

# IDENTIFYING FOREIGN LANGUAGE ANXIETY WHEN USING AN E-LEARNING SYSTEM

Daneih Ismail<sup>1</sup> and Peter Hastings<sup>2</sup>  
*DePaul University, USA*

## ABSTRACT

In addition to the cognitive, developmental, and pedagogical challenges of learning a second language, foreign language learners also face emotional obstacles. Negative emotions such as confusion, frustration, and anxiety interfere with the cognitive channels, which in return, moderates the learning. These emotions also diminish the sense of self-efficacy, producing a disinclination toward beneficial practice. In this research, we are focusing on foreign language anxiety (FLA). Our long-term goal is to recognize FLA and overcome it within the context of an intelligent foreign language learning system. As a first step, we are focusing on the recognition of FLA within a simple e-learning system for English language learners. To identify FLA, we correlated physical measurements (eye fixation, heart rate, and blood pressure) with emotional self-report. We found that the compound predictors of self-report of language difficulty, self-report of system difficulty, and exercise score can identify language anxiety. This finding will allow us to build a sensor-free anxiety detector which we will use to help overcome foreign language anxiety in the future.

## KEYWORDS

Anxiety, Second Language, Foreign Language, Behavior, E-Learning System

## 1. INTRODUCTION

Foreign language anxiety (FLA) refers to the learner's feelings of tension, stress, and nervousness while learning a non-native language (Horwitz, Horwitz and Cope, 1986; MacIntyre and Gardner, 1994). Learners with FLA suffer from both reduced performance and achievement (MacIntyre and Gardner 1994), but an intelligent tutoring system which can take into account the affective state can reduce anxiety and increase motivation and self-esteem (Lin, Chao, and Huang 2015). Previous research has shown that there are multiple ways that anxiety hinders learning (Horwitz, Horwitz, and Cope 1986). FLA reduces self-efficacy, motivation and engagement (Van der Meij, Van der Meij, and Harmsen 2015; N. Wang and Lewis Johnson 2008). It interferes with cognitive channels (Castillejo 2018) and decreases foreign language achievement (Onwuegbuzie, Bailey, and Daley 2000). It increases cognitive load and overwhelms working memory (Chen and Chang 2009). Anxious learners shift their cognition from the learning to the threat (Pappamihel 2002), which obstructs cognitive processes and reduces performance (MacIntyre and Gardner 1994). FLA also interferes with the fluency of speaking the second language (Castillejo 2018; Horwitz, Horwitz, and Cope 1986).

To improve the process of foreign language learning, we need to recognize that anxiety plays a critical role in shaping the learning experience (Horwitz, Horwitz, and Cope 1986). Identifying FLA can alleviate and eventually overcome anxiety, which in return can promote learning (Horwitz, Horwitz, and Cope 1986).

Previous research has described different types of techniques for identifying emotions including self-report (Arevalillo-Herráez et al. 2017; Castillejo 2018; Horwitz, Horwitz, and Cope 1986; Kerns et al. 2014; MacIntyre and Gardner 1994; Monkaresi et al. 2017; Wixon et al. 2014), physical instruments (Fernandes et al. 2018; Gotardi et al. 2018; Kantor et al. 2001; Keil et al. 2018; Krejtz et al. 2018; Mucci et al. 2016; Weeks, Howell, and Goldin 2013; Zhang et al. 2011), and expert observers (Baker et al. 2012; D'Mello and Graesser 2010). The most prominent self-report tool is the Foreign Language Classroom

---

<sup>1</sup> dismail1@mail.depaul.edu

<sup>2</sup> peterh@cdm.depaul.edu

Anxiety Scale (FLCAS) (Horwitz, Horwitz, and Cope 1986), which has been shown to reliably measure the anxiety level in classroom settings (Park 2014; Rodríguez and Abreu 2003). Other researchers used sliders with emojis on each end to measure emotions such as arousal and pleasure (Betella and Verschure 2016).

Previous research using physical measures to detect anxiety has shown varying results. Some researchers used eye trackers for measuring anxiety, finding that pupil dilation was associated with anxiety (Keil et al. 2018). Anxious participants also had a higher number of fixations with shorter duration (Runswick et al. 2017). On the other hand, (Gotardi et al. 2018) found no correlation between anxiety and saccadic eye movement which should be strongly correlated with fixations. Some researchers used heart rate and facial expression to detect engagement, however, they found that heart rate does not provide significant information (Monkaresi et al. 2017). Other researchers (Gotardi et al. 2018; Kantor et al. 2001; Zhang et al. 2011) found a positive correlation between anxiety and heart rate. Some prior work found a significant association between blood pressure and anxiety (Mucci et al. 2016; Zhang et al. 2011), but others (Howell et al. 2007) found no significant relationship. To label students' emotion, some prior research used observations by expert judges (Baker et al. 2012; D'Mello and Graesser 2010) to identify emotions such as boredom and confusion. These methods were used to measure general anxiety and other emotions, but we are focused on FLA in particular.

