Responding to Learners' Cognitive-Affective States with Supportive and Shakeup Dialogues

Conference Paper · July 2009

DOI: 10.1007/978-3-642-02580-8_65 · Source: DBLP

Responding to Learners' Cognitive-Affective States with Supportive and Shakeup Dialogues

Sidney D'Mello[1], Scotty Craig[2], Karl Fike[2],
and Art Graesser[2]

[1] Department of Computer Science
[2] Department of Psychology
University of Memphis, Memphis, TN 38152, USA
{sdmello|scraig|karlfike|a-graesser}@memphis.edu

Abstract. This paper describes two affect-sensitive variants of an existing intelligent tutoring system called AutoTutor. The new versions of AutoTutor detect learners' boredom, confusion, and frustration by monitoring conversational cues, gross body language, and facial features. The sensed cognitive-affective states are used to select AutoTutor's pedagogical and motivational dialogue moves and to drive the behavior of an embodied pedagogical agent that expresses emotions through verbal content, facial expressions, and affective speech. The first version, called the Supportive AutoTutor, addresses the presence of the negative states by providing empathetic and encouraging responses. The Supportive AutoTutor attributes the source of the learners' emotions to the material or itself, but never directly to the learner. In contrast, the second version, called the Shakeup AutoTutor, takes students to task by directly attributing the source of the emotions to the learners themselves and responding with witty, skeptical, and enthusiastic responses. This paper provides an overview of our theoretical framework, and the design of the Supportive and Shakeup tutors.

Keywords: affect, emotion, affect-sensitive AutoTutor, ITS

1. Introduction

Attempts to acquire a deep level understanding of conceptual information through effortful cognitive activities such as a systematic exploration of the problem space, generating self-explanations, making bridging inferences, asking diagnostic questions, causal reasoning, and critical thinking often lead to episodes of failure and the learner experiences a host of affective responses [1, 2]. Negative emotions are experienced when expectations are not met, failure is imminent, and important goals are blocked. For example, confusion occurs when learners face obstacles to goals, contradictions, incongruities, anomalies, uncertainty, and salient contrasts [3-5]. Unresolved confusion can lead to irritation, frustration, anger, and sometimes even rage. On the other hand, a learner may experience a host of positive emotions when misconceptions are confronted, challenges are uncovered, insights are unveiled, and complex concepts are mastered. Students that are actively engaged in the learning session may have a flow-like experience when they are so engrossed in the material that time and fatigue disappears [6]. They may also experience other positive emotions such as delight, excitement, and even one of those rare eureka (i.e. “a ha”) moments. Simply put, emotions are systematically affected by the knowledge and goals of the learner, as well as vice versa [1, 2, 7]. Cognitive activities such as causal reasoning, deliberation, goal appraisal, and planning processes operate continually throughout the experience of emotion.

Given this inextricable link between emotions and learning, it is reasonable to hypothesize that an Intelligent Tutoring System (ITS) that is sensitive to the affective and cognitive states of a learner would positively influence learning, particularly if deep learning is accompanied by confusion, frustration, anxiety, boredom, delight, flow, surprise and other affective experiences [8-11]. An affect-sensitive ITS would incorporate assessments of the students’ cognitive, affective, and motivational states into its pedagogical strategies to keep students engaged, boost self-confidence, heighten interest, and presumably maximize learning. For example, if the learner is frustrated, the tutor would need to generate hints to advance the learner in constructing knowledge, and make supportive empathetic comments to enhance motivation. If the learner is bored, the tutor would need to present more engaging or challenging problems for the learner to work on.

However, a number of technological challenges need to be overcome before the benefits of an affect-sensitive ITS can be fully realized. An affect-sensitive ITS must be fortified with sensors and signal processing algorithms. Further, these elements must be robust enough to detect the affective states of a learner within real-time constraints. The tutoring system also needs to select pedagogical and motivational moves that maximize learning while positively influencing the learners’ affect. The system might also synthesize affect through facial expressions and modulated speech.

