Embodied Affect in Tutorial Dialogue: Student Gesture and Posture

Joseph F. Grafsgaard[1], Joseph B. Wiggins[1], Kristy Elizabeth Boyer[1],
Eric N. Wiebe[2], and James C. Lester[1]

[1] Department of Computer Science, North Carolina State University
[2] Department of STEM Education, North Carolina State University
Raleigh, North Carolina, USA
{jfgrafsg,jbwiggi3,keboyer,wiebe,lester}@ncsu.edu

Abstract. Recent years have seen a growing recognition of the central role of affect and motivation in learning. In particular, nonverbal behaviors such as posture and gesture provide key channels signaling affective and motivational states. Developing a clear understanding of these mechanisms will inform the development of personalized learning environments that promote successful affective and motivational outcomes. This paper investigates posture and gesture in computer-mediated tutorial dialogue using automated techniques to track posture and hand-to-face gestures. Annotated dialogue transcripts were analyzed to identify the relationships between student posture, student gesture, and tutor and student dialogue. The results indicate that posture and hand-to-face gestures are significantly associated with particular tutorial dialogue moves. Additionally, two-hands-to-face gestures occurred significantly more frequently among students with low self-efficacy. The results shed light on the cognitive-affective mechanisms that underlie these nonverbal behaviors. Collectively, the findings provide insight into the interdependencies among tutorial dialogue, posture, and gesture, revealing a new avenue for automated tracking of embodied affect during learning.

Keywords: Affect, gesture, posture, tutorial dialogue.

1 Introduction

Recent years have seen a growing recognition of the central role of affect and motivation in learning. In particular, nonverbal behaviors such as posture and gesture provide key channels signaling affective and motivational states. Insights into how systems may leverage these nonverbal behaviors for intelligent interaction are offered by a growing body of literature [1–5]. Within the intelligent tutoring systems literature, nonverbal behaviors have been linked to cognitive-affective states that impact learning [6–8].

A rich body of work has explored the moment-by-moment effects of these learning-centered affective states. Numerous techniques and tools have been applied to recognize affect, including human judgments [6, 9], computer vision techniques [4, 9, 10], sensors

K. Yacef et al. (Eds.): AIED 2013, LNAI 7926, pp. 1–10, 2013.

[11], and speech [8]. There has even been work toward identifying affect in the absence of rich data streams, instead using interaction log data [12]. The abundant utility of these techniques has been illustrated by their use in a number of affectively adaptive tutoring systems [7, 8].

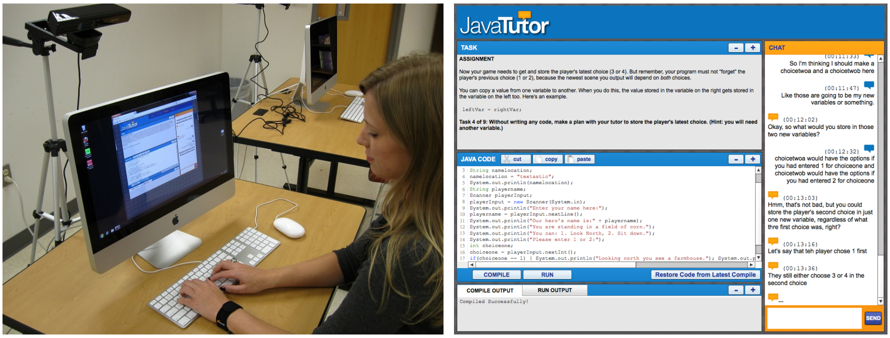
Although there has been substantial progress toward integrating affective data streams into intelligent learning environments, the field does not yet have a clear understanding of affective expression across multiple modalities. Some modalities, such as facial expression, are relatively well-explored [1, 3], while others are subjects of significant active research. For instance, posture has been used as an affective feature in multiple systems, but interpretation of postural movements is very complex [2, 9]. Early work focused on postural movement as a signal; for example, pressure-sensitive chairs have long been used for fine-grained measurement of posture [7, 13]. Early studies of posture have indicated that the signal is involved in numerous cognitive-affective states, such as boredom, focus, and frustration [7, 13]. Over the years, a replicated result in analyses of postural movement has arisen: increases in postural movement are linked with negative affect or disengagement [6, 7, 9, 14, 15]. There have also been recent developments in techniques for tracking postural movement. Posture can now be tracked in both two-dimensional [9, 14] and three-dimensional video [15] using computer vision. These computer vision-based approaches have the advantage of directly identifying postural components such as body lean angle and slouch factor [14] that were indirectly measured in the signals from pressure-sensitive chairs.

