1385261	WAITING	BeBe	three
0330	1397915	?	BeBe	three

Figure 4: Pattern of 4/5 (In this example NWWWW)

Student	Observmin	Behavior	Coder	class
0605	1527510	?	MoMo	three
0904	1529694	WAITING	MoMo	three
0714	1543121	WAITING	MoMo	three
0607	1546024	WAITING	MoMo	three
1205	1547940	WAITING	MoMo	three

Data organized by time.

Figure 5: Pattern of 3/3 (In this example WWW)

Student	Observmin	Behavior	Coder	class
1114	549570	WAITING	BeBe	one
1114	554160	WAITING	MoMo	one
1126	620020	WAITING	BeBe	one

Figure 6: Pattern of 2/3 (In this example WWN)

Student	Observmin	Behavior	Coder	class
0125	300749	WAITING	BeBe	one
0822	305858	WAITING	MoMo	one
03141	307638	?	BeBe	one

Figure 7: Pattern of 3/4 (In this example NWWW)

Student	Observmin	Behavior	Coder	class
01010	529993	?	MoMo	one
1114	549570	WAITING	BeBe	one
1114	554160	WAITING	MoMo	one
1126	620020	WAITING	BeBe	one

Figure 8: Pattern of 4/5 (In this example NWWWW)

Student	Observmin	Behavior	Coder	class
1114	549570	WAITING	BeBe	one
1114	554160	WAITING	MoMo	one
1126	620020	WAITING	BeBe	one
1126	623792	WAITING	MoMo	one
0710	627725	?	BeBe	one

The equations used to calculate z are:

- $z = \frac{x - \mu}{\sigma}$
- $x = o^{(c-1)} * (1 - o) * c * (t - c)$
- $\sigma = \sqrt{\frac{x(t-c-x)}{t-c}}$

Where "x" is the expected amount of times the defined pattern would happen, "µ" is the actual amount of times the defined pattern happened, and "σ" is the standard deviation. Standard deviation used values "o" which is the overall percentage of waiting, "c" which is the size of cluster we observed, and "t" which is the total amount of cluster observations. We calculated the Z tests in order to find the corresponding p value, which indicates the probability of the results seen if the results were due just to chance. If the p value is less than or equal to 0.05, then we state that our results are significantly significant.As seen below, Table.3 shows our calculated Z values.

Table.3: shows the value of p at each pattern

Pattern	By Coder	By Time
Pattern of 3 (3W)	**0.0001**	**0.043**
Pattern of 3 (2W, 1N)	**0.25**	**0.053**
Pattern of 4 (3W, 1N)	**0.45**	**0.43**
Pattern of 5 (4W, 1N)	**0**	**0.41**
Pattern of 6 (5W, 1N)	**0.82**	**0.82**

Table.3 shows that the patterns 3W organized by coder, 3W organized by time, and 2W with 1N organized by time are all statistically significant. This means the behavior when organized in these patterns do not happen by chance, but most likely have a reason.

Discussion

By looking at the results we can see that students waiting on the teacher happened much more than the other two behaviors observed. Though the percentages of reentering and counting on fingers are small, the actual ratio of students that did the behavior at least onceis greater than we anticipated. We thought that most instances of each behavior would happen with very few students, but many times. Instead, more students performed the action than we would have guessed without making observations. Perhaps there is a relation to these behaviors and specific problem sets. For instance, a multiple choice question seems less likely to cause the behavior of re-entering.This is an area for more research.Due to a higher percentage of Waiting on Teacher, we chose to focuson this behavior.

As seen in Table.2, the first class held the highest percentage of waiting which may be connected with the higher amount of rounds performed. The first class was very punctual, immediately logged in, and began working. This allowed us to do more rounds than with other classes. During class two, some of the computers were not working properly as students attempted to log in to their accounts. Because we waited until all students were logged in to begin our observations, we had less time to complete our observational rounds. Note that while students were waiting for help to get their computer working properly, this was not part of the data set. Class three was the period after a special testing time, so students were late to class and we did not know how many to expect. As we were well into the class period we decided to start the process and obtained as many rounds as in class two due to the very few students present. Class four appeared to be the least motivated to work. Unlike the first class that immediately went to their assigned computer and started working, most of class four waited for specific instructions to start working, whereas the teacher expected them to work without being told. Some students took advantage of the teacher's being busy at the start of a new class, and did not log in to their computer. After she had time to evaluate the classroom, she noticed this and directly told each student to log in. After this, we were able to start our observation process. Also, in this class, many students logged out earlier than in other classes, so we were not able to get as many observation rounds as in class one.

The percentage of students that waited on the teacher at least once is 50%. This may be due to class size. If the class is large, one teacher is not able to get to every person instantaneously. However, it is interesting to see that the second highest percentage of waiting was found in class three, which had significantly fewer students than the other three classes. This leads to the idea that perhaps students are working together or helping one another, regardless of the program being designed as a personal tutor (Schofield, 1995). This idea comes from the fact that though class three had only four students, they were all spaced apart from each other. This makes it more difficult to help one another and is more noticeable, so it seems natural that instead of going to another student, they would go to the teacher for help. This is a possible explanation as to why its percentage of waiting is almost as high as in class one which has six times as many students.

This brings us into the clumping pattern that was noticed when the total data was organized together. As seen in Table.3, we organized our data into two different ways, by coder and by time, and looked at different patterns of students waiting within these two categories. By organizing the data by coder and calculating the p value, we decide to either “fail to reject” orreject the assumption.The "fail to reject" terminology highlights the fact that theassumption is presumed to be true from the start of the test; if there is a lack of evidence against it, it simply continues to be assumed true.