Through this study, we want to discover how to detect FLA while using an e-learning system. We addressed four research questions:

1. What is the relationship between classroom anxiety and interaction with the system?
2. Are physical measurements useful to measure FLA?
3. What is the relationship between FLA and the difficulty of the exercise?
4. What sensor-free measures are effective for detecting FLA?

Question 1 addresses the relationship between anxiety in the classroom and anxiety while using an e-learning system. When a participant is anxious in one situation, he is likely to be anxious in a similar context (MacIntyre and Gardner 1994). Understanding the interaction between different anxiety-producing situations could help to overcome anxiety (Huang 2018). We hypothesized that students who are anxious in a foreign language classroom would have the same level of anxiety while using an e-learning system. That would result in differences in interaction behavior which would enable an adaptive tutoring system to detect FLA. This would also validate our choice of emotion self-report as our operationalization of FLA, because we compare it to a previously validated measure of anxiety. In the context of Question 2, some research (Kazdin 2000) has claimed that using physiological measurements is more accurate than using self-report because it does not depend on the truthfulness of the participants and that there is a low correlation between emotion self-report and physiological instruments, but we agree with (Nagane 1990) that using a combination of physical measurements and self-report is better than using one tool. Thus, our hypothesis was that using physical measurements alone would not provide an accurate reading of FLA. For Question 3, (Robinson 2007) showed that language complexity is positively correlated with anxiety. Although (Kim and Tracy-Ventura 2011) found no interaction between task complexity and anxiety, in this study, we hypothesized that more difficult exercises would increase FLA, decreasing participants' assessment of their emotional state. In relation to Question 4, previous research used various metrics of interaction to identify emotion, including mouse behaviors (Azcarra and Suarez 2012), time on task (Baker et al. 2012; Wixon et al. 2014), the number of wrong answers (Wixon et al. 2014) and emotion self-report (Petrovica, Anohina-Naumeca, and Ekenel 2016; Wixon et al. 2014). We hypothesized that using self-report of language difficulty, self-report of system difficulty, exercise score, time on task, and the number of mouse clicks could predict FLA.

We built a simple e-learning system which provides practice on vocabulary, grammar, listening, and pronunciation at differing levels of difficulty. We observed the learners while using the system and measured their anxiety level using emotion self-report. To detect anxiety, we correlated multiple channels such as fixations, heart rate, and blood pressure with emotion self-report. According to (Cunha-Perez et al. 2018), adopting affect detection on intelligent tutoring system would increase the motivation and learning. The eventual goal of this research is to build an intelligent tutoring system that can recognize FLA and alter its behavior to help the learner overcome their anxiety.

## 2. METHOD

### 2.1 Participants

The study consisted of 30 non-native English speakers, 77% (N= 23) female and 23% (N= 7) male, between the ages of 18 and 54 years old. There were 40% (N=12) native Arabic speakers, 37% (N=11) native Chinese speakers, 13% (N=4) native Spanish speakers and 10% (N=3) native Thai speakers. Based on the participants' self-report, their education level was 7% high school, 80% bachelor's degree, 13% master's degree. Their English level was 17% beginner, 53% intermediate and 30% advanced.

### 2.2 Procedure

We performed three rounds of pilot tests. Each consisted of 1-2 participants. During the pilot test, we tested the physical measurements and self-reports. After each round, we validated the physical measurements settings and self-reports to assure they provide an accurate reading and avoid errors. We observed the participants while doing English exercises on the e-learning system. The participants spent an average 43:41 minutes per session. All sessions were audio and video recorded, and log sheets were used to note any unusual occurrences. To do the observational study, we first asked the participant to read then sign the informed consent, ensuring that they understood. After that, we introduced the physical measurements and the use of each. Next, we requested the participant to complete the demographic survey and FLCAS. Then they followed the directions on the computer while the experimenter was in the observation room. After the participants finished all the exercises and the self-reports, we gave them a \$20 Amazon gift card in compensation.

### 2.3 E-Learning System

The e-learning system consists of 27 exercises concentrated on three topics; greeting, transportation, and emotions. The system was designed to provide practice in English with a focus on vocabulary, grammar, listening and pronunciation. To provide an effective e-learning system, we applied the multimedia, modality, redundancy, coherence, segmentation, and leveraging examples guidelines based on the Cognitive Theory of Multimedia Learning (Mayer 2005). It was implemented in HTML, JavaScript, and included agents using animated agents from Media Semantics.