In contrast to posture, affective gestural displays have recently begun to be investigated. There is abundant cultural and anecdotal evidence for the importance of gestures [16], yet empirical research results on the cognitive-affective states underlying gesture are sparse. A system trained on acted expressions of cognitive-affective states relied on combinations of facial expression and gesture features [4], with meaning ascribed by human judges. Gestures have also been tangentially reported on in the intelligent tutoring systems community [6, 7, 17], but other phenomena were the primary focus of those studies. A recent study investigated different categories of hand-over-face gestures, with the researchers providing possible interpretations ranging over cognitive-affective states such as thinking, frustration, or boredom [5]. More recently, a hand-to-face gesture tracking algorithm was developed using the Kinect depth camera [15]. This algorithm distinguishes between one or two hands contacting the lower face. Initial analyses of these hand-to-face gestures indicated that one-hand-to-face gestures may be associated with less negative affect, while two-hands-to-face gestures may be indicative of reduced focus [15].

This paper presents an analysis of posture and gesture within computer-mediated textual tutorial dialogue. Utilizing automated algorithms that measure postural quantity of motion, one-hand-to-face gestures, and two-hands-to-face gestures, we examine the interdependencies between dialogue acts and student posture and gesture in order to identify ways in which the nonverbal behaviors may influence or be influenced by dialogue. Additionally, we report groupwise differences in nonverbal behavior displays, finding that students with lower self-efficacy tend to produce more two-hands-to-face gestures. We discuss the implications of these findings as a step toward understanding the embodied affect that intertwines with tutorial dialogue.

2 Corpus Annotation and Nonverbal Behavior Tracking

The corpus consists of computer-mediated tutorial dialogue for introductory computer science. Students (*N*=42) and tutors interacted through a web-based interface that provided learning tasks, an interface for computer programming, and textual dialogue. The participants were university students in the United States, with average age of 18.5 years (*stdev*=1.5). The students voluntarily participated for course credit in an introductory engineering course, with no computer science knowledge required. Substantial self-reported prior programming experience was an exclusion criterion. Each student was paired with a tutor for a total of six sessions on different days, limited to forty minutes each session. Recordings of the sessions included database logs, webcam video, skin conductance, and Kinect depth video. The Kinect recording rate was set to approximately 8 frames per second to reduce storage requirements. The student workstation configuration and tutoring interface are shown in Figure 1.



ASSIGNMENT

Now your game needs to get and store the player's latest choice (3 or 4). But remember, your program must not "forget" the player's previous choice (1 or 2), because the newest scene you output will depend on *both* choices.

You can copy a value from one variable to another. When you do this, the value stored in the variable on the right gets stored in the variable on the left too. Here's an example.

```
leftVar = rightVar;
```

Task 4 of 9: Without writing any code, make a plan with your tutor to store the player's latest choice. (Hint: you will need another variable.)

```
String namelocation;
namelocation = "textastic";
System.out.println(namelocation);
String playername;
Scanner playerInput;
playerInput = new Scanner(System.in);
System.out.println("Enter your name here:");
playername = playerInput.nextLine();
System.out.println("Our hero's name is:" + playername);
System.out.println("You are standing in a field of corn.");
System.out.println("You can: 1. Look North, 2. Sit down.");
System.out.println("Please enter 1 or 2:");
int choiceone;
choiceone = playerInput.nextInt();
if(choiceone == 1) { System.out.println("Looking north you see a farmhouse."); System.out.p
```

```
Compiled Successfully!
```

(00:11:33)
So I'm thinking I should make a choicetwoa and a choicetwob here

(00:11:47)
Like those are going to be my new variables or something.

(00:12:02)
Okay, so what would you store in those two new variables?

