Addressing Student Behavior and Affect with Empathy and Growth Mindset

Shamya Karumbaiah
University of Massachusetts Amherst
140 Governors Drive
Amherst, MA 01003-9264
shamya@cs.umass.edu

Rafael Lizarralde
University of Massachusetts Amherst
140 Governors Drive
Amherst, MA 01003-9264
rezecib@cs.umass.edu

Danielle Allessio
University of Massachusetts Amherst
140 Governors Drive
Amherst, MA 01003-9264
allessio@umass.edu

Beverly Woolf
University of Massachusetts Amherst
140 Governors Drive
Amherst, MA 01003-9264
bev@cs.umass.edu

Ivon Arroyo
Worcester Polytechnic Institute
100 Institute Rd
Worcester, MA 01609
iarroyo@wpi.edu

Naomi Wixon
Worcester Polytechnic Institute
100 Institute Rd
Worcester, MA 01609
mwixon@wpi.edu

ABSTRACT

We present results of a randomized controlled study that compared different types of affective messages delivered by pedagogical agents. We used animated characters that were empathic and emphasized the malleability of intelligence and the importance of effort. Results showed significant correlations between students who received more *empathic messages* and those who were *more confident, more patient*, exhibited *higher* levels of *interest*, and *valued* math knowledge more. Students who received more *growth mindset* messages, *tended to get more problems correct* on their first attempt but *valued* math knowledge *less* and had lower *posttest scores*. Students who received more *success/failure* messages tended to make *more mistakes*, to be *less learning-oriented*, and stated that they were *more confused*. We conclude that these affective messages are powerful media to influence students' perceptions of themselves as learners, as well as their perceptions of the domain being taught. We have reported significant results that support the use of empathy to improve student affect and attitudes in a math tutor.

Keywords

student affect, empathy messages, growth mindset, pedagogical agents, intelligent tutor, confidence

1. INTRODUCTION

Students experience many emotions while studying and taking tests [16]. Students' emotions (such as confidence, boredom, and anxiety) can influence achievement outcomes [10, 18] and predispositions (such as low self-concept and pessimism) can diminish academic success [5, 14].

Pekrun's Control-Value Theory of emotion has been experimentally validated by classroom experiments that used student self-reports (answers to 5-point scale survey questions). These experiments provide evidence that educational interventions can reduce students' anxiety [16, 19].

Dweck's Growth Mindset Theory suggests that students who believe that intelligence can be increased through effort and persistence tend to seek out academic challenges, compared to those who view their intelligence as immutable [8, 9]. Students who are praised for their effort (as opposed to performance) are more likely to view intelligence as being malleable, and their self-esteem remains stable regardless of how hard they have to work to succeed at a task.

Hattie and Timperley [13] studied which types of feedback and conditions enable learning to flourish and which cases stifle growth. According to their study feedback is intended to help a student get from where they are to where they need to be. Graesser et al., [12] reported that there are significant relationships between the content of feedback dialogue and the emotions experienced during learning. They found significant correlations between dialog and the affective states of confusion, eureka (delight) and frustration.

Pekrun et al., [17] tested a theoretical model positing that a student's anticipated achievement feedback in a classroom setting influences his/her achievement goals and emotions. For example, *self-referential feedback*, in which a student's competence is defined in terms of self-improvement, had a positive influence on a student's mastery goal adoption. On the other hand, *normative feedback*, in which student competence is defined relative to other students' mastery goals and performance goals, had a positive influence on *performance-approach* and *performance-avoidance* goal adoption. Furthermore, feedback condition and achievement goals predicted test-related emotions (i.e., enjoyment, hope, pride, relief, anger, anxiety, hopelessness, and shame).

Teachers have limited opportunities to recognize and respond to individual student's affect in typical classrooms.

Ideally, digital learning environments can manage the delicate balance between motivation and cognition, promoting both interest and deep learning. The overwhelming majority of work on affect-aware virtual tutors has focused on modeling affect, i.e., designing computational models capable of detecting how students feel while they interact with intelligent tutoring systems [2]. While modeling affect is a critical first step, very little research exists on systematically exploring the impact of interventions on students' performance, learning, and attitudes, i.e., how an environment might respond to students emotions (e.g., frustration, anxiety, and boredom) as they arise. D'Mello and Graesser carried out close research work on empathic characters in AutoTutor, a conversational tutor that uses 3D companions to conduct dialogs in natural language with students [6, 7, 11].