### 2.4 Metrics

To measure FLA, we used self-report plus physical monitors. We used two types of self-report in the study, each in the participant's native language (Arabic, Chinese, Spanish, or Thai). The first self-report consisted of demographic information and the FLCAS. The second self-report was collected after each exercise. It consisted of three items: one on system difficulty, one on language difficulty, and one on current emotion. The system difficulty item was *'I felt I knew how to use the interface'* with a five-point Likert scale from strongly agree to strongly disagree. The language difficulty item was *'I felt I knew how to answer the question'* with a five-point Likert scale from strongly agree to strongly disagree. The current emotion question was *'I felt'* with a continuous-valued slider from *'Calm'* to *'Anxious'* with appropriate accompanying emojis.

The physical monitors were the Tobii eye tracker, Polar A370 fitness tracker, and QardioArm blood pressure monitor. The eye tracker was connected to the bottom frame of the screen. It provided us with the number of fixations, the time on task and the number of mouse clicks. The fitness tracker was worn on the left wrist to measure the heart rate. It provided continuous reading for the heart rate for the whole session. We exported the heart rate information from the fitness tracker to a log file and based on the timestamp we paired the heart rate to the exercise. The blood pressure monitor was attached to the left upper arm. We tried to get a blood pressure reading for each of the exercises but were limited in how frequently we could take measurements. On some exercises, we got a couple of measurements while on others we got none based on how long the learner spent per exercise.

## 2.5 Analyzing the Data

We first calculated the FLCAS score, giving the participants an initial classification as anxious or non-anxious. The participants who scored below 90 were non-anxious and those scoring 90 or above were classified as anxious (Guo, Xu, and Liu 2018).

Language difficulty and system difficulty were coded from the participants' Likert scale responses. The emotion self-report was coded from the participants' continuous-valued slider response. To measure the aggregate exercise difficulty, we calculated the percentage of the correct answers across all participants. Then, we classified the aggregate exercise difficulty into three levels (easy, medium, and hard).

While the participants answered the FLCAS, we measured their blood pressure to record their baseline level. To analyze blood pressure, we calculated the level of change between the baseline and the systolic and diastolic pressure reading for each participant on each exercise. The heart rate was calculated as the difference between the maximum and minimum heart rate per exercise to allow for different base heart rates, and to capture changes that might occur from the exercise itself. We figured that using the average, maximum, or minimum heart rate would not provide us with an accurate reading of the change in heart rate, while the difference would give us an indication of the variance in heart rate. The number of fixations, number of mouse clicks, and time on tasks were provided by the eye tracker device. The score for each exercise for each participant was calculated as the percentage of correct answers for that exercise.

## 3. RESULTS

### 3.1 What is the Relationship Between Classroom Anxiety and Interaction with the System?

We did a Mann Whitney U-test comparing the anxious and non-anxious groups as determined by FLCAS with respect to the other measures and found that there was a significant difference between the groups with respect to the emotion self-report, language difficulty self-report, the systolic and diastolic blood pressure relative to baseline. The results of the Mann Whitney U-Test and the median values are shown in Table 1. There was no significant difference between the two groups in the difference of heart rate, score, number of fixation, system difficulty self-report, time on task, nor the number of mouse clicks.

Table 1. Differences between anxious and non-anxious learners

| Variable                           | Anxious Median | Non-anxious Median | Mann-Whitney U | <i>p</i> -value |
|------------------------------------|----------------|--------------------|----------------|-----------------|
| Emotion self-report                | 14             | 2                  | 57404.500      | < 0.001         |
| Self-report language difficulty    | 2              | 1                  | 67585.000      | < 0.001         |
| Systolic relative to the baseline  | -5             | -8                 | 49766.500      | < 0.001         |
| Diastolic relative to the baseline | -2.667         | -5.2               | 51734.000      | < 0.001         |
| Difference in heart rate           | 10             | 10                 | 74523.000      | 0.067           |
| Self-report system difficulty      | 1              | 1                  | 76941.000      | 0.193           |
| Score                              | 100            | 95                 | 43037.500      | 0.554           |
| Time on task                       | 62.61          | 61.17              | 78668.500      | 0.641           |
| Number of mouse click              | 9              | 9                  | 78570.500      | 0.679           |
| Fixations                          | 164            | 161                | 80458.500      | 0.977           |

### 3.2 Are Physical Measurements Useful to Measure FLA?

To understand the relationship between FLA as measured by emotion self-report and the physical measurements, we did a Pearson correlation between emotion self-report, number of fixations, difference in heart rate, and the systolic and diastolic blood pressure relative to baseline. We found a significant weak

positive correlation between the emotion self-report and the number of fixations  $p < 0.001$ ,  $r = 0.171$ . Also, we found a significant weak positive correlation between the emotion self-report and difference in heart rate  $p < 0.001$ ,  $r = 0.166$ . We found a significant weak positive correlation between the emotion self-report and the systolic blood pressure relative to baseline  $p < 0.001$ ,  $r = 0.174$ . We found a significant weak positive correlation between the emotion self-report and the diastolic blood pressure relative to baseline  $p < 0.001$ ,  $r = 0.149$ .

