Adding Teacher-Created Motivational Video to an ITS

Kim Kelly[1], Neil Heffernan[1], Sidney D'Mello[2], Jeffrey Namais[1], & Amber Chauncey Strain[3]
[1]Worcester Polytechnic Institute, [2]University of Notre Dame, [3]University of Memphis
nth@wpi.edu

Abstract

Many intelligent tutoring system (ITS) researchers are looking at ways to detect and to respond to student emotional states (for instance animated pedagogical agents that mirror student emotion). Such interventions are complicated to build, and do not take advantage of the potential for teachers to be part of the process. We present two studies that intervene when a student is having trouble by presenting the student with a YouTube video that is recorded by their own teacher and that delivers a motivational message to help them to persist with the learning session. We experimentally compared two different motivational interventions, which are both grounded in the literature on student affect and motivation. We also had a control condition that had no video. We found that when looking at students' self-reports on the value of mathematics, we found a main effect of condition for the value-video. In Study 2 we examined whether these 60-second videos could impact homework completion rates and found that in fact homework completion rates were higher for students in the value-video condition. The present research is suggestive of a somewhat novel use of teacher-generated content that could easily be incorporated into other ITSs.

Introduction

In recent years, connections between complex learning and emotions have received increasing attention in the intelligent tutoring literature (Arroyo et al., 2009; Conati & Maclaren, 2009; Forbes-Riley & Litman, 2011; Robison et al., 2009). Deeper understanding of affect-learning connections is needed to design more engaging educational technologies, creating the potential for affect-sensitive ITSs and more engaging educational games (De Vicente & Pain, 2002; Graesser et al., 2008;). The focus on affect is motivated by the realization that affect and cognition are complementary processes in learning environments.

In light of the affect-cognition relationship, a number of research groups have been addressing the problem of building learning environments that detect and respond to affective states such as boredom, confusion, frustration, and anxiety; Burleson & Picard, 2007; Chaffar et al. 2009; Conati & Maclaren, 2009; D'Mello et al., 2010; Robison et al. 2009). These systems use state-of-the art sensing technologies and machine learning methods to automatically detect student affect by monitoring facial features, speech contours, body language, interaction logs, language, and peripheral physiology (e.g., electromyography, galvanic skin response) (see (Calvo & D'Mello, 2010) for an overview). These affect-sensitive systems then alter their pedagogical and motivational strategies in a manner that is dynamically responsive to the sensed affective states. Some of the implemented responses to student affect include affect mirroring (Burleson & Picard, 2007), empathetic responses, and a combination of politeness, empathy, and encouragement (D'Mello et al., 2010).

The above studies represent some of the contemporary approaches to recognizing and responding to affect. Some are multi-million dollar efforts. This study takes a different approach in attempting to improve student affect using teacher-generated video messages. Our study does not attempt to create a sophisticated emotional model of the student. Instead we attempt a novel approach of leveraging the "skill builder" infrastructure in the intelligent tutoring system, ASSISTments to deliver the motivational interventions.

The present interventions are grounded in the control-value theory of academic emotions (Pekrun et al., 2006) and on perspectives that highlight the importance of students' attitudes on the malleability of intelligence (Dweck, 2006).

According to Pekrun's *control-value* theory of academic emotions, subjective appraisals of control and value of a learning activity give rise to student affect (Pekrun et al., 2006,). Subjective control pertains to the perceived influence that a learner has over the activity, while subjective value represents the perceived value of the outcomes of the activity. Engagement is increased when task value is high, but boredom dominates when learners feel like they are being forced to exert effort on an activity they do not value (Pekrun et al., 2010). Therefore, one would expect that students completing a homework assignment might experience boredom when task value is perceived as being low or when problem difficulty is considerably below or above their skill level.

