Pilot Testing of the Pathway Active Learning Environment

Christopher M. Nakamura[1], Sytil K. Murphy[1], Dean A. Zollman[1], Michael Christel[2], and Scott Stevens[2]

[1]*Kansas State University, Physics Department, 116 Cardwell Hall, Manhattan, KS 66506, USA*
[2]*Entertainment Technology Center, Carnegie Mellon University, Pittsburgh, PA 15123-3891, USA*

Abstract. We present an initial analysis of data taken to test the technical functionality and student usability of an interactive synthetic tutoring system administered online. The system allows students to ask questions and receive pre-recorded video responses from knowledgeable tutors in real-time. It logs student interactions with a timestamp and username to generate a time-resolved picture of students' use of the system. The tutoring interaction is structured by lessons covering Newton's laws. Time on-task estimates indicate that students spent about 2.5 hours working through our materials, about as much as intended. Data show students' reluctance to query the tutor or that their focus is on other aspects of the system. This suggests modifications to the system that may encourage students to take advantage of its interactive capabilities. The system combines lessons, images, and video technology designed to emulate conversation to produce a supplemental teaching tool that may be useful for studying multimedia effects on learning.

Keywords: multimedia, Newton's laws, online instruction, pathway, physics education, tutoring, video measurement
PACS: 01.40.Fk

INTRODUCTION

Some evidence indicates that different types of multimedia may affect student learning in physics [1]. We are developing an online interactive synthetic tutoring system to investigate how multimedia and technology that may simulate a social interaction may support student learning of physics. The system is aimed at high school physics students and undergraduate algebra-based physics students. Here we discuss a pilot study conducted with algebra-based physics students at Kansas State University to test the functionality of our system and establish research protocols for a full study.

THE PATHWAY ACTIVE LEARNING ENVIRONMENT

Our synthetic tutoring system, shown in Figure 1, is composed of several components. Central to the design is the interactive video tutor interface which uses Synthetic Interview (SI) technology to match students' typed natural language questions to appropriate pre-recorded video responses from experienced physics tutors [2]. The question set was made by logical extension of questions determined by student interviews.

There is evidence that a student's perceptions about the opportunity for social interaction can have positive implications for learning [3]. A goal of this project is to determine whether we can create a simulated social interaction that promotes knowledge construction.

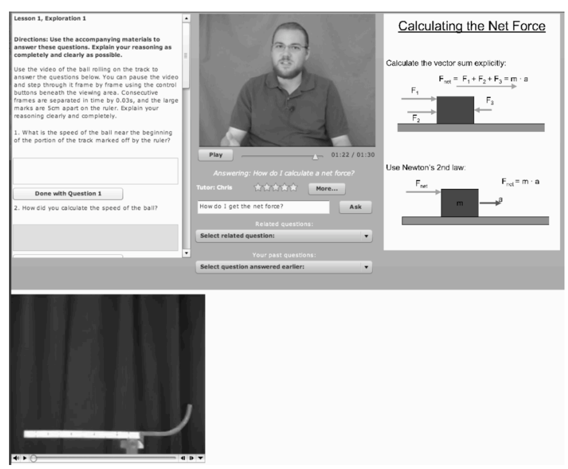

FIGURE 1. These screen captures show the PALE. Lesson questions are answered in the upper left. The questions refer to the video in the lower left. Students can query the SI tutor in the center. The tutor's responses can make use of a static image next to the tutor to illustrate the tutor's point.

Additional media can be used to supplement the tutor's video responses. When this study was conducted the supplemental media was restricted to static images, but video clips will also be used in the future.

Lesson activities utilizing video measurement and observation provide structure for the tutoring interaction, giving students a context to ask questions. The lessons were developed using three-stage learning cycles [4]. The system currently has three lessons addressing each of Newton's laws. Each lesson is designed such that we estimate students in the target populations will be able to complete it in approximately one hour. We call this system the Pathway Active Learning Environment (PALE) because of its evolution from the successful Physics Pathway system, designed to aid pre-service and in-service teachers in teaching physics topics [5].

