Modeling the Relationships between Basic and Achievement Emotions in Computer-Based Learning Environments

Anabil Munshi[1], Shitanshu Mishra[1], Ningyu Zhang[1], Luc Paquette[2], Jaclyn Ocumpaugh[3], Ryan Baker[3], and Gautam Biswas[1]

[1] Vanderbilt University, Nashville, TN, USA
{anabil.munshi, shitanshu.mishra, ningyu.zhang,
gautam.biswas}@vanderbilt.edu
[2] University of Illinois at Urbana-Champaign, Champaign, IL, USA
lpaq@illinois.edu
[3] University of Pennsylvania, Philadelphia, PA, USA
{jlocumpaugh, ryanshaunbaker}@gmail.com

Abstract. Commercial facial affect detection software is typically trained on large databases and achieves high accuracy in detecting basic emotions, but their use in educational settings is unclear. The goal of this research is to determine how basic emotions relate to the achievement emotion states that are more relevant in academic settings. Such relations, if accurate and consistent, may be leveraged to make more effective use of the commercial affect-detection software. For this study, we collected affect data over four days from a classroom study with 65 students using Betty's Brain. Basic emotions obtained from commercial software were aligned to achievement emotions obtained using sensor-free models. Interpretable classifiers enabled the study of relationships between the two types of emotions. Our findings show that certain basic emotions can help infer complex achievement emotions such as confusion, frustration and engaged concentration. This suggests the possibility of using commercial software as a less context-sensitive and more development-friendly alternative to the affect detector models currently used in learning environments.

Keywords: affective modeling, basic emotions, achievement emotions

1   Introduction

Detecting student emotions in computer-based learning environments (CBLEs) is central to the development of pedagogical interventions that respond to students' emotional needs during learning. With the growth of the affective computing field [31] and the inclusion of *emotion* as an important aspect of self-regulated learning (SRL) [1], researchers studying SRL and affect regulation have emphasized the development of affect detection technologies [7] that can be integrated with CBLEs to obtain automated

measures of student emotions during learning. However, the focus of researchers studying affect detection in the learning context has been different from other fields.

A majority of affective computing research outside education has emphasized the detection of the prototypical *basic emotions* that are considered fundamental to human psychology [17, 20]. This has led to the development of affect detectors that capture a person's basic emotions using models trained on data from multimodal sources, such as facial expressions (cf., FACS in [14]), physiological sensors (EEG), and bodily gestures. There are commercially available affect detector models, such as Affectiva [26], that are trained and tested on large datasets (~10,000 labeled facial images - see [33]). These systems predict basic emotions from face video frames with high accuracy, using only facial features as action units (AUs) [25]. Therefore, they can be integrated with any webcam-equipped computer to offer accurate and non-intrusive emotion detection.

However, most affect detection with CBLEs has been limited in tapping into the power of commercial software. In complex learning settings, learners face achievement scenarios that are hypothesized to elicit the experience of *achievement emotion* states (such as confusion, frustration, and engaged concentration), which are more complex than basic emotions (such as joy, anger, fear, and sadness). Achievement emotions reflect how learners cope with cognitive difficulties in different learning situations as they progress towards their achievement goals [30]. Therefore, research needs to focus more on detecting achievement emotions rather than basic emotions.

In recent years, AIED researchers have built classifiers to detect students' achievement emotions from features available during learning, such as learner activities [22] or facial expressions [5]. While these non-intrusive models are usable in real classrooms [28], the models driven by activity-based features are context-sensitive, since the features are based on activities specific to the learning environment. Therefore, these models require considerable effort to develop and the features have to be recomputed (or even re-designed) every time the models are deployed in new learning environments. While facial feature-based models are more generally applicable, current models for detecting achievement emotions using facial features are limited by a lack of large testing datasets across diverse populations. In general, they have a lower accuracy for predicting specific affective states [5].