As previously stated, the most reasonable explanation for any student waiting on the teacher is that the teacher is helping another student at that particular time. So if we find a clump organized by time, this lets us know that the teacher is busy with another student, most likely on the other side of the room because both observers are always working on the same side. If we find a clump organized by coder, this tells us that there are students adjacent to each other waiting for the teacher. This could be due to the teacher getting caught up answering questions in one area of the class and not yet able to make it to the section we they are waiting together. However, this could also be due to students helping each other. When they collectively reach a place where no one knows what to do; they then collectively wait on the teacher for assistance. This would create a clump in the data organized by coder.

Conclusion

Through our study, we found that when using ASSISTments, students re-enter their answers, count on their fingers, and spend time waiting on the teacher.In regard to students waiting on the teacher, we found that demand for the teacher on one side of the room likely results in students on the other side of the room having to wait for assistance. We also find that regardless of using a tutor program, students still work with each other which is also a likely contribution to clumped waiting. We think it might be beneficial to implement a chat system or list that allows students can ask the teacher for assistance through the computer. This would enable the teacher to help students in the order they requested assistance, rather than the closest student getting help before a farther away student. Also, we believe that shy students would feel more comfortable asking for assistance in this way, which would cause fewer students "suddenly" needing assistance when the teacher is within a certain area. (Though we did not formally study this, we did notice that shy students waited for the teacher's eye contact before raising their hand.)

This study led to new areas for further research, such as why students choose to work together in groups rather than individually as ASSISTments was designed for. In this area, data could be collected to see if students that work together perform better than students that work individually. If they perform better, then perhaps this could lead to a change in design for ASSISTments. If they perform worse than students than work individually, then studies as to why they choose to work together and how to get them to work individually could then be conducted. Also, why students choose not to use the hint button, but instead ask for the teacher's assistance could be further explored. With this, how to get reluctant students to use the hint button could also be studied. Likewise, why students count on their fingers and re-enter problems are both areas that may benefit from having more research.

References

Aleven, V., McLaren, B.M., Roll, I., & Koedinger, K.R. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education, 16,* 101-128.

Albayrak**,**Mustafa**.** (2010).An Experimental study on preventing first graders from fingers counting in basic calculations**.**http://www.investigacion-psicopedagogica.org/revista/articulos/22/english/Art_22_458.pdf**.**

Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z. (2004) Off-Task Behavior in the Cognitive**.** Tutor Classroom: When Students "Game The System". Proceedings of ACM CHI 2004: Computer-Human Interaction, 383-390.

Baker, R.S.J.d. (2007) Is Gaming the System State-or-Trait? Educational Data Mining Through the Multi-Contextual Application of a Validated Behavioral Model. Complete On-Line *Proceedings of the Workshop on Data Mining for UserModeling at the 11th International Conference on UserModeling 2007*, 76-80.

Baker, R.S.J.d., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C. (2010) Better to Be Frustrated than Bored: The Incidence, Persistence, and Impact of Learners' Cognitive-Affective States during Interactions with Three Different Computer-Based Learning Environments. *International Journal of Human-Computer Studies,* 68 (4), 223-241.

Baker, R.S. (2007). Modeling and understanding students' off-task behavior in intelligent tutoring systems.Proceedings of the ACM Conference on Human Factors in Computing Systems, 1059-1068.

Baker, R.S.J.d., Moore, G., Wagner, A., Kalka, J., Karabinos, M., Ashe, C., Yaron, D. (2011) The Dynamics Between Student Affect and Behavior Occuring Outside of Educational Software. *Proceedings of the 4th bi-annual International Conference on Affective Computing and Intelligent Interaction*.

Cocea, M., Hershkovitz, A., Baker, R.S.J.d. (2009) The Impact of Off-task and Gaming Behaviors
on Learning: Immediate or Aggregate? Proceedings of the 14th International Conference on Artificial Intelligence in Education, 507-514.

Gobert, J., Heffernan, N., Ruiz, C., & Kim, R. (2007). AMI: Science ASSISTments Meets Inquiry. Proposal funded September 1, 2007 by the National Science Foundation (NSF-DRL# 0733286).

Gobert, Janice, Heffernan, Neil, Koedinger, Ken, & Beck, Joseph.(2009). Science ASSISTments Meets Science Learning (AMSL; R305A090170).Awarded February 1, 2009 from the U.S. Dept. of Education, 2009.

Karweit, N. &Slavin, R. (1981). Measurement and modeling choices in studies of time and learning.American Educational Research Journal, 18, 157-171.

Koedinger, K. R. & Corbett, A. T. (2006). Cognitive Tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.) *The Cambridge Handbook of the Learning Sciences*, (pp. 61-78). Cambridge University Press.

Mendicino, M., Razzaq, L. & Heffernan, N. T. (2009). Comparison of Traditional Homework with Computer Supported Homework: Improving Learning from Homework Using Intelligent Tutoring Systems. Journal of Research on Technology in Education (JRTE), 41(3), 331-358.

Razzaq, L., *Patvarczki*, J., *Almeida*, S.F., *Vartak*, M., *Feng*, M., Heffernan, N.T. and Koedinger, K. (2009). The ASSISTment builder: Supporting the Life-cycle of ITS Content Creation. IEEE Transactions on Learning Technologies Special Issue on Real-World Applications of Intelligent Tutoring Systems. 2(2) 157-166

Schofield, J. W. (1995). Computers and classroom culture. New York: Cambridge University Press.