We are in the process of implementing this two-phase strategy (affect detection and response) into AutoTutor. AutoTutor is an intelligent tutoring system that helps learners construct explanations by interacting with them in natural language and helping them use simulation environments [12]. AutoTutor helps students learn Newtonian physics, computer literacy, and critical thinking skills by presenting challenging problems (or questions) from a curriculum script and engaging in a mixed-initiative dialog while the learner constructs an answer. AutoTutor provides *feedback* to the student on what the student types, *pumps* the student for more information, *prompts* the student to fill in missing words, gives *hints*, fills in missing information with *assertions*, identifies and corrects *misconceptions* and erroneous ideas, *answers* the student’s questions, and *summarizes* topics. While the current version of AutoTutor adapts to the cognitive states of learners, the affect-sensitive AutoTutor would be responsive to both the cognitive and affective states of learners [8].

The affect-detection phase focused on the development of computational systems that monitor conversational cues, gross body language, and facial features to detect the presence of boredom, engagement, confusion, and frustration (delight and surprise were excluded because they are extremely rare). These emotions were selected on the basis of previous empirical studies that used multiple methodologies (i.e. observational, emote-aloud, retrospective judgments by multiple judges) to monitor the emotions that learners' experienced during tutoring sessions with AutoTutor [9, 13-15]. Automated affect-detection systems that detect these emotions have been integrated into AutoTutor. They have been extensively discussed in previous publications [8, 16, 17] and will not be addressed here.

The other essential component towards affect-sensitivity is to build mechanisms that empower AutoTutor to intelligently respond to these emotions, as well as to their states of cognition, motivation, social sensitivity, and so on. In essence, how can an affect-sensitive AutoTutor respond to the learner in a fashion that optimizes learning and engagement? Therefore, the next phase of our research focused on fortifying AutoTutor with the necessary pedagogical and motivational strategies to address the cognitive and affective states of the learner. This paper provides a synthesis of these research efforts.

2. Foundations of Affect Sensitivity

Boredom, confusion, and frustration are negative emotions, and are states that, if addressed appropriately, can have a positive impact on engagement and learning outcomes. Flow, on the other hand, is a highly desirable positive affective state that is beneficial to learning. Although most tutoring environments would want to promote and prolong the state of flow, any intervention on the part of the tutor runs the risk of adversely interfering with the flow experience. Therefore, the current version of the affect-sensitive AutoTutor does not respond to episodes of flow. Instead, we focus on addressing the affective states of boredom, frustration, and confusion.

At this point in science, there are no empirically proven strategies to address the presence of boredom, frustration, and confusion. Therefore, possible tutor reactions to student emotions were derived from two sources: theoretical foundations of pedagogy/affect, and recommendations made by pedagogical experts.

2.1. Theoretical perspectives

An examination of the literature provided some guidance on how best to respond to the states of boredom, confusion, and frustration. We focused on two major theoretical perspectives that address the presence of these negative emotions. These included attribution theory [18-20] and cognitive disequilibrium during learning [3-5, 13].

Attribution theory to address boredom and frustration. Attribution theory is based on the *explanations* people make to explain their success or failure. According to this theory, the cause of the success or failure can be based on three dichotomous factors: *internal* or *external; stable* or *unstable;* and *controllable* or *uncontrollable.* A basic principle of attribution theory is that a person's attributions for success or failure determine the amount of effort the person will expend on that activity in the future, and that people tend to make attributions that allow them to maintain positive views of themselves. So, success will be attributed to stable, internal, and controllable factors and major failures will be attributed to external, uncontrollable factors.

However, it is important to get learners to change this failure attribution so that their *failures* are attributed to internal, unstable factors over which they have control (e.g., effort) [19, 20]. In order to change this attribution, learners must be encouraged to focus on learning goals. People who emphasize learning goals are likely to seek challenges if they believe the challenges will lead to greater competence, and they tend to respond to failure by increasing their effort [7].