### 3.3 What is the Relationship Between FLA and the Difficulty of the Exercise?

We measured the exercise difficulty in two ways: learner's "self-report of language difficulty", and "aggregate exercise difficulty" which we calculated based on the score of all the participants. To understand the correlation between the self-report difficulty of the language and anxiety level, we did a Spearman correlation. We found a significant moderate positive correlation between the emotion self-report and the language difficulty self-report  $p < 0.001$ ,  $r = 0.582$ .

We also did a Spearman correlation between the aggregate exercise difficulty, and the anxiety level. We found a significant weak positive correlation between the aggregate exercise difficulty and the emotion self-report  $p = 0.036$ ,  $r = 0.086$ .

### 3.4 What Sensor-Free Measures are Effective for detecting FLA?

To answer this question, we did a multiple regression analysis. The independent variables were language difficulty self-report, system difficulty self-report, score, aggregate exercise difficulty, time on task, and the number of mouse clicks. The dependent variable was the emotion self-report. We found that the model's predictions was not improving with the addition of more predictors. So, we compared all the possible models to find the best fit based on the  $R^2$ , adjusted  $R^2$ , and predicted  $R^2$ . We determined that the best predictors were language difficulty self-report, system difficulty self-report and scores. The other variables did not add power to the prediction. The multiple correlation between emotion and the three predictors provided a good fit  $R = 0.554$ . The combination of the three predictors accounts for about 30% of the variation in emotion,  $R^2 = 30.79\%$ ,  $R^2(adj) = 30.45\%$ , and  $R^2(pred) = 29.10\%$ . The regression equation, which indicated that the model was a significant predictor of emotion ( $F(3, 599) = 88.40$ ,  $p < 0.001$ ) is shown in Table 2.

Table 2. Multiple regression coefficients

| Term                            | Coefficient | Standard Error Coefficient | <i>t</i> -value | <i>p</i> -value |
|---------------------------------|-------------|----------------------------|-----------------|-----------------|
| Constant                        | 15.305      | 0.663                      | 23.07           | < 0.001         |
| Self-report language difficulty | 9.659       | 0.781                      | 12.37           | < 0.001         |
| Self-report system difficulty   | 1.772       | 0.769                      | 2.30            | 0.022           |
| Score                           | -0.580      | 0.683                      | -0.85           | 0.396           |

## 4. DISCUSSION

### 4.1 Question 1: The Relationship Between Classroom Anxiety and Interaction with the System

To initially categorize the participants as anxious or non-anxious in classroom settings, we used the FLCAS score as described above. We did a Mann-Whitney U-test to see if there was a difference between the two groups (anxious/non-anxious) with respect to their system interaction and other measurements. We found that there was a significant difference in the emotion self-report, language difficulty self-report, the systolic and diastolic blood pressure relative to baseline. Our hypothesis that students who suffer from FLA in a classroom situation would also be anxious while using the system was supported.

As we expected, there was a difference in emotion self-report between the anxious and non-anxious groups, with anxious learners reporting a higher anxiety level while interacting with the system. This agrees with (MacIntyre and Gardner 1994), who found that when someone is anxious in one situation he is likely to also be anxious in a similar context. Moreover, the median and average result for the emotion self-report was higher for the anxious group. This indicates that emotion self-report is a valid metric for language anxiety in the context of an e-learning system.

Anxious learners' median self-report of language difficulty was, I *'agree'* that *'I felt I know how to answer the questions'*. While the non-anxious learner median was I *'extremely agree'* that they knew how to answer the questions. This means that anxious learners may have felt the questions were somewhat more difficult than the non-anxious learners did. Or it may indicate that the non-anxious learners were more confident about their answers.

Similar to (Mucci et al. 2016; Zhang et al. 2011), the blood pressure was higher for anxious learners than non-anxious learners. Although both median and average reading for both groups were in the normal range according to (Whelton et al. 2018), there was a slight increase in the average blood pressure reading for anxious learners. This may indicate that non-anxious learners were more relaxed during the exercise.

Similar to (Tanielian and English 2014) the relationship between the anxiety level and the score was inconclusive. This could indicate that anxiety is more related to confidence and self-efficacy than to actual competence with the language.

According to (M. Wang 2014), FLA affects self-confidence and self-esteem. Accordingly, we assumed that there would be differences between anxious and non-anxious learners in the time spent on task and the number of mouse clicks. But we did not find significant differences. Perhaps the anxious learners did not spend significantly more time on the exercises because the difficulty of the exercises was not challenging enough. The number of mouse clicks was the same for both groups which reflects clicking the target button only once when needed.