In a related vein, Dweck and colleagues discovered that people with trait-based beliefs in their own competencies (also termed the "entity theory of intelligence" or the "fixed mindset") are much more likely to disengage when they are confronted with challenges and threats of failure (Dweck, 2006). In contrast, individuals who believe that their competency can be developed with practice and effort (i.e., the "incremental theory of intelligence" or the "growth mindset") are more resilient to failure and its resultant negative effect (such as frustration or anxiety). Past research suggests that messages that help students change their attributions of failure from external, fixed, and uncontrollable factors to internal, malleable, and controllable factors are effective at re-engaging students (Shores & Smith, 2010;). Students are expected to become frustrated and disengage when they perceive that the learning task is out of their control due to adherence to entity theories of intelligence.

In line with these theories, one proposed intervention consists of displaying a video message designed to help learners appraise the learning task, so as to increase its perceived value (*value-video* condition). The message reminds learners that although they may see little value in the specific current task (e.g., surface area of 3-dimensional figures), each task helps to broaden and build upon their math skills, get a good grade, progress to the next grade level, and that the content covered has the potential to be used in a future career or to improve general problem solving abilities. The other intervention (control-video) incorporates a motivational video message to the students in order to help them alter their attributions of failure, and promote a growth mindset. This intervention emphasizes the fact that failure is part of learning and if the student continues to work hard, they will be met with success. A third condition, called the *no video* condition serves as the experimental control.

Such interventions are intended to address student beliefs; targeting attributes like "grit" rather than provide academic assistance. Therefore, immediate learning gains are not to be expected. However, it has been determined that "grit", or perseverance, is essential to success in a variety of domains (Duckworth et al., 2007). This suggests that interventions that positively impact grit and increase persistence, might lead to higher long-term learning gains. Measures of perseverance include homework completion (Fredricks et al., 2011).

Much of the prior work that attempted to manipulate the fixed mindset was done in a laboratory setting. These interventions are unique in that they were tested in the context of a real math class and over the course of an actual homework assignment with ASSISTments, a math ITS (Mendicino et al., 2009; Kelly et al, submitted), thus providing ecological validity. ASSISTments allows students to complete problem sets for homework, providing tutoring and hints at the student's request. The pre and post-tests were administered during class before and after the targeted homework assignment with ASSISTments, respectively. These students were accustomed to using ASSISTments, so, while the embedded video was novel, the homework task was not.

The research questions are:

1a. Can a teacher-generated video help change students' perceptions of the value of math or encourage a "growth mind-set"? (Study 1 and 2)

1b. Can the motivational videos increase persistence, as measured by homework completion? (Study 2)

2a. Does one video (control or value) have a greater impact on student perceptions than the other? (Study 1 & 2).

2b. Does one video (control or value) have a greater impact on homework completion rates than the other? (Study 2)

Study 1

This study attempted to use an intelligent tutoring system to deliver a motivational message in the form of a video starring the teacher. The videos targeted either students' value in or their control over the learning activity. The effects of the videos on students' perceived value, control, and learning gains were measured and analyzed.

Videos

The intervention consisted of students viewing one of two possible videos. The teacher used a video camera to record herself reading the script. The videos were then uploaded to YouTube and the link to the video was an i frame embedded into the html of the problem in ASSISTments. The videos were of their math teacher delivering a motivational intervention. The *control video* had the teacher explaining that if the student works hard, they can be successful. The *value video* had the teacher explaining how important math homework is to learning as well as to their future. (See Kelly (2012b) for the actual videos). The script for the *control video* was based on Dweck's (2006) work on mindset. However, the first author, who was the teacher in the study, revised the scripts multiple times to tailor them to middle school students. For example, to convince students that their work in math class has value, a reference to problem solving in video games was made as an appropriate connection. A primary goal was to keep the videos short in order to ensure that students were engaged.

Participants and Design

Twenty-four 7th grade students in a suburban school in Massachusetts participated in this study as part of their homework in their math class. In an attempt to balance the groups, students were blocked into groups based on prior knowledge. This was done using students' current averages in math class, ordering them from least to greatest, taking three consecutive students at a time and randomly assigning them to the *no-video condition* (n=8), *control video condition* (n=8), or *value video condition* (n=8).