Therefore, an alternate approach is to develop methods to link achievement and basic emotions so that the commercial software output can be transformed to report academic achievement emotions. After discussing our framework and the Betty's Brain system, this paper presents a first attempt at this approach. We report findings from data on learners' emotions collected from a classroom study.

2   Background

Emotions have been widely studied in psychology. Plutchik [32, p.345] describes *emotion* as "a complex chain of loosely connected events that begins with a stimulus and includes feelings, psychological changes, impulses to action and specific, goal-directed behavior", thereby suggesting that an emotion is not an isolated event but more of a human response to certain actions or situations.

2.1 Basic emotions

Over time, researchers have defined a set of *basic* emotions that deal with universal human situations [14] or have served certain fundamental biological and social functions through human evolution [20], e.g., as the basis for coping strategies and adaptation. Multiple research studies support this concept of basic emotions, e.g., how *sadness* elicits empathy [4] or *fear* elicits protection behaviors [6]. Another view of basic emotions deals with their fundamental role in *'universal human predicaments'* [15, p.46]. These emotions can be distinguished from each other and from other emotions [14]. Ekman's list of seven basic emotion states are: *anger, contempt, disgust, fear, joy, sadness,* and *surprise* [16]. Ekman [14] claimed that there is robust and consistent evidence of distinctive facial expressions for each basic emotion. Currently, commercial software, such as iMotions AffDex, predicts these basic emotions from facial features with high accuracy by using the Emotional Facial Action Coding System (EMFACS) [18] that provides a comprehensive taxonomy for coding facial behaviors. The AffDex SDK uses binary support vector machine classifiers that compute the likelihood values of the seven basic emotions by detecting facial landmarks from video frames, extracting features, classifying facial action units (AUs), and then modeling emotion expressions using EMFACS. While iMotions does not provide public information about the accuracy of their emotion prediction models, they report very high accuracy (with ROC values ranging from 0.75 to 0.96 for AU-classifiers) for the identification of AUs [26].

2.2 Achievement emotions

Achievement emotions are tied directly to students' achievement activities and outcomes in learning situations [30]. Since students' learning activities and outcomes are often judged with respect to achievement standards (e.g.., the COPES SRL model [34]), emotions pertaining to these learning situations may be seen as *achievement emotions*. Individuals experience specific achievement emotions based on their perceived control of the achievement activities and the personally meaningful outcomes (cf., control-value theory [30]). Researchers also constitute these emotions as *cognitive-affective states* [3] due to the relevance of learner cognition to these emotional experiences. Several studies (e.g., [27, 11]) have shown the relation of these emotions to cognitive activities and performance in the learning environment. These achievement emotion states include *boredom, confusion, frustration, engaged concentration,* and *delight*. D'Mello et al. [12] have explored the transitions between emotion states during learning. Affect observation methods such as BROMP [29] have facilitated the observation and coding of these emotions in classrooms. Classifier models trained on BROMP affect labels can capture the probability of occurrence of the emotion states during learning from log data [22] and facial AUs [5]. However, these models may not be as robust as commercial models that detect basic emotions (cf. Section 1). In this paper, we apply methods for basic and achievement emotion detection to collect both types of affect data for students working in Betty's Brain, a learning-by-teaching environment [24].

3 The Betty's Brain Learning Environment

Betty's Brain adopts a learning-by-teaching method to teach complex scientific processes, such as climate change or thermoregulation, to middle school students. Students teach the virtual pedagogical agent Betty by building causal (cause-and-effect) relationships between concepts, and they have access to a set of hyperlinked *science book* pages to learn the science topic. A *causal map* equipped with a palette of editing tools helps them build and annotate their causal map. A *quiz* module offers students the ability to let Betty take a quiz on causal relationships she has been taught. A *mentor agent* named Mr. Davis helps students evaluate the quiz results by comparing their causal model to an *expert model* that is hidden from the student's view. These tools allow the student to constantly refine their maps and engage in learning and understanding of the scientific process as they teach Betty.