Empathy has been indicated to be an important emotional response for attributions [21]. In this case empathy serves two functions. First, displaying empathy portrays an awareness of blocked goals and a willingness to help. When displays of empathy are observed, the learner is more likely to anticipate the goals of the other displaying empathy [18]. So an example from a tutoring context would be the tutor displaying empathy for the student will cause the student to understand the tutor is attempting to help and will make the student more likely adopt the learning goals put forth by the tutor. Therefore, both boredom and frustration can be handled in similar ways using empathetic responses by the tutor.

Cognitive disequilibrium theory to address confusion. Cognitive disequilibrium is believed to play an important role in comprehension and learning processes [4, 5]. Deep comprehension occurs when learners confront contradictions, anomalous events, obstacles to goals, salient contrasts, perturbations, surprises, equivalent alternatives, and other stimuli or experiences that fail to match expectations [1, 22]. Cognitive disequilibrium has a high likelihood of activating conscious, effortful cognitive deliberation, questions, and inquiry that aim to restore cognitive equilibrium.

When a learner enters a state of confusion due to the content they are learning it is equivalent to entering cognitive disequilibrium. The tutor's first step should be to encourage the tutee to continue working so they can reach a state of equilibrium again and by doing so reach the full benefit of the state of disequilibrium. However, if the learner persists in a state of cognitive disequilibrium for too long the tutor should display empathy with the learner's attempts, thereby acknowledging their attempts to reach their goals and direct them out of the state of confusion before they give up.

2.2. Recommendations by pedagogical experts

In addition to theoretical considerations, the assistance of experts was enlisted to help create the set of tutor responses. Two experts in pedagogy, with approximately a decade of related experience each, were provided with excerpts from real AutoTutor dialogues (including both the tutor and student dialogue content, screen capture of the learning environment, and video of the student's face as illustrated in Figure 1). There were approximately 200 excerpts averaging around 20 seconds in length, each of which included an affective response by the student. The experts were instructed to view each of the excerpts and provide an appropriate follow-up response by the tutor. These example responses were placed into similar groups that loosely resembled production rules. For example, if a student is frustrated then the tutor should provide encouragement to continue and establish a small sub-goal, perhaps a hint or simplified problem. The tutor might also provide motivational and empathetic statements to alleviate frustration because this approach has been shown to be quite effective in reducing frustration [21].

3. Strategies to Respond to Learners' Affective States

We created a set of production rules that addressed the presence of boredom, confusion, and frustration by amalgamating perspectives from attribution theory and cognitive disequilibrium theory with the recommendations made by the experts. Although the rules created by the pedagogical experts allowed for any possible action on the part of the tutor, AutoTutor can only implement a portion of those actions. For example, one possibility to alleviate boredom would be to launch an engaging simulation or a seductive, serious game. However, the current version of the tutor does not support simulations or gaming, so such a strategy is not immediately realizable. Consequently, we only selected production rules that could be implemented by AutoTutor's current actions which include feedback delivery (positive, negative, neutral), a host of dialogue moves (hints, pumps, prompts, assertions, and summaries), and facial expressions and speech modulation by AutoTutor's embodied pedagogical agent (EPA).

The production rules were designed to map dynamic assessments of the students' cognitive and affective states with tutor actions to address the presence of the negative emotions. There were five parameters in the student model and five parameters in the tutor model. The parameters in the student model included: (a) the current emotion detected, (b) the confidence level of that emotion classification, (c) the previous emotion detected, (d) a global measure of student ability (dynamically updated throughout the session), and (e) the conceptual quality of the student's immediate response. AutoTutor incorporates this five-dimensional assessment of the student and responds with: (a) feedback for the current answer, (b) an affective statement, (c) the next dialogue move, (d) an emotional display on the face of the EPA, and (e) emotionally modulating the voice produced by AutoTutor's text-to-speech engine.

As a complete example, consider a student that has been performing well overall (high global ability), but the most recent contribution wasn't very good (low current contribution quality). If the current emotion was classified as boredom, with a high probability, and the previous emotion was classified as frustration, then AutoTutor might say the following: "Maybe this *topic* is getting old. I'll help you finish so we can try something new". This is a randomly chosen phrase from a list that was designed to indirectly address the student's boredom and to try to shift the topic a bit before the student becomes disengaged from the learning experience. This rule fires on several different occasions, and each time it is activated AutoTutor will select a dialogue move from a list of associated moves. In this fashion, the rules are context sensitive and are dynamically adaptive to each individual learner.