(00:12:32)
choicetwoa would have the options if you had entered 1 for choiceone and choicetwob would have the options if you had entered 2 for choiceone

(00:13:03)
Hmm, that's not bad, but you could store the player's second choice in just one new variable, regardless of what thre first choice was, right?

(00:13:16)
Let's say that teh player chose 1 first

(00:13:36)
They still either choose 3 or 4 in the second choice

...

Fig. 1. JavaTutor student workstation and tutoring interface

Prior to the first session, students completed a main pretest and pre-survey, which included an instrument for domain-specific self-efficacy (six Likert-scale items adapted from [18]). Before each session, students completed a content-based pretest. After each session, students answered a post-session survey and posttest (identical to the pretest). This paper presents analyses of data from the first session.

Dialogue acts were annotated using a parallel coding scheme that was applied to both tutor and student utterances. The coding scheme used here is an update to a prior task-oriented dialogue annotation scheme [19]. Three annotators tagged a subset of the corpus (*N*=36). Fourteen percent of these annotated sessions were doubly annotated, with a resulting average agreement across dialogue acts of Cohen's κ=0.73. The dialogue act tags and frequencies in the corpus are shown in Table 1.

Table 1. Dialogue act tags ordered by frequency in the corpus (*S*=student, *T*=tutor)

Act	Example Tutor Utterances	*S*	*T*
STATEMENT	*"java does things in the order you say."*	282	1255
QUESTION	*"Any questions so far?"*	213	630
POSITIVE FEEDBACK	*"great debugging!"*	2	539
DIRECTIVE	*"change that in all three places"*	-	252
HINT	*"it is missing a semicolon."*	-	223
ANSWER	*"yes, now line 1 is a comment."*	547	162
ACKNOWLEDGMENT	*"alright" "okay" "Yes"*	323	68
LUKEWARM FEEDBACK	*"Right, nearly there"*	-	32
NEGATIVE FEEDBACK	*"no" "nope"*	-	19
CORRECTION	Repairing a prior utterance: *"*can use"*	11	15
REQUEST CONFIRMATION	*"Make sense?" "okay?"*	6	14
OTHER	*"LOL"*	11	6
REQUEST FOR FEEDBACK	*"How does that look?"*	11	1

Recently developed posture and gesture tracking techniques were applied to the recorded Kinect depth images. The posture tracking algorithm compares depth pixels in three regions at the center of the depth image (head, mid torso, and lower torso) and selects depth pixel distances representative of each region. The gesture detection algorithm performs a surface propagation across the head and connected surfaces to identify hand-to-face gestures. The posture tracking algorithm was previously evaluated to be 92.4% accurate, while gesture tracking was found to be 92.6% accurate [15]. The tracking algorithms were run on all sessions, but four sessions had no Kinect recordings due to human error (*N*=38). The combined corpus of dialogue acts and nonverbal tracking data contains 32 sessions. Sample output of posture and gesture tracking is shown in Figure 2.

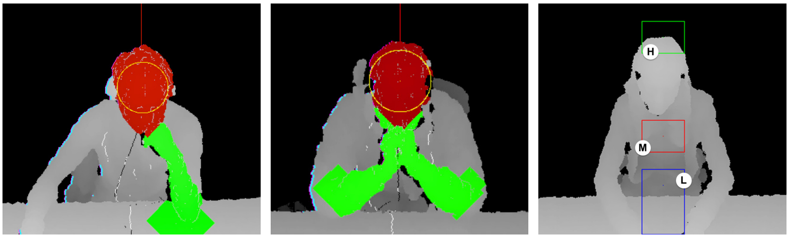


Fig. 2. Tracked gestures (one-hand-to-face, two-hands-to-face) and posture

The posture tracking values were converted into a "postural shift" feature, a discrete representation of *quantity of motion* [14]. Postural shifts were identified through tracked head distances as follows. The median head distance of students at each workstation was selected as the "center" postural position. Distances at one standard deviation (or more) closer or farther than "center" were labeled as "near" or "far," respectively. Postural shifts were labeled when a student moved from one positional category to another (e.g., from "near" to "center"). Both postural shift and gesture events were smoothed by removing those with duration of less than one second. This smoothing mitigated the problem of jitter at decision boundaries (e.g., slight movements at the boundary between "center" and "far" postural positions that cause rapid swapping of both labels). The nonverbal behaviors will hereafter be referenced with the labels ONEHAND, TWOHANDS, and PSHIFT.