The learning environment supports SRL as students engage in cognitive activities and develop strategies to teach Betty a correct causal model. This enables achievement scenarios, which elicit the experience of achievement emotions. Prior research [27] has explored the relationships between students' cognitive and affective experiences in Betty's Brain and emphasized how automated affect detector models can be beneficial for providing students with personalized guidance that respond to their affective-cognitive states during learning.

In the following section, we describe our classroom study and data collection procedures. We analyze and relate two types of emotion data obtained from separate automated affect detector models in Betty's Brain.

4 Methodology

4.1 Study Design and Data Collection

The classroom study involved 65 sixth-grade students in an urban public school in the southeastern USA. The study was conducted over a period of 7 days. Day 1 included a pre-test of domain knowledge and causal reasoning skills. Day 2 familiarized students with the features of Betty's Brain. For the next four days, students built causal models of climate change, and then took a post-test (identical to pre-test) on Day 7.

In addition to pre-post test scores that showed statistically significant learning gains ($p < 0.05, Cohen's\ d = 1$), we collected timestamped logs of students' activities in Betty's Brain over 4 days. Our action-view logging mechanism (based on the cognitive task model in [23]) captured and categorized student activities. Affect detector models (binary classifiers trained on BROMP affect labels aligned to learners' activity sequences - cf., [22]) were used to measure students' achievement emotion probabilities at a 20-second interval based on a sliding window of their cognitive activities within the learning environment. Individual students worked on their own webcam-enabled laptops, and their facial videos were processed *post hoc* using iMotions AffDex [26] to obtain basic emotion likelihoods at a 30Hz frequency. (Our facial videos suffered from occasional data loss when students moved or changed their laptop orientations.)

4.2 Data Analysis

Data Processing Stages. The basic emotion likelihood scores (between 0(absent) to 1(present)) for *joy, anger, surprise, contempt, fear, sadness,* and *disgust* were obtained from AffDex [26] at the frame rate of 30Hz. Separate classifiers detect likelihood values for each facial AU, and emotion likelihoods are calculated by weighting averages of the relevant AU likelihood scores for each basic emotion.

The achievement emotions (*confusion, frustration, engaged concentration, boredom, delight*) were obtained at a 0.05 Hz frequency (i.e., one set of emotion likelihood values every 20sec), as probability scores (between 0 to 1) from the affect detector models (originally validated using BROMP data) integrated with Betty's Brain.

Data synchronization. We aligned the two affect data streams using logged timestamps. Since achievement emotions were available at a coarser time scale *(one set of likelihood values every 20 sec)* than basic emotions *(30 likelihood values per emotion every 1 sec)*, the two data streams were aligned at the coarser granularity, i.e., one set of emotion likelihood values every 20-sec. We extracted the sets of basic emotion likelihoods for 20-sec intervals and picked the set with the highest sum of likelihoods for that interval. (This set represented the most *pronounced* likelihood predictions from the iMotions software for that time interval). Assuming at $time = t\ secs$, the set of likelihood values of the 7 basic emotions is denoted by $L_{B_{time=t}} = [L_{B1}, L_{B2}, \ldots., L_{B7}]_{time=t}$, then the representative set of basic emotion likelihoods for the time interval $\{t, t + 20\} secs$ can be obtained as the set $L_{B_{time=T}}$, where $\sum_{i=1}^{7} L_{Bi_{time=T}} = max_{t \in \{t,t+20\}} [\sum_{i=1}^{7} L_{Bi_t}]$ and $t \leq T \leq t + 20$. The joined likelihood set $\{L_{B_{time=T}}, L_{A_{time=T}}\}$ is the representative set of basic and achievement emotions after the selection and merging of data at the 20-sec time interval. This set was then aligned with the set of achievement emotion likelihoods that the BROMP detector provided for the same interval.