PALE has extensive data logging capabilities, which are necessary to effectively study students' actions while using it. Students who volunteer to work with our system create a user account and must log in with a username and password to use the system. The PALE records information about each student's use of the system with a timestamp. This enables us to get a time-resolved picture of student interactions with PALE. Most important amongst these interactions are queries to the tutor and typed responses to lesson questions. Our ultimate goal is to seek out patterns in these interactions that give clues to which component combinations best help students learn physics. Before doing so we must characterize the most basic facets of student interactions with PALE.

PALE is designed to supplement students' regular instruction and to be used online, therefore characterizing unsupervised students' willingness to use the system with minimal incentive and determining the time it takes most students to work through the materials is an important step.

RESEARCH DESIGN

To collect the data discussed in this paper we recruited volunteers from an algebra-based physics course at Kansas State University. This was a one-semester course taken by architecture students in which both mechanics and electricity and magnetism are studied. The instructor offered a modest amount of extra-credit to students who used our materials. At the time that we solicited volunteers there were 188 students enrolled in the course. Forty-eight of the 188 students worked through all of the materials; 63 other students created accounts. Volunteers worked with the materials on their own time (though they could contact us for assistance) using their own (or university) computer facilities. The students were given three weeks to work through the materials, during which over 10,000 student interactions with PALE were logged. They had already studied Newton's Laws. The system's design as a tutoring system is consistent with aiding students after they have studied a concept. Pre-test scores, discussed below, indicate that mastery had not yet been achieved.

When they first logged on to PALE, students were assigned to one of three experimental groups (with sub-groups in some cases) in a round-robin fashion. The first experimental group only worked with our lesson materials, without the SI tutor. The second experimental group worked with the lesson materials and the SI tutor, but the SI tutor's responses were not supplemented by an image. Students assigned to this group were assigned to one of three SI tutors. The third experimental group worked with our lesson materials, the SI tutor, and the tutor's responses were supplemented with images. Students in this group were assigned to one of two SI tutors.

Students were given a short survey prior to working with our system in which they provided demographic information and information about their prior experience working with computers and the Internet. A 10 question multiple-choice test that used questions from the Force Concept Inventory was administered as a pre-test and post-test as well [6].

DATA ANALYSIS

A few statistics can describe students' willingness to use the PALE and the ease with which they completed the materials. These include the amount of time students take to work through the three lessons and the number of questions students are posing to the synthetic tutor as they work with the system. These will help us to ascertain whether students will actively use our system. If students are unwilling to work with the system, serious revision is necessary.

A precise measurement of the time students spent working with our system is complicated by the fact that while each interaction with PALE is time stamped and we can sum up the differences between these timestamps to get the total time the student was logged in, we cannot infallibly discern between time spent working and time spent on other activities. To attempt to address this problem we exploit the fact that, beyond some point, longer time intervals are more likely to be of the latter type, while shorter time intervals are more likely to be of the former type. We examined the time intervals between all interactions for seven of the 48 students who completed all materials to determine the maximum duration of an interaction that should be considered time on-task. Data for these seven students indicate that 96.3% of the time intervals between interactions were less than 10 minutes in duration. This saturation is shown in Figure 2. We see that long time intervals comprise a small portion of student interaction times, but could skew time on-task estimates towards the high end. We therefore set a cut-off and posit that interaction time intervals longer than the cut-off include time off-task and interactions shorter than the cut-off are time on-task. Setting the cut-off is somewhat arbitrary, but a reasonable number is between five and ten minutes. We chose a seven-minute cut-off corresponding to inclusion of about 95% of the time intervals. We then summed all student interaction time intervals considered to be on task. We found that on average students spent 149 minutes working through all the lessons. The standard deviation of the distribution was 46 minutes. The completion times ranged between 47 and 222 minutes. The histogram of student completion times is shown in Figure 3. No differences in time on-task were observed between the experimental groups.