## 4.2 Question 2: Identifying FLA Using Physical Instruments

Our hypothesis of having physical measurements alone would not provide an accurate reading of FLA was supported. The physical measurements had low correlations with the emotion self-report. Our results agreed with (Kazdin 2000) that there is low correlation between emotion self-report and the physical measurements.

Our findings agreed with (Runswick et al. 2017) which showed that anxious learners had more eye fixations than calm learners. The results showed us that the increase in the anxiety level would increase the difference between the highest and lowest heart rate per exercise. So, if the learners were calm before starting the exercise their heart rate would be relatively low. Then through the exercise, if the learners become more anxious their heart rate would increase (Gotardi et al. 2018; Kantor et al. 2001; Zhang et al. 2011). We used the blood pressure reading for each exercise relative to the learner's baseline to standardize the measurements. We found a positive correlation between the blood pressure relative to the baseline and the emotion self-report. This finding agrees with (Mucci et al. 2016) that showed a positive correlation between anxiety and blood pressure. Based on the blood pressure, anxious participants had a higher rate than non-anxious.

Overall, the physical measurements alone would not suffice to identify FLA within the context of an e-learning system because of the very weak correlation with the emotion self-report. Measuring FLA using physical instrument is inconclusive, some type of self-report is needed to help identify FLA.

## 4.3 Question 3: The Relationship Between FLA and the Difficulty of the Exercise

We measured the difficulty of the exercise using the language self-report and the average score across participants. The language difficulty self-report represents the user's impression of the exercise while the aggregate exercise difficulty provides a more objective value.

We found that both measurements for difficulty of the exercise language are positively correlated with the emotion self-report. This supports our hypothesis that language difficulty would affect FLA, and replicates the findings of (Robinson 2007) about the relationship between exercise difficulty and anxiety. Similarly,

(Chen and Chang 2009) showed that increased difficulty would increase anxiety, and (Lin, Chao, and Huang 2015) found that providing an easy exercise could decrease anxiety.

#### **4.4 Question 4: Identifying FLA by the Interaction of the Learner with the System**

To identify FLA, we did a multiple regression analysis by adapting some of the features used by (Wixon et al. 2014) to identify emotion. The dependent variable was the emotion self-report. The independent variables were language difficulty self-report, system difficulty self-report, exercise score, time on task, number of mouse clicks, and aggregate exercise difficulty. We found, however, that this model was not ideal because the power of the prediction did not improve with the addition of more predictors. So we looked for the best model which used a subset of the predictors.

The factors which appeared to be the most effective predictors for FLA were language difficulty self-report, system difficulty self-report, and exercise score. The other predictors were not dominant. Our explanation of this finding is that participants are aware of their emotions and reported their feelings accurately (Petrovica, Anohina-Naumeca, and Ekenel 2016). They felt more anxious when they believed they didn't know how to answer the question and vice versa. We confirmed that there is a relationship between the complexity of the task and anxiety (Robinson 2007). This relationship led to having language difficulty self-report as the most precise FLA predictor.

The second predictor was the self-report of system difficulty. We assumed this predictor would affect language anxiety because the lack of technical knowledge or of how the e-learning system worked would make the learner confused, which in turn could advance to frustration and anxiety. The system difficulty would also reduce the sense of self-efficacy which in turn increases anxiety (Saadé and Kira 2009). These two predictors replicate the value of self-report as a sensor-free emotion detector as shown by (Wixon et al. 2014). According to (MacIntyre and Gardner 1994), students' performance is affected by FLA. Thus, the third predictor was the learner's score on the exercise. Although the score was not shown to the student in the e-learning system, it predicts their anxiety level. Through this analysis, we could conclude that the calm learners would do better than the anxious learners. More relevant to our long-term goals, it shows that we could use learners' performance in the exercises to help determine their level of FLA.

The other predictors were less important than expected. One of the predictors used by (Wixon et al. 2014) was time on task. We replicated that test, assuming that the time on task would affect language anxiety because we thought the longer the learner took to answer a question, the more anxious he would become. However, we found that time on task had little power in predicting FLA.

We thought that using both measures of exercise language difficulty would add power to the language anxiety predictor. However, the aggregate exercise difficulty did not help predict language anxiety. According to (Robinson 2007), task complexity is determined by the cognitive processing needed for the exercise, which varies from person to person. So although the individual language difficulty self-report did improve the power of prediction of FLA, adding the aggregate exercise difficulty did not.

We hypothesized that anxious learners would click the mouse more than the non-anxious learners. We thought that anxious learners would not be sure about their answer, so they would change their answer which affects the number of the mouse clicks. However, this assumption was wrong. There was no significant correlation between the anxiety level and the number of mouse click, and adding the number of mouse clicks did not affect the power of the prediction of the regression model. This means the number of the mouse clicks would not be a useful component of a sensor-free FLA detector.

