The Affective Impact of Tutor Questions: Predicting Frustration and Engagement

Alexandria K. Vail, Joseph B. Wiggins, Joseph F. Grafsgaard, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester
Introduction
Overview

Student
Tutor
Tutorial Interface
Question

Student Engagement
Student Frustration

NC STATE UNIVERSITY
Computer Science
Introduction
Related Work

• Recognized need for modeling affect during learning
  (Paquette & Baker & D'Mello; Rowe & Lester, Forbes-Riley & Litman)

• Focus on frustration and engagement
  (Grafsgaard & Lester; Baker; Heffernan)
Introduction

Related Work

- Significance of short-term over long-term nonverbal patterns (Grafsgaard, Calvo, D'Mello)
- Potential critical moments: tutor questions (Piaget, Stein and Levine, Hardy)
Introduction

Research Question

What is the relationship between student engagement and frustration and his or her nonverbal behavior during critical moments of the tutorial interaction?
Multimodal Corpus
Study Overview

**Domain:** Introductory Java programming
- Computer-mediated human-human interactions
- N = 67 student-tutor pairs
  - 67 unique students
  - 5 unique tutors
Multimodal Corpus
Tutorial Interface

JavaTutor

**TASK**

**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 8. Without writing any code, make a plan with your tutor to store the player’s latest choice. (Hint: you will need another variable.)

**JAVA CODE**

```java
3 String namelocation;
4 namelocation = "fantastic!";
5 System.out.println(namelocation);
6 String playersName;
7 Scanner playInput; 8 playInput = new Scanner(System.in);
9 System.out.println("Enter your name here! ");
10 playname = playInput.nextLine();
11 System.out.println("Our hero’s name is: " + playname);
12 System.out.println("You are standing in a field of corn.");
13 System.out.println("You can: 1. Look North, 2. Hit down.");
14 System.out.println("Please enter 1 or 2.");
15 System.out.println("Looking north you see a farmhouse."); System.out.
16 choiceone = playInput.nextInt();
17 if(choiceone == 1) System.out.println("Looking north you see a farmhouse."); System.out.
```

**CHAT**

(00:11:13)

So I’m thinking I should make a choiceone and a choicetwo 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?

(00:12:30)

choiceone would have the options if you had entered 1 for choiceone and choiceetwo would have the options if you had entered 2 for choiceone.

(00:13:03)

Hm, that’s not bad, but you could store the player’s second choice in just one new variable, regardless of what the first choice was, right?

(00:13:16)

Let’s say that teh player chose 1 first.

(00:13:56)

They still either choose 3 or 4 in the second choice.

...
Multimodal Corpus
Multimodal Collection

- Posture & Gesture
- Facial Expression
- Task & Dialogue
- Electrodermal Activity
Multimodal Corpus
Task & Dialogue Traces

Logs including dialogue messages, coding changes, task progress

**Tutor Inference Questions**
- Questions that require reasoning about content knowledge or formulating a plan
  - *e.g.*, *How do you think this problem can be solved?*

**Tutor Evaluative Questions**
- Questions that require the student to evaluate his or her own understanding
  - *e.g.*, *Do you understand so far?*
Multimodal Corpus
Facial Expression Features

• Facial expression recognition by FACET commercial software
• 19 Facial Action Units
• AU Evidence: Measure of confidence in the presence of each expression
Multimodal Corpus
Gesture Features

- Kinect depth camera images
- Detects amount of time one or two hands touching the lower face
- Algorithm based on surface propagation from center of head (Grafsgaard et al., 2012)
Analysis
Overview

Session Time

Tutor
How can we solve this problem?

3-second interval

Multimodal Features

<table>
<thead>
<tr>
<th>Avg AU4 = 1.20</th>
<th>Avg AU7 = -0.21</th>
<th>Avg AU9 = 0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>1H Gesture = 57%</td>
<td>2H Gesture = 3%</td>
<td>Avg Distance = 1.34</td>
</tr>
</tbody>
</table>

SCRs = 5

$X_0$

$X_1$

$X_n$
Analysis
Overview

Session Time

$\begin{bmatrix}
  x_{00} \\
  x_{10} \\
  \vdots \\
  x_{n0}
\end{bmatrix} \\
\begin{bmatrix}
  x_{01} \\
  x_{11} \\
  \vdots \\
  x_{n1}
\end{bmatrix} \\
\begin{bmatrix}
  x_{02} \\
  x_{12} \\
  \vdots \\
  x_{n2}
\end{bmatrix} \\
\cdots \\
\begin{bmatrix}
  x_{0m} \\
  x_{1m} \\
  \vdots \\
  x_{nm}
\end{bmatrix}$

Conditional Features

$\begin{bmatrix}
  \overline{x_0} \\
  \overline{x_1} \\
  \vdots \\
  \overline{x_n}
\end{bmatrix}$

Unconditional Features

$\begin{bmatrix}
  y_0 \\
  y_1 \\
  \vdots \\
  y_n
\end{bmatrix}$
Analysis
Features Summary

1. Average measure of each facial action unit
2. Percentage of time in which one- or two-hands-to-face gesture is present
3. Average student distance from workstation
4. Average difference between highest and lowest posture point (i.e., leaning)
5. Number of skin conductance responses observed
Analysis
Features Summary

1. Average measure of each facial action unit
2. Percentage of time in which one- or two-hands-to-face gesture is present
3. Average student distance from workstation
4. Average difference between highest and lowest posture point (i.e., leaning)
5. Number of skin conductance responses observed
Analysis
Stepwise Regression Modeling Procedure