1.1 MathSpring

The testbed for this research is MathSpring, an intelligent tutor that personalizes mathematics problems, provides help using multimedia, and effectively teaches students to improve in standardized test scores [4]. Learning companions (Figure 1) in MathSpring suggest to students that their effort contributes to success, and that making mistakes only means more effort is needed. Companions use about 20 different messages focused on effort and growth mindset (Table 2).

To date, MathSpring learning companions have provided positive significant effects for the overall population of students and were more effective for lower achieving students and for female students in general [2]. However, characters seemed to have been harmful to some students (e.g., high-achieving males), who had higher affective baselines at pretest time and seem to have been distracted by the characters. These results suggest that affective characters should probably be different for students who are not presently frustrated or anxious (often high achieving males). One possibility is that the behavior of the characters be adaptive to the affective state of the student.

1.2 Recognize and Respond to Affect

Previously, we evaluated the hypothesis that **tailored affective messages delivered by digital animated characters may positively impact students emotions, attitude, and learning performance**. Specifically, we identified concrete prescriptive principles about how to respond to student emotion as it occurs during online learning [1, 3]. With models of student emotion, we explored mechanisms to address negative emotions. Our models predict confidence, interest, frustration, and excitement in real-time, based on data from hundreds of students. The gold standard was students' self-reported responses to questions, such as "How confident do you feel right now?"

We found that **growth mindset messages** based on Dweck's theory [9] provide an apparent **boost in student math learning** [3], resulted in **less performance-oriented goals** (e.g., beating classmates, instead of a self-referenced focus), and **less boredom** reported on the posttest. Typically online educational systems only report correctness: "Your answer is correct/incorrect." We discovered that such **success/failure** messages are correlated to higher reported **anxiety** and **boredom**, and appear to increase **performance-oriented goals**[3]. Other results indicate that empathic characters can help decrease students' anxiety and boredom. Our results showed that: a) student anxiety and boredom can be reduced using simple 2D characters, as did D'Mello et al., (2007); b) these benefits are due primarily to empathy, and secondarily to growth mindset messages; and c) indicating only success or failure is actively **harmful** to students, in comparison to emphasizing the learning process and the importance of effort.

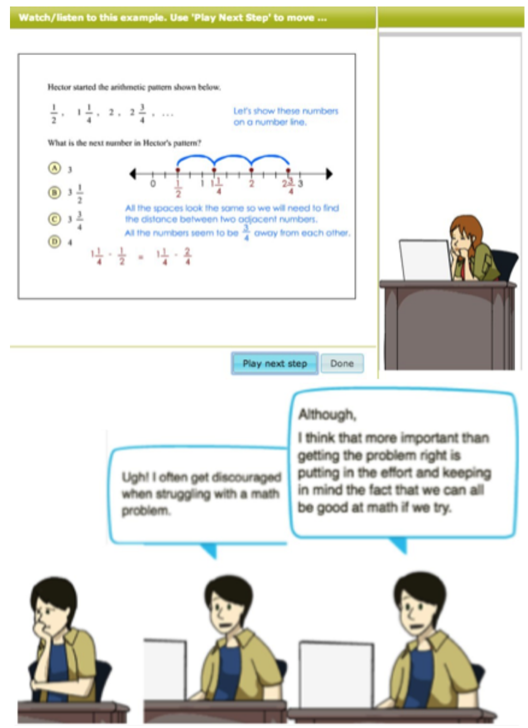


Figure 1: Learning companions respond to student actions with gestures and messages shown both as text and audio. *Above:* Companion shows high interest while the student views an example problem with solution steps shown. *Below:* Companion provides a growth mindset message, encouraging the student to put in effort to become good at math.

1.3 Research Goals

The research questions in this paper focus on identifying messages that support students' motivation to persist working on a task. Which messages (see Table 2) should a tutoring system send to students to encourage them to persist? How should agents respond to negative emotions? Should students be praised when they do well? Are the benefits to student learning and emotion due to empathic or motivational aspects of the companion? What are the results on learning and emotion of using an empathic or less empathic companion in comparison to a companion that indicates only success or failure?