## **5. CONCLUSION**

FLA is a major obstacle for foreign language learners. This research showed that emotion self-report is a valid metric for detecting FLA in the context of an interactive system. But we believe that constantly asking the learner how they are feeling is not the best way to help them get over their anxiety. So we have attempted to uncover other measures which will allow an interactive system to detect FLA. We found that the strongest predictors for FLA were the self-report of difficulty of the language in an exercise, self-report of system

difficulty, and the exercise score. Finally, we confirmed that physical measurements were not optimal for identifying FLA.

One limitation of the current study is the small range of difficulties of the exercises. We were also unable to evaluate some of the exercises due to technical issues which reduced the total number of available scores and kept us from using the previous exercise score as predictor for FLA.

We need to extend the e-learning system to include a wider range of challenging levels to verify that there are similar responses. Also, we need to determine the best way to intervene when we know the learner is feeling anxious. This research will allow us to build an adaptive tutoring system which includes a sensor-free intelligent FLA detector. We will use the language difficulty self-report, system difficulty self-report, and the score as predictors of the anxiety level. These will trigger the system's anxiety-reduction behaviors, which should help the learners improve the effectiveness of their efforts.

## ACKNOWLEDGEMENT

This project was funded in part by a grant from the College of Computing and Digital Media at DePaul University. Also, we are grateful to Media Semantics for providing us with a free license for the animated agent application.

## REFERENCES

- Arevalillo-Herráez, Miguel, Luis Marco-Giménez, David Arnau, and José A. González-Calero. 2017. "Adding Sensor-Free Intention-Based Affective Support to an Intelligent Tutoring System." *Knowledge-Based Systems* 132: 85–93. <https://doi.org/10.1016/j.knsys.2017.06.024>.
- Azcarraga, Judith, and Merlin Teodosia Suarez. 2012. "Predicting Academic Emotions Based on Brainwaves, Mouse Behaviour and Personality Profile." In , 728–33. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-32695-0\\_64](https://doi.org/10.1007/978-3-642-32695-0_64).
- Baker, Ryan S.J.d., Sujith M Gowda, Michael Wixon, Jessica Kalka, Angela Z Wagner, Aatish Salvi, Vincent Alevan, Gail W Kusbit, Jaclyn Ocumpaugh, and Lisa Rossi. 2012. "Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra." *Proceedings of the Fifth International Conference on Educational Data Mining*, 126–33. <https://doi.org/10.1007/s00192-014-2606-4>.
- Betella, Alberto, and Paul F.M.J. Verschure. 2016. "The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions." *PLoS ONE* 11 (2): 1–11. <https://doi.org/10.1371/journal.pone.0148037>.
- Castillejo, Susana Pérez. 2018. "The Role of Foreign Language Anxiety on L2 Utterance Fluency during a Final Exam." *Language Testing*. <https://doi.org/10.1177/0265532218777783>.
- Chen, I-Jung, and Chi-Cheng Chang. 2009. "Cognitive Load Theory: An Empirical Study of Anxiety and Task Performance in Language Learning." *Electronic Journal of Research in Educational Psychology* 7 (2): 729–46. <https://doi.org/10.25115/ejrep.v7i18.1369>.
- Cunha-Perez, Cristina, Miguel Arevalillo-Herraez, Luis Marco-Gimenez, and David Arnau. 2018. "On Incorporating Affective Support to an Intelligent Tutoring System: An Empirical Study." *Revista Iberoamericana de Tecnologías Del Aprendizaje* 13 (2): 63–69. <https://doi.org/10.1109/RITA.2018.2831760>.
- D'Mello, Sidney K., and Arthur Graesser. 2010. "Multimodal Semi-Automated Affect Detection from Conversational Cues, Gross Body Language, and Facial Features." *User Modeling and User-Adapted Interaction* 20 (2): 147–87. <https://doi.org/10.1007/s11257-010-9074-4>.
- Fernandes, Catarina, Susana Silva, Joana Pires, Alexandra Reis, Antónia Jimenez Ros, Luís Janeiro, Luís Faisca, and Ana Teresa Martins. 2018. "Eye-Tracking Evidence of a Maintenance Bias in Social Anxiety." *Behavioural and Cognitive Psychotherapy* 46 (1): 66–83. <https://doi.org/10.1017/S1352465817000418>.
- Gotardi, Gisele, Paulo Schor, John van der Kamp, Martina Navarro, Dominic Orth, Geert Savelsbergh, Paula F. Polastri, Raoul Oudejans, and Sergio T. Rodrigues. 2018. "The Influence of Anxiety on Visual Entropy of Experienced Drivers." *Proceedings of the 3rd Workshop on Eye Tracking and Visualization - ETVIS '18*, 1–4. <https://doi.org/10.1145/3205929.3205936>.
- Guo, Yan, Jinfen Xu, and Xudong Liu. 2018. "English Language Learners' Use of Self-Regulatory Strategies for Foreign Language Anxiety in China." *System* 76: 49–61. <https://doi.org/10.1016/j.system.2018.05.001>.