Procedure

Students were given a pre-test and a pre-conception survey in class. The pre-test consisted of five content questions (adding and subtracting fractions). The actual pre-test can be found in Kelly (2012a). The pre-conceptions survey consisted of five questions to measure attitudes towards value and control. The survey questions were adapted from the Academic Emotions Questionnaire (Pekrun, Goetz & Perry 2005) and responses were measured via a 5-point Likert scale. See Kelly (2012) for the questions and other study materials.

Students were then given a homework assignment in ASSISTments. As part of the students' typical routine, this assignment was completed at home as their math homework. The homework started with three content related questions that were morphologically similar to the pre-test. Students in the experimental conditions, who got at least one of these questions wrong were given a link to either the *control* video or the *value video.* The assignment continued with content related questions until the student independently and correctly answered three consecutive questions. Correctness feedback was provided and hints were available upon student request to assist in answering questions. Finally, after completing the content questions, students were given the same five survey questions that had been asked during the pre-test. To see the entire assignment, as a student would experience it, see Kelly (2012c).

The next day, students were given a post-test in class. The post-test consisted of five morphologically similar, content related questions. Five days later, a retention post-test was given consisting of five content related questions (see Kelly (2012d) for the items). Students also completed the 5-item survey at both these testing points.

Results

Five students answered the first three questions on the homework assignment correctly. They did not receive the intervention and were therefore excluded from the analysis. Only students who completed all of the tasks were included (n=16). There were no significant differences found on the pre-test or conceptions survey prior to the intervention.

The results of the survey questions are shown in Table 1. An ANOVA showed a significant main effect of *value video* ($F(2,13)=4.51$, $p=0.03$) after completing the homework assignment. Specifically, post hoc tests showed that students viewing the *value video* showed marginally higher appraisals of the perceived value of math than the students in the *no video* condition immediately after the homework assignment ($F(1,14)=3.01$, $p=0.084$). The effect size, using Hedges correction (CEM 2013), was estimated to be 1.22.. This effect was stronger the following day on the immediate post-test ($F(1,14)=4.30$, $p=0.037$), with an effect size of 1.52 and a 95% confidence interval of 0.22 to 2.82. This effect lasted to a lesser degree (effect size of 0.92), several days later on the delayed post-test ($F(1,14)=2.09$, $p=0.12$)..

There were no significant differences of meaningful trends for the *control-video*, which suggests that this intervention was less successful than the value video. We were also unable to measure perseverance through homework completion rates since all students completed the assignment.

Table 1: Mean and standard deviations (in parenthesis) for survey questions asked during each task by condition.

		Value	**Control**
Pre Test	No Video	3.6 (0.6)	4.2 (0.4)
	Value Video	3.8 (0.8)	4.4 (0.4)
	Control Video	4.3 (0.3)	4.4 (0.5)
Homework	No Video	3.6 (0.6)	4.3 (0.4)
	Value Video	4.4 (0.7)	4.4 (0.2)
	Control Video	4.1 (0.5)	4.4 (0.5)
Immediate Post	No Video	3.3 (0.6)	4.1 (0.5)
	Value Video	4.3 (0.7)	4.2 (0.5)
	Control Video	3.5 (0.6)	4.0 (1.1)
Delayed Post	No Video	3.4 (0.5)	4.2 (0.5)
	Value Video	4.1 (1.0)	3.8 (1.1)
	Control Video	3.1 (0.9)	3.8 (0.9)

Additionally, due to the small sample size and scale of the intervention, differences in learning gains were not expected. However, for completeness, mean scores were calculated and can be found in Kelly (2012). As expected, learning gains were not reliably different on the immediate post-test ($F(2,13)=0.036$, $p=0.9$), or the delayed post-test ($F(2,13)=0.067$, $p=0.9$).