Table 1: Outcomes variables measured in the experiment. The questions on the pre- and posttest were answered in a 5-point scale, going from "not at all" to "very much".

Interest - Students' interest in math. "Are you interested when solving math problems?"
Excitement - How exciting students find math. "Do you feel that solving math is exciting?"
Confusion - How confused students feel while solving math problems. "Do you feel confident that you will eventually be able to understand the Mathematics material?"
Frustration - How frustrating students find math. Average of "Do you get frustrated when solving math problems?" and "Does solving math problems make your feel frustrated?"
Learning Orientation - How much students focus on learning as opposed to performance. Average of "When you are doing math exercises, is your goal to learn as much as you can?" and "Do you prefer learning about things that make you curious even if that means you have to work harder?"
Performance Approach Goals - "Do you want to show that you are better at math than your classmates?"
Math Value - How important do students think math is. "Compared to most other activities, how important is it or you to be good at math?"
Math Liking - Measure of how much students like math. "Do you like your math class?"
Math Test Performance - Student's score on math questions that are representative of the content covered in MathSpring.

2. METHOD

We conducted a randomized controlled study to evaluate three different types of affective messages delivered by pedagogical agents (Table 2). The study took place in an urban school district in Southern California with sixty-four 6th grade students in three math classes for four class sessions, during December 2016. On part of the first and last day, students completed a pretest and posttest including questions related to various affective states, and questions about mathematics. Outcome variables measured from these questions are provided in Table 1.

Three conditions of learning companion messages were randomly assigned to students and delivered in both audio and written form in order to increase the likelihood of exposure: 1) **Empathy Condition** for 24 students, 2) **Growth Mindset Condition** for 20 students and 3) **Success/Failure Condition** for 20 students; see Table 2 for examples of the different types of messages. For all conditions, students were asked to self-report their frustration or confidence in a five-point scale every five minutes or every eight problems, which ever came first, but only after a problem was completed. The prompts were shown on a separate screen and invited students to report on their frustration or confidence.

The **Empathy** condition was set to visually reflect positive emotion with a certain probability for each math problem if the last student emotion report had a positive valence. When the most recent emotion report had a negative valence, and with a certain probability, the character first visually reflected the negative emotion; then it reported an empathy message such as "Sometimes these problems make me feel [frustrated]", and finally a connector such as "on the other hand", connected with a growth mindset message such as "I know that putting effort into problem solving and learning from hints will make our intelligence grow." Note that only students experiencing negative emotions were exposed to growth mindset messages, as opposed to the following condition.

The **Growth Mindset** condition emphasized messages that accentuate the importance of effort and perseverance in achieving success. The growth mindset condition was set to provide one of many growth mindset messages after a second incorrect attempt was made (the first incorrect attempt caused the hint button to flash), regardless of students' emotions. This condition also provided occasional growth mindset messages at the beginning of a new problem.

The **Success/Failure** condition provided both traditional success/failure messages and some more basic meta-cognitive support for when students made mistakes (e.g., acknowledging that their answer was not correct while encouraging them to use a hint). The success/failure condition provided students with a response if they answered a problem correctly and also after they made a second mistake.