The subsequent section discusses each of the major components of the affect-sensitive AutoTutor. These include the short feedback, an emotional or motivational expression that is sensitive to the learners' affective and cognitive states, an emotionally expressive facial display, and emotionally modulated speech.

3.1. Short Feedback

AutoTutor provides short feedback to each student response. The feedback is based on the semantic match between the response and the anticipated answer. There are five levels of feedback: positive, neutral-positive, neutral, neutral-negative, and negative. Each feedback category has a set of predefined expressions that the tutor randomly selects from. “Good job” and “Well done” are examples of positive feedback, while “That is not right” and “You are on the wrong track” are examples of negative feedback. In addition to articulating the textual content of the feedback, the affective AutoTutor also modulates its facial expressions and speech prosody. Positive feedback is delivered with an *approval* expression (big smile and big nod). Neutral positive feedback receives a *mild approval* expression (small smile and slight nod). Negative feedback is delivered with a *disapproval* expression (slight frown and head shake), while the tutor makes a *skeptical* face when delivering neutral-negative feedback (see Figure 1). No facial expression accompanies the delivery of neutral feedback.

Fig. 1. Affect synthesis by embodied pedagogical agents

3.2. Emotional Response

After delivering the feedback, the affective AutoTutor delivers an emotional statement if it senses that the student is bored, confused, or frustrated. A non-emotional discourse marker (e.g. “Moving on”, “Try this one”) is selected if the student is neutral. We are currently implementing two pedagogically distinct variants of the affect-sensitive AutoTutor. These include a Supportive and a Shakeup AutoTutor.

Supportive AutoTutor. The supportive AutoTutor responds to the learners’ affective states via empathetic and motivational responses. These responses always attribute the source of the learners’ emotion to the material instead of the learners’ themselves. So the supportive AutoTutor might respond to mild boredom with “This stuff can be kind of dull sometimes, so I'm gonna try and help you get through it. Let's go”. A more encouraging response is required for severe boredom (“Let's keep going, so we can move on to something more exciting”). An important point to note is that the supportive AutoTutor never attributes the boredom to the student. Instead, it always blames itself or the material.

A response to confusion would include attributing the source of confusion to the material (“Some of this *material* can be confusing. Just keep going and I am sure you will get it”) or the tutor itself (“I know *I* do not always convey things clearly. I am always happy to repeat myself if you need it. Try this one”). If the level of confusion is low or mild, then the pattern of responses entails: (a) acknowledging the confusion, (b) attributing it to the material or tutor, and (c) keeping the dialogue moving forward via hints, prompts, etc. In cases of severe confusion, an encouraging statement is included as well.

Similarly, frustration receives responses that attribute the source of the frustration to the material or the tutor coupled with an empathetic or encouraging statement. Examples include: “*I* may not be perfect, but I'm only human, right? Anyway, let's keep going and try to finish up this problem.”, and “I know this *material* can be difficult, but I think you can do it, so let's see if we can get through the rest of this problem.”

Shakeup AutoTutor. The major difference between the shakeup AutoTutor and the supportive AutoTutor lies in the source of emotion attribution. While the supportive AutoTutor attributes the learners’ negative emotions to the material or itself, the shakeup AutoTutor directly attributes the emotions to the learners. For example, possible shakeup responses to confusion are “This material has got *you* confused, but I think you have the right idea. Try this…” and “*You* are not as confused as you might think. I'm actually kind of impressed. Keep it up”.

Another difference between the two versions lies in the conservational style. While the supportive AutoTutor is subdued and formal, the shakeup tutor is edgier, flaunts social norms, and is witty. For example, a supportive response to boredom would be “Hang in there a bit longer. Things are about to get interesting.”. The shakeup counterpart of this response is “Geez this stuff sucks. I'd be bored too, but I gotta teach what they tell me”.