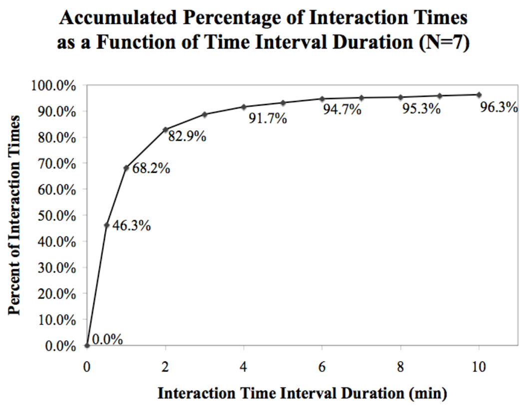

FIGURE 2. This plot of the accumulated fraction of student interaction time-intervals as a function of time-interval duration shows a clear saturation indicating the majority of interaction time-intervals are less than 10 minutes long.

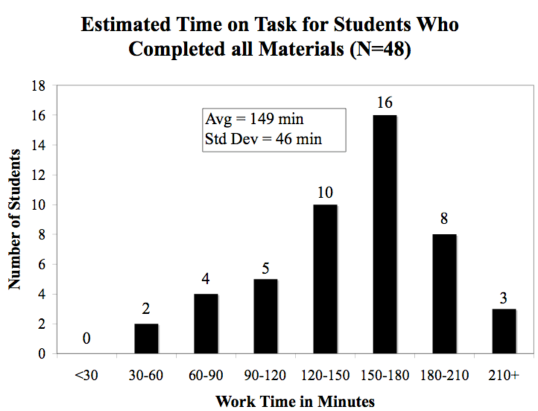

FIGURE 3. This histogram shows the distribution of estimated completion times for the 48 students who completed all materials. The average is about 2.5 hours, which is very near our target of three hours.

Eleven students were assigned to the group that saw only our lesson materials and no SI tutor. Therefore the tutor-querying behaviors can be analyzed only for the other 37 students. We observed that this group of students asked 145 questions. Figure 4 depicts a histogram that shows students' querying behaviors. We see that a significant portion of the students posed no questions at all. At the same time, more than half of the students did ask questions, though most of them asked only between one and five questions over the course of the three lessons. The number of questions posed by students in each of the experimental groups was proportional to the number of students in that group. We see nothing in the data that suggests that the specific tutor interacted with or having access to static images, promoted different question-asking behavior.

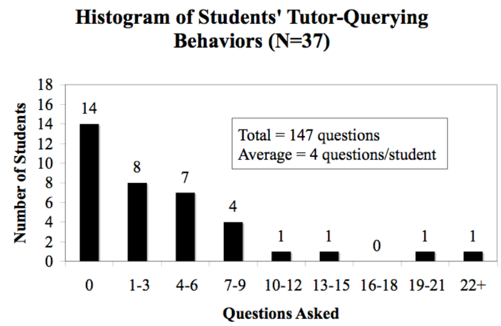

FIGURE 4. This histogram shows the number of students versus how many queries they posed to the SI tutor while using the PALE. While many students asked no questions, the majority asked at least one; the most asked was 22.

The PALE allows students to evaluate the SI tutor's responses via a five star rating system (essentially a 5-point Likert scale) as well as by submitting typed comments. Students made minimal use of this system and we have very little information about how useful they found the SI responses to be.

While a short 10 question pre-test and post-test was administered through the PALE website, we did not observe large gains from pre to post for any of the students who completed both tests. Thirty-two out of the 48 students exhibited a change of either 0, a gain of one question, or a loss of one question. The group of 48 students as a whole averaged 5.4 out of 10 on the pre-test, suggesting that although the group had studied Newton's Laws, they had not mastered the relevant material. The pre-test and post-test averages for all experimental groups were the same to well within error bars of one standard deviation and none of the differences were of practical significance.

Attrition was a problem in this pilot study. While 111 of the 188 students enrolled in the course created an account on our system, only 48 of those students finished. Attrition was somewhat evenly divided across the different experimental groups and across genders. The gender composition of the students who completed the materials matched the composition of the group that started the materials. The initial composition was 54% male and 46% female while the final composition was 52% male and 48% female. The 2% change in the final composition is a difference of 1 person, which is not practically significant.