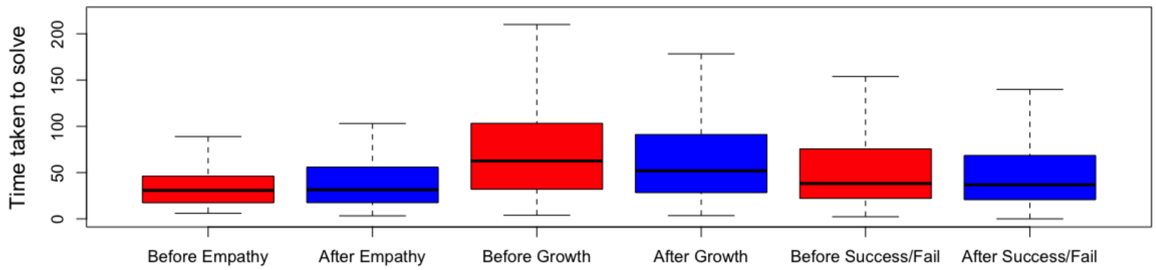
3. RESULTS

Out of the 64, three students' data were discarded due to minimal interaction with MathSpring. Across the $N = 61$ students, 21066 event log rows were recorded for three classes over four separate days, from which several behavioral features were derived and used throughout the analysis; our data and processing scripts can be found on GitHub [15]. All the students completed a pretest and posttest. Students in empathy, growth mindset and success/failure conditions received a total of 978, 763, and 882 messages respectively. Means, standard deviations and percentage shares for each type of message are given in Table 3. It is important to note that students received messages from all categories but their condition emphasized the corresponding message type. For example, a student in growth mindset condition received significantly more growth mindset messages than a student in empathy condition. This distribution of messages means that different students saw different amounts of each type of message, which allows us to perform partial correlations with respect to the counts of each message type, separating their effects.

Table 2: **Examples of messages spoken by characters.**

Condition	Message
Empathy	"Don't you sometimes get frustrated trying to solve math problems? I do. But guess what. Keep in mind that when you are struggling with are new idea or skill you are learning something and becoming smarter."
Growth Mindset	"Hey, congratulations! Your effort paid off, you got it right!" "Did you know that when we practice to learn new math skills our brain grows and gets stronger?" "Let's click on help, and I am sure we will learn something."
Success/ Failure	"Very good, we got another one right!" "Hmm. Wrong. Shall we work it out on paper?"

Figure 2: Time spent on a problem immediately before and after receiving the different categories of messages.



3.1 Partial Correlations

First, we attempted to replicate the results of our previous exploratory work [3]. For the three message types, partial correlations of the total number of each messages were measured for the nine posttest measures, controlling for the corresponding pretest measure, time spent in the tutor, and message frequency (total messages heard / time spent).

Table 4 shows the result of this analysis. We observe that with exposure to more **empathic** messages, students exhibited **higher levels of interest** and **valued math knowledge more** (rows 1 and 7). Increased interest can be viewed as analogous to the high negative correlation with boredom reported in our earlier work. With **growth mindset** messages, students **valued math knowledge less** and had **lower post test performance scores** (rows 7 and 9). With **success/failure** messages, students were **less learning-oriented** and claimed to be **more confused** (rows 6 and 3).

To further understand the dynamics, we derived some in-tutor variables and performed partial correlations shown in Table 5. The data for this analysis was derived as per student metrics based on their interaction with MathSpring. We observed that students tend to answer significantly more questions when in the **success/failure** condition and end up making more mistakes as well (rows 4 and 5). It is important to note that they also **avoid asking for hints** (row 6). It seems like these students tend to rush through the problems while being more careless. They also make **more mistakes** when they receive more growth mindset messages (row 5). This leads to simpler questions which they tend to get right in the first attempt (row 1). It appears that the students in **empathy** condition continue to **invest more time** on solving problems than rushing through the problem set. The number of problems seen by these students is significantly less (row 4).

As we see in Figure 2, students tend to spend less time on problems immediately after they receive growth mindset or success/failure messages. In contrast, the time spent on a problem increases slightly after receiving empathic messages. Students who received more empathic and growth mindset messages tend to answer fewer questions than do students who received mostly success/failure message (Figure 3). Combined with the last plot, it looks like the students in the empathy condition continue to invest more time on solving problems than rushing through the problem set.

Figure 3: Problems seen per minute across different pedagogies

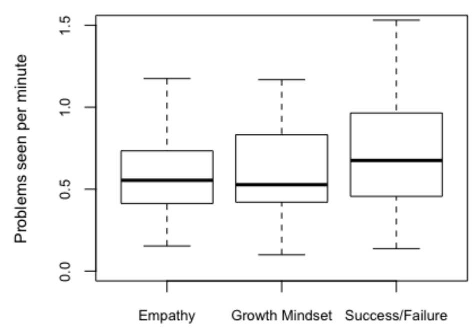


Table 3: The distribution of messages seen by students in each pedagogical conditions.