- Horwitz, Elaine K., Michael B. Horwitz, and Joann Cope. 1986. "Foreign Language Classroom Anxiety." *The Modern Language Journal* 70 (2): 125–32. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>.
- Howell, Carol C., Marti H. Rice, Myra Carmon, and Roxanne Pickett Hauber. 2007. "The Relationships among Anxiety, Anger, and Blood Pressure in Children." *Applied Nursing Research* 20 (1): 17–23. <https://doi.org/10.1016/j.apnr.2005.10.006>.
- Huang, Heng Tsung Danny. 2018. "Modeling the Relationships between Anxieties and Performance in Second/Foreign Language Speaking Assessment." *Learning and Individual Differences* 63: 44–56. <https://doi.org/10.1016/j.lindif.2018.03.002>.
- Kantor, Ludmila, Norman S. Endler, Ronald J. Heslegrave, and Nancy L. Kocovski. 2001. "Validating Self-Report Measures of State and Trait Anxiety against a Physiological Measure." *Current Psychology* 20 (3): 207–15. <https://doi.org/10.1007/s12144-001-1007-2>.
- Kazdin, Alan E. 2000. *Encyclopedia of Psychology*. Washington, D.C., USA: American Psychological Association.
- Keil, Verena, Robert Hepach, Severin Vierrath, Detlef Caffier, Brunna Tuschen-Caffier, Christoph Klein, and Julian Schmitz. 2018. "Children with Social Anxiety Disorder Show Blunted Pupillary Reactivity and Altered Eye Contact Processing in Response to Emotional Faces: Insights from Pupillometry and Eye Movements." *Journal of Anxiety Disorders* 58: 61–69. <https://doi.org/10.1016/j.janxdis.2018.07.001>.
- Kerns, Caroline E., Douglas S. Mennin, Frank J. Farach, and Christopher C. Nocera. 2014. "Utilizing an Ability-Based Measure to Detect Emotion Regulation Deficits in Generalized Anxiety Disorder." *Journal of Psychopathology and Behavioral Assessment* 36 (1): 115–23. <https://doi.org/10.1007/s10862-013-9372-3>.
- Kim, YouJin, and Nicole Tracy-Ventura. 2011. "Chapter 11. Task Complexity, Language Anxiety, and the Development of the Simple Past." In , 287–306. John Benjamins Publishing Company. <https://doi.org/10.1075/tblt.2.18ch11>.
- Krejtz, Krzysztof, Katarzyna Wisiecka, Izabela Krejtz, Paweł Holas, Michał Olszanowski, and Andrew T. Duchowski. 2018. "Dynamics of Emotional Facial Expression Recognition in Individuals with Social Anxiety." In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications - ETRA '18*, 1–9. New York, New York, USA: ACM Press. <https://doi.org/10.1145/3204493.3204533>.
- Lin, Hao Chiang Koong, Ching Ju Chao, and Tsu Ching Huang. 2015. "From a Perspective on Foreign Language Learning Anxiety to Develop an Affective Tutoring System." *Educational Technology Research and Development* 63 (5): 727–47. <https://doi.org/10.1007/s11423-015-9385-6>.
- MacIntyre, Peter D, and R C Gardner. 1994. "The Subtle Effects of Language Anxiety on Cognitive Processing in the Second." *Language Learning*. Vol. 44. <https://doi.org/10.1111/j.1467-1770.1994.tb01103.x>.
- Mayer, Richard E. 2005. "Cognitive Theory of Multimedia Learning." In *The Cambridge Handbook of Multimedia Learning*, edited by Richard Mayer, 31–48. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.004>.
- Meij, Hans Van der, Jan Van der Meij, and Ruth Harmsen. 2015. "Animated Pedagogical Agents Effects on Enhancing Student Motivation and Learning in a Science Inquiry Learning Environment." *Educational Technology Research and Development* 63 (3): 381–403. <https://doi.org/10.1007/s11423-015-9378-5>.
- Monkaresi, H, N Bosch, R A Calvo, and S K D'Mello. 2017. "Automated Detection of Engagement Using Video-Based Estimation of Facial Expressions and Heart Rate." *IEEE Transactions on Affective Computing* 8 (1): 15–28. <https://doi.org/10.1109/TAFFC.2016.2515084>.
- Mucci, Nicola, Gabriele Giorgi, Stefano De Pasquale Ceratti, Javier Fiz-Pérez, Federico Mucci, and Giulio Arcangeli. 2016. "Anxiety, Stress-Related Factors, and Blood Pressure in Young Adults." *Frontiers in Psychology* 7 (October): 1–10. <https://doi.org/10.3389/fpsyg.2016.01682>.
- Nagane, Mitsuo. 1990. "Development of Psychological and Physiological Sensitivity Indices to Stress Based on State Anxiety and Heart Rate." *Perceptual and Motor Skills* 70 (2): 611–14. <https://doi.org/10.2466/pms.1990.70.2.611>.
- Onwuegbuzie, Anthony J, Phillip Bailey, and Christine E Daley. 2000. "Cognitive, Affective, Personality, and Demographic Predictors of Foreign-Language." *The Journal of Educational Research*. Vol. 94. <https://doi.org/10.1080/00220670009598738>.
- Pappamihel, N Eleni. 2002. "English as a Second Language Students and English Language Anxiety: Issues in the Mainstream Classroom." *Research in the Teaching of English*. Vol. 36. <http://www.jstor.org/stable/40171530>.
- Park, Gi-Pyo. 2014. "Factor Analysis of the Foreign Language Classroom Anxiety Scale in Korean Learners of English as a Foreign Language." *Psychological Reports* 115 (1): 261–75. <https://doi.org/10.2466/28.11.PR0.115c10z2>.
- Petrovica, Sintija, Alla Anohina-Naumeca, and Hazim Kemal Ekenel. 2016. "Emotion Recognition in Affective Tutoring Systems: Collection of Ground-Truth Data." *Procedia Computer Science* 104: 437–44. <https://doi.org/10.1016/j.procs.2017.01.157>.
- Robinson, Peter. 2007. "Task Complexity, Theory of Mind, and Intentional Reasoning: Effects on L2 Speech Production, Interaction, Uptake and Perceptions of Task Difficulty." *IRAL - International Review of Applied Linguistics in Language Teaching* 45 (3). <https://doi.org/10.1515/iral.2007.009>.