DISCUSSION

We find that based on our method of estimating the amount of time students spent working with our system, the lessons are roughly the correct length. Students spent on average 2.5 hours on the three lessons, which nearly matches our target of three hours. At the same time we observe that most students asked far fewer questions than we had hoped. Observing students using the system in our interview facility should provide more insight into these issues.

We are taking several approaches to promote student interaction with the SI tutor. We are including a training lesson, which will prompt students to interact with the SI, and hopefully encourage them to continue using it. We are developing video clips in which the SI tutor introduces the lesson activities. By making the SI tutor initiate the conversation we hope to increase its presence in the interface and in students' minds. We are modifying the second stage of the learning cycles, in which the exploration activities are discussed in the context of the content material, so that the SI tutor delivers that material as well. This is currently done via text. This again aims at increasing the visibility of the SI tutor to encourage interaction. Further study, by interviews and by more-detailed observations of students' online use of the system, is needed to determine to what extent the revised system gives students a perception of social interaction, which [3] suggests may support learning.

Assessing how useful students perceive the SI responses is challenging. As noted in the Data Analysis section the students did not use our feedback system, and while we want information about the utility of individual SI responses, we don't want the students to spend too much time rating responses. Studying students' use of the system in our presence, during interviews should help us to resolve this issue.

Reducing attrition is also difficult. To do so requires a better understanding of why students stop using the system, which may also be gained through observing student use in an interview setting.

CONCLUSIONS & FUTURE WORK

This pilot study has given us our first look at how students interact with our synthetic tutoring system. We've seen evidence that students are neither spending an undue amount of time working with the system nor exhausting its offerings too quickly. We've determined that students are not taking advantage of the interactive features, which we feel should be the most novel and useful aspects of the system. We have explored ways of modifying and improving the system to encourage students to take advantage of the interactive nature of our tutoring system, and discussed interviews to investigate these issues. Further study will help reveal to what extent the features of this system are useful to students.

Further analysis of these data is needed to look for patterns in students interactions with PALE and to determine, in the absence of large pre-test to post-test gains, what observations may indicate effects due to the various interactive media components. Once we complete our revisions and adjustments, further studies – including interviews – will need to be done as we address the questions related to how different interactive multimedia components can aid students in learning physics.

These results give us confidence in the technical functionality of our system and give reason for cautious optimism that students will willingly use it to aid them in studying physics. Analysis of relationships between queries to the tutor, interactions with multimedia, and responses to the lesson questions may provide insight into learning benefits not revealed by pre/post testing. The system may also be useful to study the effects of social interaction on learning when the social interaction is synthetic.

ACKNOWLEDGMENTS

We acknowledge Bryan Maher for his work on PALE and Josh Gross for his work making and editing video content. This work is supported by the U.S. National Science Foundation under Grants REC-0632587 and REC-0632657.

REFERENCES

1. T. Stelzer, G. Gladding, J. P. Mestre and D. T. Brookes, *Am. J. Phys.* **77**, 184 (2009).
2. S. Stevens, and D. Marinelli in *Proceedings of the 4th IEEE Workshop on Interactive Voice Technology for Telecommunications Applications*, (IEEE, Torino, Italy, 1998), pp. 43-48.
3. S. Y. Okita, J. Bailenson, and D. L. Schwartz in *Proceedings of the 8th International Conference for the Learning Sciences,* (Lawrence Erlbaum and Associates, Utrecht, Netherlands, 2008), pp. 132-139.
4. R. Karplus and D.P. Butts, *Res. Sci. Teach.* **14**, 169 (1977).
5. B. Adrian, D. Zollman, and S. Stevens, in *Proceedings of the Physics Education Research Conference-2005*, (AIP Conference Proceedings, Melville, NY, 2006), pp. 15-18.
6. D. Hestenes, M. Wells, and G. Swackhamer, *The Phys. Teach.* **30** (3), 141 (1992).