Emotions and Learning with AutoTutor

Arthur GRAESSER, Patrick CHIPMAN, Brandon KING,
Bethany McDANIEL, and Sidney D'MELLO
Institute for Intelligent Systems, The University of Memphis
365 Innovation Drive, University of Memphis, Memphis, TN, 38152, USA

Abstract. The relationship between emotions and learning was investigated by tracking the emotions that college students experienced while learning about computer literacy with AutoTutor. AutoTutor is an animated pedagogical agent that holds a conversation in natural language, with spoken contributions by the learner. Thirty students completed a multiple-choice pre-test, a 35-minute training session, and a multiple-choice post-test. The students reviewed the tutorial interaction and were stopped at strategically sampled points for emotion judgments. They judged what emotions they experienced on the basis of the dialogue history and their facial expressions. The emotions they judged were boredom, flow (engagement), frustration, confusion, delight, surprise, and neutral. A multiple regression analysis revealed that post-test scores were significantly predicted by pre-test scores and confusion, but not by any of the other emotions.

Keywords: Tutorial dialogue, emotions, learner modeling

1. Introduction

A satisfactory understanding of the connections between emotions and complex learning is necessary to design engaging learning environments that motivate students to learn. In order to systematically investigate these relationships, we are in the process of developing a version of AutoTutor that is sensitive to both the cognitive and affective states of the learner [1, 2]. Such an affect-sensitive tutor would presumably enhance the intelligent learning environment [1, 2, 3].

AutoTutor is an intelligent tutoring system that helps students learn by holding a conversation in natural language [4]. An automated emotion classifier is necessary for AutoTutor to be responsive to learner emotions. We have previously reported some studies that collect the dialogue history, facial action units, position of their body, and other sensory channels while they learn and emote aloud [1, 5]. The features from the various modalities can be detected in real time automatically on computers, so we are currently integrating these technologies with AutoTutor.

The present study investigated the relationship between learning and the emotions that college students experience while interacting with AutoTutor on the topic of computer literacy. Whereas our previous studies of AutoTutor had learners enter their contributions through keyboard, this is the first AutoTutor study that had students make their contributions through speech. Properties of speech should be diagnostic of the learners' emotions [6].

2. Methods

The participants were 30 undergraduates at the University of Memphis who participated for extra course credit. The experiment consisted of a pre-test, an interaction with AutoTutor, a post-test, and judgments of emotions the learner experienced during the session with AutoTutor. The participants were tutored with AutoTutor on one of three major computer literacy topics: hardware, operating systems, or the internet. The pre-test and post-tests consisted of multiple choice questions that had been used in previous research on AutoTutor [7]. The 10 questions on each test tapped deep levels of reasoning, causality, and explanations. Performance in these tests was simply the proportion of questions answered correctly.

Participants interacted with AutoTutor for 35 minutes on one of the three randomly assigned topics in computer literacy. A detailed discussion of the architecture, strategies, and effectiveness of AutoTutor are provided in previous publications, as cited above. Each major topic had 6 tutoring questions that required about a paragraph of information in a good answer. We used the commercially available Dragon Naturally Speaking™ (v 6) speech recognition system for speech-to-text translation.

Students viewed their own session with AutoTutor after interacting with the tutor and completing the post-test. The judgments for a learner's tutoring session proceeded by playing a video of the learner's face along with the dialogue history. The students were instructed to make judgments on what affective states were present at three different points during the tutorial dialogue: (1) immediately after AutoTutor gave the short feedback (positive, neutral, negative) during a turn, (2) immediately before the learner started expressing his/her spoken turn, and (3) other randomly selected points in the dialogue. The students also had the option of going back in between these points and making emotion judgments. The data collection program provided a checklist of emotions for them to mark at these points. The participant was instructed to mark the affect state that was most pronounced at each point.

A list of the affective states and definitions was provided for the learners. The states were boredom, confusion, flow, frustration, delight, surprise, and neutral. These were the affective states that were most frequently experienced in previous studies of AutoTutor [5, 8, 9] that investigated emotions with alternative methods: Emote-aloud procedures during learning and observations of trained judges or peers.