3.3. Emotional Facial Expressions

Seven facial expressions were generated for the affective AutoTutor. These include: approval, mild approval, disapproval, empathy, skepticism, mild enthusiasm, and high enthusiasm. The *Short Feedback* section lists some of the conditions upon which these expressions are triggered. The supportive and shakeup responses are always paired with the appropriate expression, which can be neutral in some cases.

Example affective displays are illustrated in Figure 1. The facial expressions in each display were informed by Ekman's work on the facial correlates of emotion expression [23]. For example, empathy is a sense of understanding displayed to the user. This is manifested by an inner eyebrow raise, eyes open, and lips slightly pulled down at the edges (action units 1, 5, 15; [24]). Skepticism is a combination of confusion and curiosity, characterized by a furrowing of the brow, an eye squint, and one outer eyebrow is raised (action units 2, 4, 7)[25]. These displays were created with the Haptek Software Development Kit.

3.4. Emotionally Modulated Speech

The facial expressions of emotion displayed by AutoTutor are augmented with emotionally expressive speech synthesized by the agent. The emotional expressivity is obtained by variations in pitch, speech rate, and other prosodic features. Previous research has led us to conceptualize AutoTutor's affective speech on the indices of pitch range, pitch level, and speech rate [26].

4. Conclusions

We have described a new version of AutoTutor that aspires to be responsive to learners' affective and cognitive states. The affect-sensitive AutoTutor aspires to keep students engaged, boost self-confidence, and presumably maximize learning by narrowing the communicative gap between the highly emotional human and the emotionally challenged computer. We are currently conducting a study that evaluates the pedagogical effectiveness of the two affect-sensitive versions of AutoTutor when compared to the original tutor. This original AutoTutor has a conventional set of fuzzy production rules that are sensitive to cognitive states of the learner, but not to the emotional states of the learner. Both versions of the improved AutoTutor are sensitive to the learners' affective states in distinct ways. The obvious prediction is that learning gains and the learner's impressions should be superior for the affect-sensitive versions of AutoTutor. In addition to testing for learning gains, we will also compare learners' engagement levels while interacting with the different versions of AutoTutor. We will also test if personality differences predict preference for Supportive versus Shakeup AutoTutor.

The affect-sensitive AutoTutor represents one out of a handful of related efforts made by a number of researchers who have a similar vision [21, 27-30]. Our unified vision is to advance education, intelligent learning environments, and human-computer interfaces by optimally coordinating cognition and emotions. Whether the affect-sensitive AutoTutor positively influences learning and engagement awaits further development and empirical testing.

Acknowledgments

This research was supported by the National Science Foundation (REC 0106965 and ITR 0325428). 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] G. Mandler, "Another theory of emotion claims too much and specifies too little," *Current Psychology of Cognition,* vol. 4, pp. 84-87, 1984.

[2] N. L. Stein, & Levine, L. J. , "Making sense out of emotion," in *Memories, thoughts, and emotions: Essays in honor of George Mandler* A. O. W. Kessen, & F, Kraik (Eds.), Ed. Hillsdale, NJ: Erlbaum, 1991, pp. 295-322.

[3] L. Festinger, *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press, 1957.

[4] A. C. Graesser and B. A. Olde, "How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down," *Journal of Educational Psychology,* vol. 95, pp. 524-536, Sep 2003.

[5] J. Piaget, *The origins of intelligence*. New York: International University Press, 1952.

[6] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York: Harper and Row, 1990.

[7] C. S. Dweck, "Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways)." in *Improving academic achievement: Impact of psychological factors on education* J. Aronson, Ed. Orlando, FL: Academic Press, 2002, pp. 61-87.

[8] S. D'Mello, A. Graesser, and R. W. Picard, "Towards an Affect-Sensitive AutoTutor," *Intelligent Systems, IEEE,* vol. 22, pp. 53-61, 2007.

[9] A. C. Graesser, Chipman, P., King, B., McDaniel, B., & D'Mello, S. , "Emotions and learning with AutoTutor," in *13th International Conference on Artificial Intelligence in Education* 2007, pp. 569-571.