• Standardized all features and outcomes:
\[ z = \frac{x - \mu}{\sigma} \]

• Optimized leave-one-student-out cross-validation

• Required strict \( p < 0.05 \) cut-off value (after a Bonferroni correction)
Predictive Model
Student Frustration

• Self-reported outcome collected at end of tutorial session
• NASA-TLX workload survey
  – Frustration Level
• Rating from 1 to 100
# Predictive Model

## Student Frustration – Tutor Evaluative Questions

<table>
<thead>
<tr>
<th>Frustration =</th>
<th>$R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.7039 \times \text{AU12 Lip Corner Puller}$</td>
<td>0.0764</td>
<td>0.014</td>
</tr>
<tr>
<td>$-0.6279 \times \text{AU28 Lip Suck}$</td>
<td>0.2471</td>
<td>0.030</td>
</tr>
<tr>
<td>$-0.1635 \times \text{(intercept)}$</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

Leave-One-Out Cross-Validated $R^2 = 0.3235$
Predictive Model
AU12 Lip Corner Puller

- Associated with less frustration after evaluative questions
- Typically related to smiling
Predictive Model
AU28 Lip Suck

• Associated with **less** frustration after evaluative questions

• Sometimes related to fidgeting, perhaps a “self-manipulator” for emotion regulation
Predictive Model
Student Frustration – Tutor Inference Questions

\[
\text{Frustration} = +0.5660 \times \text{AU6 Cheek Raiser} + 0.3635 \times \text{AU20 Lip Stretcher} - 0.0174 \text{ (intercept)}
\]

<table>
<thead>
<tr>
<th></th>
<th>$R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU6 Cheek Raiser</td>
<td>0.2893</td>
<td>0.022</td>
</tr>
<tr>
<td>AU20 Lip Stretcher</td>
<td>0.0499</td>
<td>0.019</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

Leave-One-Out Cross-Validated $R^2 = 0.3392$
Predictive Model

AU6 Cheek Raiser

• Associated with **more** frustration after inference questions
• Related to pain expressions
Predictive Model
AU20 Lip Stretcher

- Associated with more frustration after inference questions
- Previously found to coincide with moments of embarrassment or “awkwardness”
Predictive Model
Student Engagement

• Self-reported through survey at end of tutorial session

• Parts of User Engagement Survey
  – Focused Attention, Felt Involvement, Endurability
  – Excluded: Perceived Usability, Aesthetics, Novelty

• Rating from 1 to 85
# Predictive Model

## Student Engagement – Tutor Evaluative Questions

<table>
<thead>
<tr>
<th>Engagement =</th>
<th>$R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+0.4422 \times \text{One Hand-to-Face Gesture}$</td>
<td>0.1815</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$-0.5989 \times \text{AU10 Upper Lid Raiser}$</td>
<td>0.1831</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$+0.5770 \times \text{AU12 Lip Corner Puller (session-wide)}$</td>
<td>0.2280</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$+0.5097 \times \text{AU26 Jaw Drop}$</td>
<td>0.0514</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>$-0.2941 \times \text{AU2 Outer Brow Raiser (session-wide)}$</td>
<td>0.1923</td>
<td>0.003</td>
</tr>
<tr>
<td>$+0.2467 \times \text{AU5 Upper Lid Raiser (session-wide)}$</td>
<td>0.0295</td>
<td>0.002</td>
</tr>
<tr>
<td>$+0.1792 \times \text{AU24 Lip Pressor}$</td>
<td>0.0566</td>
<td>0.018</td>
</tr>
<tr>
<td>$+0.4100 \text{ (intercept)}$</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

Leave-One-Out Cross-Validated $R^2 = 0.9224$
Predictive Model
One Hand-to-Face Gesture

• Associated with more engagement after evaluative questions
• Potential indicator of thoughtful contemplation
Predictive Model
AU10 Upper Lip Raiser

• Associated with less engagement after evaluative questions
• Component of prototypical disgust
Predictive Model
AU12 Lip Corner Puller

- Associated with **more** engagement in general
- Often related to smiling

Aaron Lee / YouTube
Predictive Model
AU26 Jaw Drop

- Associated with more engagement after evaluative questions
- May be indication of focus
Predictive Model
AU2 Outer Brow Raiser

- Associated with **less** engagement in general
- Component of the “fear brow”, a display of anxiety
Predictive Model

AU5 Upper Lid Raiser

- Associated with more engagement in general
- Previously found to indicate focused attention
Predictive Model
AU24 Lip Pressor

- Associated with more engagement after evaluative questions
- Component of prototypical anger
Discussion

Frustration
• Facial expression predictive
• Predictive power resulting from only short-term reactions

Engagement
• Facial expression and gesture highly predictive
• Mix of short-term and session-wide behaviors
Conclusion

• Multimodal traces can provide insight into student affective outcomes
• Short-term nonverbal behavior during tutoring can be significantly predictive of student frustration and engagement
• Facial expression and hand-to-face gestures at critical moments were highly predictive
Future Work

**Goal:** Understanding nonverbal behavior with respect to affective outcomes

- Student nonverbal behavior after other important tutorial events
- Combinations of multiple event types for increased prediction
- Implications of these observations for adaptive systems
Acknowledgments

Collaborators
Eunyoung Ha
Christopher Mitchell
Alok Baikadi
Aysu Ezen-Can
Joseph Wiggins

Contact
akvail@ncsu.edu

Funding
CNS-1453520
IIS-1409639