- Rodríguez, Maximo, and Orangel Abreu. 2003. "The Stability of General Foreign Language Classroom Anxiety across English and French." *The Modern Language Journal* 87 (3): 365–74. <https://doi.org/10.1111/1540-4781.00195>.
- Runswick, Oliver R, André Roca, · A Mark Williams, Neil E Bezodis, and Jamie S North. 2017. "The Effects of Anxiety and Situation-Specific Context on Perceptual-Motor Skill: A Multi-Level Investigation." *Psychological Research* 82: 708–19. <https://doi.org/10.1007/s00426-017-0856-8>.
- Saadé, Raafat George, and Dennis Kira. 2009. "Computer Anxiety in E-Learning: The Effect of Computer Self-Efficacy." *Journal of Information Technology Education: Research* 8 (1): 177–91. <https://doi.org/10.28945/166>.
- Tanielian, Adam R, and Disney English. 2014. "Foreign Language Anxiety in a New English Program in Thailand." *The International Education Journal: Comparative Perspectives* 13 (1): 60–81. <https://eric.ed.gov/?id=EJ1035913>.
- Wang, Meihua. 2014. "An Empirical Study on Foreign Language Anxiety of Non-English Major Students: Take the Sophomores in Inner Mongolia University of Technology as an Example." *Studies in Literature and Language* 9 (3): 128–35. <http://www.cscanada.net/index.php/sll/article/view/5930>.
- Wang, Ning, and W Lewis Johnson. 2008. "The Politeness Effect in an Intelligent Foreign Language Tutoring System." *LNCs* 5091: 270–80. [https://doi.org/10.1007/978-3-540-69132-7\\_31](https://doi.org/10.1007/978-3-540-69132-7_31).
- Weeks, Justin W., Ashley N. Howell, and Philippe R. Goldin. 2013. "Gaze Avoidance in Social Anxiety Disorder." *Depression and Anxiety* 30 (8): 749–56. <https://doi.org/10.1002/da.22146>.
- Whelton, Paul K., Robert M. Carey, Wilbert S. Aronow, Donald E. Casey, Karen J. Collins, Cheryl Dennison Himmelfarb, Sondra M. DePalma, et al. 2018. "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines." *Hypertension* 71 (6). <https://doi.org/10.1161/HYP.0000000000000065>.
- Wixon, Michael, Ivon Arroyo, Kasia Muldner, Winslow Burleson, Dovan Rai, and Beverly Park Woolf. 2014. "The Opportunities and Limitations of Scaling Up Sensor-Free Affect Detection." *Proceedings of the 7th International Conference on Educational Data Mining*, 145–52. <https://doi.org/10.1110/ps.073398508>.
- Zhang, Zhihong, Hai Su, Qiang Peng, Qing Yang, and Xiaoshu Cheng. 2011. "Exam Anxiety Induces Significant Blood Pressure and Heart Rate Increase in College Students." *Clinical and Experimental Hypertension* 33 (5): 281–86. <https://doi.org/10.3109/10641963.2010.531850>.