[10] M. Lepper and M. Woolverton, "The wisdom of practice: Lessons learned from the study of highly effective tutors," in *Improving academic achievement: Impact of psychological factors on education,* J. Aronson, Ed. Orlando, FL: Academic Press, 2002, pp. 135-158.

[11] R. Picard, *Affective Computing*. Cambridge, Mass: MIT Press, 1997.

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

[13] S. D. Craig, A. C. Graesser, J. Sullins, and J. Gholson, "Affect and learning: An exploratory look into the role of affect in learning," *Journal of Educational Media* vol. 29, pp. 241-250, 2004.

[14] S. K. D'Mello, S. D. Craig, J. Sullins, and A. Graesser, "Predicting Affective States expressed through an Emote-Aloud Procedure from AutoTutor's Mixed-Initiative Dialogue," vol. 16, pp. 3-28, 2006.

[15] A. C. Graesser, McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. , "Detection of Emotions during learning with AutoTutor," in *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Mahwah, NJ: , 2006, pp. 285-290.

[16] S. K. D'Mello, S. D. Craig, A. Witherspoon, B. McDaniel, and A. C. Graesser, "Automatic detection of learner's affect from conversational cues," *User Modeling and User-Adapted Interaction,* vol. 18, pp. 45-80, Feb 2008.

[17] S. D'Mello and A. Graesser, "Automatic Detection of Learners' Affect from Gross Body Language," vol. 23, pp. 123 - 150, 2009.

[18] C. D. Batson, C. L. Turk, L. L. Shaw, and T. R. Klein, "Information Function of Empathic Emotion - Learning That We Value the Others Welfare," *Journal of Personality and Social Psychology,* vol. 68, pp. 300-313, Feb 1995.

[19] F. Heider, *The Psychology of Interpersonal Relations*. New York: John Wiley & Sons, 1958.

[20] B. Weiner, *An attributional theory of motivation and emotion*. New York: Springer-Verlag, 1986.

[21] W. Burleson and R. Picard, "Evidence for Gender Specific Approaches to the Development of Emotionally Intelligent Learning Companions," *IEEE Intelligent Systems,* vol. 22, pp. 62-69, 2007.

[22] R. C. Schank, *Explanation Patterns: Understanding Mechanically and Creatively*. Hillsdale, NJ: Erlbaum, 1986.

[23] P. Ekman and W. V. Friesen, *The Facial Action Coding System: A Technique For The Measurement Of Facial Movement*. Palo Alto: Consulting Psychologists Press, 1978.

[24] N. Chovil, "Discourse-oriented facial displays in conversation," *Research on Language and Social Interaction,* vol. 25, pp. 163-194, 1991.

[25] S. D. Craig, S. D'Mello, A. Witherspoon, and A. Graesser, "Emote aloud during learning with AutoTutor: Applying the facial action coding system to cognitive-affective states during learning," *Cognition & Emotion,* vol. 22, pp. 777-788, 2008.

[26] T. Johnstone and K. R. Scherer, "Vocal communication of emotion," in *Handbook of Emotions*, 2nd ed, M. Lewis and J. Haviland-Jones, Eds. New York: Guilford Press, 2000, pp. 220-235.

[27] C. Conati, "Probabilistic assessment of user's emotions in educational games," *Applied Artificial Intelligence,* vol. 16, pp. 555-575, Aug-Sep 2002.

[28] D. J. Litman and K. Forbes-Riley, "Predicting student emotions in computer-human tutoring dialogues," in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics* Barcelona, Spain: Association for Computational Linguistics, 2004.

[29] S. W. McQuiggan, B. W. Mott, and J. C. Lester, "Modeling self-efficacy in intelligent tutoring systems: An inductive approach," *User Modeling and User-Adapted Interaction,* vol. 18, pp. 81-123, Feb 2008.

[30] B. Woolf, W. Burleson, and I. Arroyo, "Emotional Intelligence for Computer Tutors," in *Workshop on Modeling and Scaffolding Affective Experiences to Impact Learning at 13th International Conference on Artificial Intelligence in Education*, Los Angeles, USA, 2007, pp. 6-15.