3 Tutorial Dialogue and Nonverbal Behavior

Tutorial dialogue and nonverbal behavior have both been extensively examined separately from each other, but there are few investigations of their interactions [20]. We focused on a series of analyses to identify co-dependencies between tutorial dialogue and nonverbal behavior. First, we ran a series of comparisons between overall dialogue act frequencies and dialogue act frequencies conditioned on presence of nonverbal displays. Then, a series of groupwise comparisons identified whether differences existed between students based on gender, prior knowledge, and domain-specific self-efficacy. Statistically significant results are shown in bold.

The first analyses consider the frequency of dialogue acts given that a nonverbal behavior occurred either before or after a dialogue act. An empirically determined fifteen-second interval was used to tabulate occurrence of nonverbal behavior events both before and after dialogue acts. The frequencies were normalized for individuals and averaged across the corpus. Thus, the values shown in the analyses below are average relative frequencies. Dialogue acts with overall average relative frequency below 1% were excluded from the analyses.

The analyses of student dialogue acts consider two situations for each nonverbal behavior. The first examines student dialogue acts given that a nonverbal behavior occurred prior to a dialogue act. This may show how student dialogue moves are affected by the nonverbal behaviors. The second situation considers student dialogue acts given that a nonverbal behavior followed. This represents differences in how a student proceeded following their own dialogue act. In both situations, the nonverbal context may provide insight into the dialogue.

The analyses of student dialogue acts conditioned on prior ONEHAND events revealed a statistically significantly lower frequency of student QUESTIONS following ONEHAND gestures. There was also a trend of more student answers following ONEHAND gestures (Table 2).

Table 2. Analyses of student dialogue acts preceded by ONEHAND gesture

Student Dialogue Act	**Relative Freq. of Stud. Act (*stdev*)**	**Rel. Freq. of Stud. Act with ONEHAND Prior (*stdev*)**	***p*-value (paired *t*-test, two-tailed, *N*=30)**
ANSWER	0.42 (*0.16*)	0.50 (*0.27*)	0.114
ACKNOWLEDGMENT	0.22 (*0.08*)	0.22 (*0.23*)	0.878
QUESTION	**0.14 (*0.09*)**	**0.08 (*0.16*)**	**0.048**
STATEMENT	0.18 (*0.09*)	0.18 (*0.22*)	0.896

The analyses of student dialogue acts followed by PSHIFT events showed a statistically significant lower frequency of student questions followed by PSHIFT (Table 3).

Table 3. Analyses of student dialogue acts followed by PSHIFT postural event

Student Dialogue Act	**Relative Freq. of Stud. Act (*stdev*)**	**Rel. Freq. of Stud. Act Followed by PSHIFT (*stdev*)**	***p*-value (paired *t*-test, two-tailed, *N*=24)**
ANSWER	0.40 (*0.13*)	0.43 (*0.33*)	0.649
ACKNOWLEDGMENT	0.23 (*0.09*)	0.29 (*0.29*)	0.296
QUESTION	**0.15 (*0.09*)**	**0.08 (*0.12*)**	**0.019**
STATEMENT	0.20 (*0.11*)	0.16 (*0.20*)	0.246

The analyses of tutor dialogue acts are conditioned on student nonverbal behaviors present after a tutor move, which may show how students reacted to tutor moves. The analyses of tutor dialogue acts followed by posture identified statistically significant lower frequencies of tutor DIRECTIVEs and tutor POSITIVE FEEDBACK followed by PSHIFT (Table 4). The analyses of tutor dialogue acts followed by TWOHANDS revealed statistically significant lower frequencies of tutor ANSWERs and tutor DIRECTIVEs followed by TWOHANDS (Table 5). Additionally, there was a trend of greater frequency of questions followed by TWOHANDS.