Data filtering. We applied norm-based thresholding to the aggregated data to filter out the instances that had a very low likelihood of detecting a basic or achievement emotion. This was achieved by filtering out data points where the norm of emotion likelihoods for basic or achievement emotions was below the first quartile, i.e. keeping only those instances at which $Norm_{bB} > Q_1(Norm_B)$ and $Norm_{aA} > Q_1(Norm_A)$, where $Norm_B = \sqrt{\sum_{i=1}^{7} L_{Bi}^2}$ and $Norm_A = \sqrt{\sum_{i=1}^{5} L_{Ai}^2}$.

The norm-filtered data contained 5152 of the original 9198 data points. This included 4607 instances where the dominant achievement emotion was *confusion*, 157 instances of dominant *engaged concentration,* 360 instances of dominant *frustration*, 28 instances of *boredom, and* 0 instances of *delight*. (The dominant achievement emotion at each data-point was obtained as $max\ (L_{Ai})$.) Due to the lack of sufficient training instances to model delight or boredom, data instances with dominant delight or boredom were excluded from subsequent analyses. We re-sampled the three other labels to remove class imbalance biases and then proceeded to build binary classifier models (with target class prediction label = TRUE or FALSE for each classifier).

We note here that the distribution of data instances above is dependent on prediction rates of the BROMP-trained affect detector models, and while these models predict the

classes with high AUC ROC, they are not representative of the exact frequency of each affective state occurring in the classroom. These detectors attempt to identify rare situations – using re-sampling to succeed in this goal [22], and may be more biased towards preferring false positives to false negatives, which may lead to over prediction in certain situations. This limitation, likely caused by re-sampling, is addressed by using more sophisticated re-sampling to address the imbalance created.

Specifically, classifier bias due to imbalanced target classes (i.e., a large difference in the proportion of target class labels) in the training data for each binary classifier was handled by (1) under-sampling majority-class cases using random sampling and (2) synthetic oversampling minority-class cases using the SMOTE algorithm [8]. The re-sampled data, containing 7 numeric features (implying likelihood of basic emotions) and one nominal binary target class, was used to train classifier models.

Training Classifier Models. We used 10-fold stratified cross-validation to build binary classifiers for predicting *engaged concentration*, *frustration*, and *confusion*. The classifier models selected for this purpose included Random Forest (RF), Decision Tree (DT), Neural Network (NN), Naïve Bayes (NB), and Logistic Regression (LR). The Naïve Bayes and Logistic Regression models were selected to serve as baselines for establishing model performance. While our model selection criteria considered both interpretability and performance, since our intended research objective with this analysis was to interpret how the basic emotion features predict achievement emotion classes, our selection of classifier models was biased towards interpretable models like logistic regression and decision trees over more complex and less interpretable models (cf., [9]). In practice, complex models (viz., Neural Networks) did not produce very notable differences in predictive accuracy for this data set (see Table 1), likely due to the relatively small sample size. In the next section, we present our findings from five different classifier models and study the best-performing interpretable model to determine relations between basic and achievement emotions in learning.

5 Results and Discussion

5.1 Model Performance

Table 1 lists the performance metrics for five classifier models (Random Forest (RF), Decision Tree (DT), Neural Network (NN), Naïve Bayes (NB), and Logistic Regression (LR)). NB and LR served as baselines for establishing model performance. AUC was used as the primary performance metric, since it provides a better measure of model performance than classification accuracy in a skewed dataset. From Table 1, Random Forest outperformed other models for all three prediction classes.

5.2 Model Interpretation

Random forest was the highest performing algorithm, followed closely by the decision tree with forward pruning (Table 1). The high performance of random forest can be attributed to averaging over multiple generated random trees, thereby achieving a

model with low bias and low variance. Despite its high predictive efficiency, interpreting a random forest model is considerably difficult, especially compared to 'glass-box' approaches like decision tree. In Table 1, we observe that the decision tree model (with Gini-index for feature selection and forward pruning to prevent overfitting the feature space) achieved the second-highest performance for all target classes. Since decision trees provide better interpretability, we choose to study and interpret the decision tree models for each predicted class in greater detail, given that the purpose of this research is to relate the predicted achievement emotions to the basic emotion detectors.