Study 2

Based on the results of Study 1, we refined the procedure before replicating the study. The homework task was lengthened and made more challenging so that differences in homework completion rates could be calculated, the pre-test was embedded in the homework, and the preconception survey was administered twice instead of four times.

Participants and Design

Seventy-six 8th grade students in a suburban school in Massachusetts participated in this study as part of their homework in their algebra math class. As in Study 1, groups were balanced using blocking on prior knowledge, and were then randomly assigned to a *no-video condition* (n=29), *control video condition* (n=38), or *value video condition* (n=38).

Procedure

Students were given a homework assignment in ASSISTments. The first six questions of the assignment were used as a pre-test to assess how fluent students were in finding the surface area of cylinders. Students who answered all six questions correctly (n= 16) were then given survey questions but were not given the intervention and were therefore not included in the study.

Following the pre-test, students in the experimental condition, who answered at least one question incorrectly, were given a link to one of two possible videos on YouTube. The videos were the same from Study 1. Students were then asked to indicate how many questions they though they answered correctly. Students in the *no-video* condition were simply asked this question and were not given a link to a video.

Students then began the homework assignment, which consisted of three sections. The first section required students to find the surface area of a cylinder. Students completed this section by answering three questions in a row correctly. The second section required students to, given the surface area, find the height, radius, diameter, base area, or lateral surface area. Again, students completed this section by answering three questions in a row correctly. ASSISTments provided correctness feedback as well as offered hints on demand to students. The third section of the assignment consisted of eight survey questions to measure students' attitudes towards value and control. A 5-point Likert scale was used. Based on student feedback from Study 1, we limited the number of times we asked students these questions and therefore included additional questions for a more robust measure (see Kelly 2012e).

The following day, in-class, students were given a post-test. This test consisted of six questions that were morphologically similar to the ones given on the pre-test. Finally, six days later students were given a retention post-test, again consisting of six morphologically similar questions. Following the retention post-test, students were given the same eight survey questions. To preview an actual post-test see Kelly (2012f).

Results

Sixteen students who answered the first six questions correctly did not receive the intervention and were therefore not included in the analyses. To analyze homework completion rates, only those students who attempted the homework assignment were included in the analysis (n=60). Students in the *value video* condition had substantially higher homework completion rates (95%) when compared to students in the *control video* (70%) and *no video* conditions (69%). A Chi-Square test revealed that these differences were marginally significant ($X^2(2)=4.611$, $p=0.099$).

For the analysis regarding survey questions and learning gains, only students who completed all three tasks, and who received the intervention (in the case of the experimental condition) were included in the analysis (n=27). Responses to the value and control questions were averaged together, with question 7 being reversed scored. There was no condition effect for the *control video* or *value video*, thereby failing to replicate a finding from Study 1. It is important to note, that students had to complete the homework to receive the survey questions. Therefore, there is the potential for selection bias in these results; specifically, if students that got the *value video* were more likely to complete their homework, the sample of students in the value condition might have a different perception of the survey questions. This possible source of bias might account for our failure to replicate the findings from Study 1 pertaining to the value-video.

Similar to Study 1, no significant differences were found in percent correct on the immediate post-test ($F(2,25)=1.67$, $p=0.20$) or the delayed post-test ($F(2,25)=1.27$, $p=0.30$). See Table 2 for average perceptions and test scores by condition.

Table 2. Means and standard deviations (in parenthesis) of survey questions on each task by condition and for percent of correct answers on knowledge tests

	Control	**Value**	**No Video**
Survey Questions			
HW Control	4.0 (0.5)	3.6 (0.3)	4.1 (0.5)
HW Value	3.7 (1.1)	4.1 (0.6)	4.4 (0.8)
Post-Test Control	3.6 (0.8)	3.4 (0.5)	3.8 (0.3)
Post-Test Value	3.9 (0.9)	3.9 (0.8)	4.1 (0.7)
Test Performance			
Pretest	65% (22)	64% (21)	56% (33)
Immediate Post	87% (12)	77% (23)	92% (9)
Delayed Post	95% (5)	89% (13)	95% (4)

Contributions and Discussion

This study suggests a novel use of ITS in a highly ecologically valid situation (i.e., use in a real math classroom for homework). We hypothesized that providing motivational videos from students' own teacher might increase their motivation and problem completion rates. We found a positive effect for the value-focused video on students' perceptions of value in math in Study 1 and on homework completion rates in Study 2. The control-video did not yield any benefits over not receiving a video.