Table 4. Analyses of tutor dialogue acts followed by PSHIFT postural event

Tutor Dialogue Act	**Relative Freq. of Tutor Act (*stdev*)**	**Rel. Freq. of Tutor Act Followed by PSHIFT (*stdev*)**	***p*-value (paired *t*-test, two-tailed, *N*=24)**
ANSWER	0.04 (*0.03*)	0.04 (*0.07*)	0.722
ACKNOWLEDGMENT	0.03 (*0.03*)	0.06 (*0.13*)	0.162
DIRECTIVE	**0.08 (*0.04*)**	**0.05 (*0.06*)**	**0.012**
HINT	0.07 (*0.05*)	0.11 (*0.20*)	0.350
POSITIVE FDBK	**0.18 (*0.05*)**	**0.13 (*0.10*)**	**0.033**
QUESTION	0.21 (*0.07*)	0.26 (*0.24*)	0.359
STATEMENT	0.36 (*0.10*)	0.32 (*0.23*)	0.419

Table 5. Analyses of tutor dialogue acts followed by TWOHANDS gesture

Tutor Dialogue Act	**Relative Freq. of Tutor Act (*stdev*)**	**Rel. Freq. of Tutor Act Followed by TWOHANDS (*stdev*)**	***p*-value (paired *t*-test, two-tailed, *N*=23)**
ANSWER	**0.05 (*0.03*)**	**0.01 (*0.03*)**	**<0.001**
ACKNOWLEDGMENT	0.03 (*0.03*)	0.01 (*0.04*)	0.258
DIRECTIVE	**0.08 (*0.03*)**	**0.03 (*0.05*)**	**<0.001**
HINT	0.06 (*0.05*)	0.04 (*0.11*)	0.382
POSITIVE FDBK	0.18 (*0.05*)	0.21 (*0.18*)	0.524
QUESTION	0.19 (*0.07*)	0.26 (*0.25*)	0.135
STATEMENT	0.39 (*0.09*)	0.39 (*0.30*)	0.977

The primary focus of the above analyses was to investigate the relationships between tutorial dialogue and student nonverbal behaviors. However, the broader nature of nonverbal behavior in tutoring can be explored through analyses conditioned upon student characteristics. For this purpose, three groupwise analyses were conducted to examine gender and domain-specific self-efficacy. First, students were grouped into categories of male (*N*=28) and female (*N*=10). Comparisons of PShift, OneHand, and TwoHands yielded no significant differences (*t*-tests with unequal variance, two-tailed). Second, students were grouped through a median split on pretest score, with high prior knowledge (*N*=19) and low prior knowledge (*N*=19). Comparisons of PShift, OneHand, and TwoHands yielded no significant differences (*t*-tests with unequal variance, two-tailed). Finally, a median split on domain-specific self-efficacy was performed to create groups of high self-efficacy (*N*=19) and low self-efficacy (*N*=19). No differences were found in OneHand or PShift across the groups (*t*-tests with unequal variance, two-tailed). However, students who reported low self-efficacy were found to display more TwoHands gestures (*t*-test with unequal variance, two-tailed). Students in the low self-efficacy group had an average of 0.53 TwoHands displays per minute (*N*=19, *stdev*=0.52), while the high self-efficacy group had an average of 0.20 TwoHands displays per minute (*N*=19, *stdev*=0.34). This result was statistically significant with *p*=0.029.

4 Discussion

The hand-to-face gestures examined here are in a class different from those involved in social conversation and face-to-face tutoring. In face-to-face interaction, social communication guides the nonverbal interaction [16]. Objects in the surrounding environment and spoken concepts form a common substrate that is referenced in conversational gestures. In the case of computer-mediated tutoring, social displays are greatly reduced [15]. Thus, hand-to-face gestures may be more representative of the cognitive-affective states that accompany them compared to communicative or social gestures.

One-hand-to-face gestures are often thought of as embodiments of a thoughtful state.[1] Here, student questions were found to be less frequent following a one-hand-to-face gesture. It may be that students who presented one-hand-to-face gestures had fewer questions to ask. Only fifteen percent of one-hand-to-face gestures occurred before student utterances. Additionally, one-hand-to-face gestures most frequently occurred before student answers. Students are likely to think before providing an answer and in work on task outside of the dialogue. The occurrence of one-hand-to-face gestures coincides with both of these thought-provoking events. Thus, our corpus supports interpretation of one-hand-to-face gestures as a nonverbal behavior with an underlying thoughtful state.