Condition	N	Empathy Messages			Growth Mindset Messages			Success/Failure Messages		
		mean	std	%	mean	std	%	mean	std	%
Empathy	21	7.48	7.0	16%	9.95	7.2	21%	29.1	22	62%
Growth Mindset	20	0.2	0.5	0.5%	10	5	26%	27.9	19.2	73%
Success/ Failure	20	1.2	1.7	2.7%	4.6	4.8	10%	38.3	26.6	86%

Table 4: Partial correlations between different types of messages seen and posttest variables (Table 1), accounting for the corresponding pretest value, time spent in tutor and message frequency.

	Variable	Empathy Messages		Growth Mindset Messages		Success/Failure Messages	
		corr	p	corr	p	corr	p
(1)	**Interest**	**0.28***	0.03	0.19	0.15	-0.20	0.14
(2)	**Excitement**	0.00	1.00	-0.07	0.60	-0.08	0.54
(3)	**Confusion**	-0.05	0.74	-0.05	0.74	**0.32***	0.02
(4)	**Frustration**	0.10	0.43	-0.08	0.54	-0.18	0.18
(5)	**Performance Approach**	-0.19	0.14	-0.05	0.70	0.20	0.12
(6)	**Learning Orientation**	0.02	0.85	0.02	0.88	**-0.24**$^+$	0.06
(7)	**Math Value**	**0.25***	0.05	**-0.22**$^+$	0.09	-0.10	0.45
(8)	**Math Liking**	0.01	0.96	0.01	0.96	0.05	0.72
(9)	**Performance**	-0.01	0.93	**-0.23**$^+$	0.07	-0.13	0.33

$^+$ $p \leq 0.10$, * $p \leq 0.05$

Table 5: Partial correlations between different types of messages seen and within-tutor variables, accounting for time spent in the tutor and message frequency.

	Variable	Empathy Messages		Growth Mindset Messages		Success/Failure Messages	
		corr	p	corr	p	corr	p
(1)	**% Problems Solved on First Attempt**	0.06	0.62	**0.34****	0.007	-0.01	0.94
(2)	**Avg Problem Difficulty**	0.07	0.61	-0.05	0.69	0.19	0.14
(3)	**Learning Gain**	-0.10	0.50	-0.07	0.63	-0.14	0.34
(4)	**Problems Seen**	**-0.23**$^+$	0.07	-0.04	0.78	**0.77****	4E-13
(5)	**Mistakes Made**	-0.01	0.91	**0.59****	6E-7	**0.30***	0.02
(6)	**Hints Per Problem**	0.10	0.43	0.16	0.22	**-0.22**$^+$	0.10

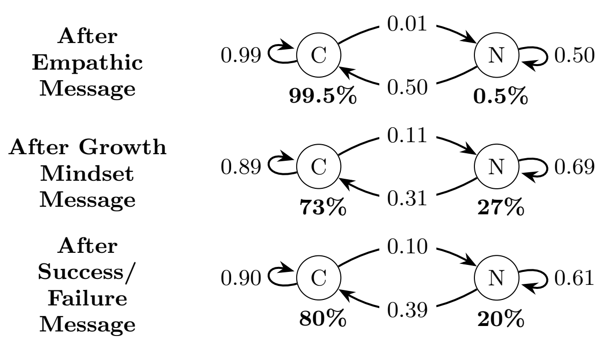
$^+$ $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$

3.2 Markov Chain Analysis

As students solve problems in the tutoring system, the learning companion comments on their attempts; the effect of these messages on student affect is sequential, but the partial correlations do not capture this. To analyze this effect, we built Markov Chain models using in-tutor student self-reports of confidence and frustration. Each model describes transitions in affective states, from one self-report to the next, where students received a particular type of character messages (empathy, growth mindset, and success/failure) between self-reports. To reduce the state space, the 5-point scale used in the self-reports was simplified to two values - confident (≥ 3), not confident (< 3); similarly for frustration.

The goal of the Markov models was not to predict emotional changes, but rather to examine whether different messages had significant effects on affect. Markov models can show the probability of transitioning between affective states, but also have a stationary distribution, which represents the amount of students that would be in each state after undergoing many transitions. For example, a group of students were to use the system for many hours and receive only empathic messages, our model suggests that 99.5% of them would be confident about learning math (Figure 4).