3. Results and Discussion

A number of scores were computed for each of the 30 college students. These included pre-test scores, post-test scores, and the proportion of first-choice emotion judgments in each of the 7 emotion categories. The test scores varied from 0 to 1.

The post-test scores were significantly predicted by confusion (r = .490), but not the other emotion categories. We performed a multiple regression analysis that assessed the extent to which post-test scores were predicted by pre-test scores and confusion. The regression equation was significant, $F(2, 27) = 6.50$, $p < .05$, $R^2 = .326$. Confusion was a significant predictor ($p < .05$, two-tailed test, $\beta = .421$), as was also the pre-test scores ($p < .05$, one-tailed test, $\beta = .300$). This significant effect of confusion replicates a previous study that had trained judges observe student-AutoTutor interactions and record emotions every 5 minutes [8]. Confusion is a signal that the learner is experiencing cognitive disequilibrium and thinking.

This is yet another study that substantiates the importance of confusion (perplexity) in complex learning [8, 9]. When the learner is confused, they are in the state of cognitive disequilibrium, heightened physiological arousal, and more intense thought. In contrast, post-test scores were not significantly predicted by the other emotions (flow, boredom, frustration, delight, surprise) that were retrospectively identified by the learners when they viewed a recording of their learning experience. These other emotions presumably play a more prominent role in other learning environments and other populations of learners.

Our next step is to build an emotion-sensitive AutoTutor that will promote both learning gains and more engagement in the learner. AutoTutor should have different strategies and dialogue moves when the learner is confused, frustrated, bored, versus experiencing flow – the four most frequent emotions we have found in our work with AutoTutor. Since confusion is tightly linked to learning, it will be important to have the dialogue moves manage the learner's confusion productively. AutoTutor's mechanisms will need to be sensitive to cognitive and motivational characteristics of the learners in addition to their emotional state. We have already designed a computational architecture and algorithms to automatically sense the four major emotions (confusion, frustration, boredom, and flow) on the basis of the dialogue history, facial expressions, and body posture [1]. Whether this new emotion-sensitive AutoTutor will help learning awaits future research.

Acknowledgements

This research was supported by the National Science Foundation (ITR 0325428, REC 0633918). Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

References

[1] S. K. D'Mello, R. Picard and A. C. Graesser, Toward an affect-sensitive AutoTutor, *IEEE Intelligent Systems* (in press).

[2] A. C. Graesser, G. T. Jackson and B. McDaniel, AutoTutor holds conversations with learners that are responsive to their cognitive and emotional states, *Educational Technology* **47** (2007), 19-22.

[3] M. R. Lepper and J. Henderlong, Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation, *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (2000), 257-307.

[4] 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** (2005), 612-618.

[5] S. K. D'Mello, S. D. Craig, and A. C. Graesser, Predicting affective states through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue, *International Journal of Artificial Intelligence in Education* **16** (2006), 3-28.

[6] D. J. Litman and K. Forbes-Riley, Predicting student emotions in computer-human tutoring dialogues, *Proceedings of the 42nd annual meeting of the association for computational linguistics* (2004), 352-359.

[7] A. C. Graesser, S. Lu, G. T. Jackson, H. Mitchell, M. Ventura, A. Olney, and M. M. Louwerse, AutoTutor: A tutor with dialogue in natural language, *Behavioral Research Methods, Instruments, and Computers* **36** (2004), 180-193.

[8] S. D. Craig, A. C. Graesser, J. Sullins, and B. Gholson, Affect and learning: An exploratory look into the role of affect in learning with AutoTutor, *Journal of Educational Media* **29** (2004), 241-250.

[9] A. C. Graesser, B. McDaniel, P. Chipman, A. Witherspoon, S. K. D'Mello, and B. Gholson, Detection of emotions during learning with AutoTutor, *Proceedings of the 28th Annual Meetings of the Cognitive Science Society* (2006), 285-290.