The groupwise self-efficacy analysis presented here showed that students with lower self-efficacy tend to produce more two-hands-to-face gestures. Coupled with a

[1] One such gesture has even been cast in bronze as a timeless exemplar, “The Thinker.”

prior result [15] that found two-hands-to-face gestures to be negatively correlated with focus, a picture emerges of this gesture as an embodiment of reduced focus and lower confidence. Here, tutor answers and tutor directives were less likely to be followed by two-hands-to-face displays. This appears to indicate that students were more focused after these tutor moves. Both tutor answers and directives provide responsive instruction to the student. In the case of answers, the student would have asked a question, and thus would be attentively waiting for the tutor's answer. With directives, the tutor is supplying the student with direct task solution steps that the student must then act upon. The interface did not allow tutors to edit students' computer programming code, so tutor directives imply subsequent student work.

Postural shifts have been linked with disengagement or negative affect. Studies in different contexts agree: whether it is a child playing a game with a robot [14] or a student interacting with a tutoring system [6, 7, 9], postural shifting has repeatedly been shown to co-occur with disengaged or negative cognitive-affective states. Thus, the postural shifts examined in these analyses most likely indicate a disengaged affective state. In this case, we find that less disengagement followed student questions, tutor answers, and tutor positive feedback. Each of these dialogue acts is directly related to collaborative tutorial interaction in which the student is more likely to be engaged. In the case of student questions and tutor answers, the student has posed the question and subsequently received a response. The student clearly plays an active role in this pattern, so it is not surprising that their body reflects this. With tutor positive feedback, the tutor has praised the student for completing a sub-task. The student was actively engaged in the computer programming task, so this result shows that both the student's body and tutor praise reflect the student's engagement.

4.1 Limitations

As noted in [5], there are many variants of hand-to-face and hand-over-face gestures. The hand-to-face gestures tracked here consider contact between hands and the lower face, without more detail as to how the hand is touching the face (e.g., the difference between holding one's chin and leaning on the palm of a hand). Additionally, temporal characteristics of these gestures may be important. An individual may stroke his or her chin, as opposed to resting on a hand. Thus, the present analyses aggregate an array of more specific gestures into categories of one-hand-to-face or two-hands-to-face. Further development efforts are needed to provide tracking algorithms that distinguish between the spatiotemporal subtleties of hand and face [2].

5 Conclusion

Posture and gesture are fundamental components of embodied affect, with ties to cognitive-affective states that may help or hinder learning. Posture and gesture in computer-mediated tutorial dialogue were investigated using automated techniques to track posture and hand-to-face gestures. Annotated dialogue transcripts were analyzed to identify the relationships between student posture, student gesture, and tutor and student dialogue. The results indicate that posture and hand-to-face gestures are significantly associated with student questions, tutor answers, tutor directives and tutor positive feedback. Additionally, two-hands-to-face gestures occurred significantly more frequently among students with low self-efficacy. The results shed light on the cognitive-affective mechanisms that underlie these nonverbal behaviors. Collectively, the findings provide novel insight into the interdependencies among tutorial dialogue, posture, and gesture, revealing a new avenue for automated tracking of embodied affect during learning.

An important emerging trend in intelligent tutoring systems research is that models of nonverbal behaviors are gradually being integrated into runtime diagnostic models. Gesture is a particularly promising modality for informing runtime behavior of tutoring. Gesture and posture constitute key components of a holistic model of nonverbal behavior and embodied affect during learning. Together, they provide a basis for the next generation of affect-informed personalized learning technologies.