Figure 4: State transitions between the Confident (C) and Not Confident (N) affective states. The stationary distribution is shown below each state. Only the empathy model was significant in the likelihood ratio test ($p \leq 0.05$)



We used a likelihood ratio test to analyze the significance of these models: the probability of the null model (ignoring message type) generating the data divided by the probability of the alternate model (for a particular message type) generating the data gives a p-value. Figure 4 shows the state transitions for **confidence** in the null model and the model for confidence after receiving **empathic** messages, which was significant with $p = 0.0149$ (the other models were not significant). We also examined the stationary distributions for each model (Table 6).

Table 6: Stationary distributions in the Markov models of confidence and frustration.

Message Type	Confidence		Frustration	
	Conf	Not	Frust	Not
Empathy	**99.5%***	**0.05%***	35%	65%
Growth Mindset	74%	26%	30%	70%
Success/ Failure	80%	20%	25%	75%

*$p \leq 0.05$

4. DISCUSSION

Some of our results support the hypothesis that affective messages delivered by characters can positively impact students' emotions and affective predispositions for math problem solving. This is particularly evident for empathy, as the more empathic messages a student saw the higher their interest in mathematics problem solving, as well as their beliefs that mathematics is valuable to learn (Table 4). An analysis of student behavior suggests that students who saw a high frequency of empathic messages also tended to be more patient and cautious with problem solving, suggesting that empathic messages may encourage students to persist through adversity. Exposure to empathic messages was significantly correlated to investing time in each math problem activity, leading also to fewer problems seen per session. A positive trend is exhibited between high frequency of empathic messages and hints requested, even if not significant (Table 5). Empirical temporal models generated from students' changes in self-reports of affect, within the tutor, revealed that students receiving empathic messages have a higher likelihood to become more confident and to remain confident.

The response to growth mindset messages delivered by characters yielded mixed results. As students saw more of these kinds of messages they also succeeded more often at solving problems correctly (on the first attempt); interestingly, at the same time, they also made more mistakes. This is also desirable, as growth mindset messages emphasize that making mistakes is okay and can even help learning, legitimizing a high frequency of errors. It is possible that students were using those mistakes and hints to learn and succeed later on; a (not significant) positive trend suggests that students receiving more of these kinds of messages also asked for more hints per problem. In contrast, marginally significant effects suggest that a high frequency of growth mindset messages might be detrimental to students' perception of math value, and that their posttest performance is worse when they receive more of this kind of messages. It is hard to conclude the meaning of these marginally significant effects, especially because a previous study suggested that these messages were beneficial in general [3]. Note that empathic messages used 'growth mindset' messages also, in order to resolve the negative emotion (see Table 2). One possible explanation is that the empathic condition was so positive because it was also very selective at showing growth mindset messages for only those who experienced negative emotions. It is likely that high achieving students, or those who "felt OK", rejected growth mindset messages that they might have perceived to be unnecessary.

An important comment is that we did not expect that success/failure messages could be so harmful to students. Regardless of whether messages indicated success or failure, as students received more of these messages they also exhibited lower levels of mastery/learning orientation at posttest time. They also reported higher levels of confusion at posttest time (note that the confusion can be positive for learning within the learning experience, but not after the learning experience has concluded). Regarding behavior within the tutor, the more students were exposed to success/failure messages, the more they appeared to rush through problems, make mistakes, and request fewer hints per problem.

To summarize, empathy messages were associated with variables consistent with methodical work and an increased interest/value of mathematics. However, both growth mindset and success/failure messages appeared to be associated with a greater number of mistakes. Finally, success/failure messages themselves were associated with a whole host of concerning behaviors such as confusion with the material following posttest, reduced learning orientation, hurried work, and a reduced likelihood of requesting hints. This is consistent with Dweck's findings that growth mindset messages are superior to success/failure messages [8, 9]. Whether empathic messages in fact result in improved student performance pre to posttest will likely require larger samples than this small study (N = 61). However, students in non-empathic conditions have demonstrated significantly more mistakes in their work.

5. CONCLUSIONS

This research emphasizes the importance of understanding an intervention's effect on a student's affective state, which in turn is connected to engagement, performance, and learning. Although many researchers have focused on modeling affect, very little research effort has been put into systematically measuring the impact of the intervention on the student behavior in an adaptive learning environment. Empathic messages that respond to students' recent emotions have resulted in superior results both in improving the student interaction with the system and in the overall learning experience. Growth Mindset follows next with some positive impact on in-tutor performance but its overall effect in the short-term is questionable. Success/Failure messages are generally harmful to students: reducing learning orientation, increasing confusion, and making students more careless during the learning experience.