We anticipated the difficulty in obtaining statistically significant effects given that our intervention consisted of a 60-second video and the sample sizes were quite small. Clearly, future studies should be conducted implementing more substantial and long-term interventions. A longer-term study might also address the significant attrition we encountered, as there would be more opportunities for intervention and growth.

We acknowledge that the empirical contributions of this work are modest. Instead, the main contribution of this paper is to develop a proof-of-concept of cost-effective motivational interventions in ITSs. The long-term goal of this research is to identify interventions that work (from both a cognitive and motivational standpoint) and to deploy these at the appropriate time. For example, we envision a system that identifies when and why students are struggling and displays different types of prerecorded messages. There will need to be different messages for different anticipated problems; some might be content related but others might be focused on motivation. We are in the process of augmenting ASSISTments with the infrastructure to support teachers posting their own messages for common situations, and these messages can be accompanied with videos, which would allow us to test the hypothesis that individual teacher messages might be better than generic messages.

Additionally, matching the interventions to the specific needs of the student might yield more impressive outcomes. Specifically, a learner who is struggling and therefore risks disengagement due to low appraisal of control (despite valuing mathematics) might receive supportive messages or a video to help the student have a greater sense of control. This intervention might succeed in alleviating boredom in the short-term, but boredom might emerge as the session progresses (value decreases) and as the topic is mastered (control increases). In this situation, additional messages, either text or video, may be required. Similarly, offering praise for success should also be investigated.

Simple user modeling techniques might allow for more personalized messages, however more sophisticated tracking of student affect might be needed for finer grained interventions. For example, ASSISTments is currently incorporating automated detectors of boredom, engagement, and other problematic behaviors (Sao Pedro et al. (submitted). More research is needed to see if personalizing motivational feedback to individual students goals and needs is an effective way of improving ITSs.

It is also important to compare the effectiveness of teacher-generated messages versus other motivational messages. It could be that teacher created video are more effective due to the fact they come for the students teacher rather than some generic source (maybe a paid professional actor) or an animated pedagogical agent. An added benefit might be an increase in teachers' willingness to adopt the ITS if they feel they can override system default messages with their own personalized messages.

In conclusion, Study 1 showed promising results in the ability to alter student perceptions of the value of math by showing a short, teacher-generated video during a homework assignment. Study 2 showed that this intervention might also increase motivation and therefore homework completion rates, which over time could lead to improved learning. These studies were intended to serve as a proof-of-concept to test the idea that there might be potential benefits for including teacher-delivered motivational interventions within the context of mathematics homework completion with an ITS. Further research is needed to refine the interventions, augment the infrastructure to afford easy production of additional messages for a variety of situations by a variety of teachers, understand how to tailor messages to different situations and learners, and to study the long-term efficacy of this approach to facilitate learning by increasing motivation to persist in light of difficulty when it is so much easier to simply give up.

Acknowledgements

Bill and Melinda Foundation via EDUCAUSE, IES grants R305C100024 and R305A120125, and NSF grants ITR 0325428, HCC 0834847, DRL 1235958.

References

Arroyo, I., Woolf, B., Cooper, D., Burleson, W., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. In V. Dimitrova, R. Mizoguchi, B. Du Boulay & A. Graesser (Eds.), *Proceedings of 14th International Conference on Artificial Intelligence In Education* (pp. 17-24). Amsterdam: IOS Press.