Acknowledgments. This work is supported in part by the North Carolina State University Department of Computer Science and the National Science Foundation through Grant DRL-1007962 and the STARS Alliance Grant CNS-1042468. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the participants, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

References

1. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 39–58 (2009)
2. Kleinsmith, A., Bianchi-Berthouze, N.: Affective Body Expression Perception and Recognition: A Survey. IEEE Transactions on Affective Computing (2012)
3. Calvo, R.A., D'Mello, S.K.: Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing 1, 18–37 (2010)
4. El Kaliouby, R., Robinson, P.: The Emotional Hearing Aid: an Assistive Tool for Children with Asperger Syndrome. Universal Access in the Information Society 4, 121–134 (2005)
5. Mahmoud, M., Robinson, P.: Interpreting Hand-Over-Face Gestures. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) ACII 2011, Part II. LNCS, vol. 6975, pp. 248–255. Springer, Heidelberg (2011)
6. Rodrigo, M.M.T., Baker, R.S.J.d.: Comparing Learners' Affect while using an Intelligent Tutor and an Educational Game. Research and Practice in Technology Enhanced Learning 6, 43–66 (2011)
7. Woolf, B.P., Burleson, W., Arroyo, I., Dragon, T., Cooper, D.G., Picard, R.W.: Affect-Aware Tutors: Recognising and Responding to Student Affect. International Journal of Learning Technology 4, 129–164 (2009)
8. Forbes-Riley, K., Litman, D.: Benefits and Challenges of Real-Time Uncertainty Detection and Adaptation in a Spoken Dialogue Computer Tutor. Speech Communication 53, 1115–1136 (2011)

9. D'Mello, S., Dale, R., Graesser, A.: Disequilibrium in the Mind, Disharmony in the Body. Cognition & Emotion 26, 362–374 (2012)
10. Baltrusaitis, T., McDuff, D., Banda, N., Mahmoud, M., El Kaliouby, R., Robinson, P., Picard, R.: Real-Time Inference of Mental States from Facial Expressions and Upper Body Gestures. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, pp. 909–914 (2011)
11. Brawner, K.W., Goldberg, B.S.: Real-Time Monitoring of ECG and GSR Signals during Computer-Based Training. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 72–77. Springer, Heidelberg (2012)
12. Baker, R.S.J.d., Gowda, S.M., Wixon, M., Kalka, J., Wagner, A.Z., Salvi, A., Aleven, V., Kusbit, G.W., Ocumpaugh, J., Rossi, L.: Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra. In: Proceedings of the 5th International Conference on Educational Data Mining, pp. 126–133 (2012)
13. Kapoor, A., Burleson, W., Picard, R.W.: Automatic Prediction of Frustration. International Journal of Human-Computer Studies 65, 724–736 (2007)
14. Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P.W., Paiva, A.: Automatic Analysis of Affective Postures and Body Motion to Detect Engagement with a Game Companion. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, pp. 305–311 (2011)
15. Grafsgaard, J.F., Fulton, R.M., Boyer, K.E., Wiebe, E.N., Lester, J.C.: Multimodal Analysis of the Implicit Affective Channel in Computer-Mediated Textual Communication. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, pp. 145–152 (2012)
16. McNeill, D.: Gesture & Thought. The University of Chicago Press, Chicago (2005)
17. Grafsgaard, J.F., Boyer, K.E., Phillips, R., Lester, J.C.: Modeling Confusion: Facial Expression, Task, and Discourse in Task-Oriented Tutorial Dialogue. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) AIED 2011. LNCS, vol. 6738, pp. 98–105. Springer, Heidelberg (2011)
18. Bandura, A.: Guide for Constructing Self-Efficacy Scales. In: Pajares, F., Urdan, T. (eds.) Self-Efficacy Beliefs of Adolescents, pp. 307–337. Information Age Publishing, Greenwich (2006)
19. Boyer, K.E., Phillips, R., Ingram, A., Ha, E.Y., Wallis, M., Vouk, M., Lester, J.: Characterizing the Effectiveness of Tutorial Dialogue with Hidden Markov Models. In: Aleven, V., Kay, J., Mostow, J. (eds.) ITS 2010, Part I. LNCS, vol. 6094, pp. 55–64. Springer, Heidelberg (2010)
20. Ha, E.Y., Grafsgaard, J.F., Mitchell, C.M., Boyer, K.E., Lester, J.C.: Combining Verbal and Nonverbal Features to Overcome the "Information Gap" in Task-Oriented Dialogue. In: Proceedings of the Thirteenth Annual SIGDIAL Meeting on Discourse and Dialogue, pp. 247–256 (2012)