We conclude that affective messages delivered by characters in online tutoring environments are a very important medium for building student-tutor rapport in a virtual environment, powerful signals that influence perceptions of students themselves as learners, as well as perceptions of the domain being taught. We have reported significant results that support the use of empathy to improve student affect and attitudes in a math tutor. The long-term effect of these messages needs to be studied when the novelty of this intervention wears off. In the future, we hope to study the impact of the frequency and content of these messages. To evaluate the generalizability of these results, student populations across different demographics needs to be studied as well as the applicability of the messages to domains beyond mathematics.

6. ACKNOWLEDGMENTS

This research is supported by the National Science Foundation (NSF) 1324385 IIS/Cyberlearning DIP: Collaborative Research: Impact of Adaptive Interventions on Student Affect, Performance, and Learning. Any opinions, findings, and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

7. ADDITIONAL AUTHORS

Additional authors: Winslow Burleson (New York University, 70 Washington Square South New York, New York, 10012; email: wb50@nyu.edu).

8. REFERENCES

[1] I. Arroyo, W. Burleson, M. Tai, K. Muldner, and B. P. Woolf. Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational Psychology*, 105(4):957, 2013.

[2] I. Arroyo, D. G. Cooper, W. Burleson, B. P. Woolf, K. Muldner, and R. Christopherson. Emotion sensors go to school. In *AIED*, volume 200, pages 17–24, 2009.

[3] I. Arroyo, S. Schultz, N. Wixon, K. Muldner, W. Burleson, and B. P. Woolf. Addressing affective states with empathy and growth mindset. *6th International Workshop on Personalization Approaches in Learning Environments*, 2016.

[4] I. Arroyo, B. P. Woolf, W. Burelson, K. Muldner, D. Rai, and M. Tai. A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 24(4):387–426, 2014.

[5] L. Corno and R. E. Snow. Adapting teaching to individual differences among learners. *Handbook of research on teaching*, 3(605-629), 1986.

[6] S. D'Mello and A. Graesser. Automatic detection of learner's affect from gross body language. *Applied Artificial Intelligence*, 23(2):123–150, 2009.

[7] S. D'Mello and A. Graesser. Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(4):23, 2012.

[8] C. S. Dweck. *Self-theories: Their role in motivation, personality, and development*. Psychology Press, 2000.

[9] C. S. Dweck. Beliefs that make smart people dumb. *Why smart people can be so stupid*, 24:41, 2002.

[10] D. Goleman. Emotional intelligence. why it can matter more than fq. *Learning*, 24(6):49–50, 1996.

[11] A. C. Graesser, P. Chipman, B. C. Haynes, and A. Olney. Autotutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4):612–618, 2005.

[12] A. C. Graesser, S. K. D'Mello, S. D. Craig, A. Witherspoon, J. Sullins, B. McDaniel, and B. Gholson. The relationship between affective states and dialog patterns during interactions with autotutor. *Journal of Interactive Learning Research*, 19(2):293, 2008.

[13] J. Hattie and H. Timperley. The power of feedback. *Review of educational research*, 77(1):81–112, 2007.

[14] A. N. Kluger and A. DeNisi. Feedback interventions: Toward the understanding of a double-edged sword. *Current directions in psychological science*, 7(3):67–72, 1998.

[15] R. Lizarralde and S. Karumbaiah. A collection of scripts for processing mathspring data. `https://github.com/rezecib/MathspringDataProcessing`, 2017.

[16] R. Pekrun. Emotions and learning. *International Academy of Education. Australia: International Bureau of Education*, 2014.

[17] R. Pekrun, A. Cusack, K. Murayama, A. J. Elliot, and K. Thomas. The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction*, 29:115–124, 2014.

[18] R. Pekrun, T. Goetz, L. M. Daniels, R. H. Stupnisky, and R. P. Perry. Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3):531, 2010.

[19] R. Pekrun, E. Vogl, K. R. Muis, and G. M. Sinatra. Measuring emotions during epistemic activities: the epistemically-related emotion scales. *Cognition and Emotion*, pages 1–9, 2016.