Burleson, W. & Picard, R. W. (2007) Affective Learning Companions. Special Issue on Pedagogical Agents, Educational Technology, Saddle Brook, N.J., Vol. 47, no. 1, pp. 28-32, ISSN 0013-1962.

Calvo, R., D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their

Applications, IEEE Transactions on Affective Computing, v.1 n.1, p.18-37

CEM (2013). Accessed 1/28/13 at http://www.cemcentre.org/evidence-based-education/effect-size-calculator.

Chaffar, S,, Derbali, L., Frasson, C. (2009). Inducing positive emotional state in Intelligent Tutoring Systems, Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modeling, p.716-718, July 20, 2009

Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction, 19*(3), 267-303.

D'Mello, S., Lehman, B., Sullins, J, Daigle, R., Combs, R., Vogt, K., Perkins, L., Graesser, A. (2010). A time for emoting: when affect-sensitivity is and isn't effective at promoting deep learning, Proceedings of the 10th international conference on Intelligent Tutoring Systems, June 14-18, 2010, Pittsburgh, PA

De Vicente, A., & Pain, H. (2002). Informing the detection of the students' motivational state: An empirical study. In S. A. Cerri, G. Gouarderes & F. Paraguacu (Eds.), *6th International Conference on Intelligent Tutoring Systems* (pp. 933-943). San Sebastian, Spain.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. Journal of Personality and Social Psychology; Journal of Personality and Social Psychology, 92(6), 1087.

Dweck, C. (2006). *Mindset*. New York: Random House.

Forbes-Riley, K., & Litman, D. J. (2011). Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication, 53*(9-10), 1115-1136. doi: 10.1016/j.specom.2011.02.006

Fredricks, J., McColskey, W., Meli, J., Mordica, J., Montrosse, B., and Mooney, K. (2011). *Measuring student engagement in upper elementary through high school: a description of 21 instruments.* (Issues & Answers Report, REL 2011–No. 098). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Graesser, A., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes, 45* (4-5), 298-322.

Kelly, K. (2012a-f). sites.google.com/site/flairsdata/ (study materials, data, and tables)
Study 1 Materials. Accessed 11/1/12.
a. Pre-Test: http://www.webcitation.org/6CCmWVad5
b. Control Video: http://www.webcitation.org/6CCmq7KpZ & Value Video: http://www.webcitation.org/6CCmiK5pb
c. Homework Assignment: http://tinyurl.com/cbze4nr
d. Post-Test: http://www.webcitation.org/6CCoZrovS
Study 2 Materials. Accessed 11/15/12
e. Homework Assignment: http://tinyurl.com/c6k64ys
f. Post-Test: http://www.webcitation.org/6CCoqgspN

Kelly, K., Heffernan, N., Heffernan, C., Goldman, S., Pellegrino, J., & Soffer Goldstein, D. (Submitted). Estimating the effect of web-based homework. The Artificial Intelligence in Education Conference. Memphis 2013.

Mendicino, M., Razzaq, L. & Heffernan, N. T. (2009). Comparison of Traditional Homework with Computer Supported Homework. *Journal of Research on Technology in Education, 41*(3), 331-359.

Pekrun, R., Goetz, T., & Perry, R. P. (2005). Academic Emotions Questionnaire (AEQ). User's Manuel. Department of Psychology, University of Munich.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341.

Robison, J., McQuiggan, S., & Lester, J. (2009). Evaluating the consequences of affective feedback in intelligent tutoring systems. Paper presented at the International Conference on Affective Computing & Intelligent Interaction, Amsterdam.

Shores, M. L., & Smith, T. (2010). Attribution in Mathematics: A Review of Literature. *School Science and Mathematics, 110*(1), 24-30.

Sao Pedro, M., Baker, R., *Gowda, S.,* & Heffernan, N. (submitted). Towards an Understanding of Affect and Knowledge from Student Interaction with an Intelligent Tutoring System. The Artificial Intelligence in Education Conference. Memphis. 2013.