Table 1. Performance metrics (average over classes) for classifier models predicting achievement emotion (AE) classes (class label = TRUE or FALSE) from basic emotion features

AE	Confusion					Frustration					Engaged Concentration				
Classifier	RF	DT	NN	NB	LR	RF	DT	NN	NB	LR	RF	DT	NN	NB	LR
AUC	0.73	0.70	0.60	0.56	0.58	0.76	0.73	0.59	0.58	0.54	0.85	0.83	0.66	0.62	0.58
F1	0.64	0.64	0.53	0.45	0.45	0.69	0.68	0.56	0.56	0.50	0.78	0.76	0.61	0.58	0.57
Accuracy	0.66	0.66	0.59	0.56	0.56	0.69	0.69	0.57	0.57	0.52	0.78	0.76	0.62	0.58	0.58

Figs. 1 and 2 present the visualization of the decision tree models for predicting *confusion* and *frustration*. (The figure for *engaged concentration* is not presented due to space constraints in the paper.). In each figure, the color of a tree node indicates the predicted class labels at that node (*'red' = TRUE, 'blue' = FALSE),* the strength of the colors indicates the predictive power of the model at that node, the width of the edge indicates the proportion of instances classified along a branch with respect to total instances in the training data. The root node at the top is both the most individually predictive and most meaningful for interpretation.

Fig 1 presents the decision tree model to predict confusion. From the first two splits, we find that the two most informative features are $anger$ and $disgust$. Fig. 1 shows a stark contrast between the left and right halves of the tree, right from the first split. We note how the right half of the tree, with higher values for anger and disgust, has more 'red' nodes predictive of $confusion = TRUE$, with a recall value of 87% at the decision node $L_{Anger} \geq 0.6 \,\&L_{Disgust} \geq 0.06$ at $depth = 3$. Moving further down to depth=4 gives more pronounced predictions of confusion=TRUE, where the decision nodes $L_{Anger} \geq 0.6 \,\&L_{Disgust} \geq 0.06 \,\& (L_{Sadness} \geq 0.98 \,||L_{Contempt} \leq 0.71)$ show that higher likelihood of *anger* and *disgust*, together with high *sadness* or low *contempt*, predict *confusion* with a recall value upwards of 98%. When we shift our focus to the left half of the tree, we see that low anger and low disgust are mostly predictive of a lack of confusion. The only low anger-low disgust situation that is predictive of confusion=TRUE is when this lack of anger and disgust is present together with high likelihood of sadness ($L_{Anger} \leq 0.6 \,\& L_{Disgust} \leq 0.06 \,\& L_{Sadness} \geq 0.95$).

We interpret these findings based on the affect literature in learning. *Confusion* has been linked to 'cognitive disequilibrium triggered by contradictions, conflicts, …' [13, p.10]. In an agent-based learning environment (like Betty's Brain), this disequilibrium

could be socio-cognitive [13, p.10], e.g., when the learner disagrees with agent feedback. In this context, the close mapping of *confusion* to higher *anger* likelihood make more sense, especially when we note 'interference with one's activity' is a cognitive antecedent event to anger [17]. The relation of *confusion* to *disgust* may be explained by Plutchik's circumplex model of emotions, which notes disgust as a complementary (contrasting) state to *trust/acceptance* [32]. This again relates back to cognitive disequilibrium, in an achievement scenario that may incorporate conflict due to lack of trust, perhaps in the agent's feedback or the quiz results in Betty's Brain, or a student's disappointment with himself/herself. This suggests that investigating the socio-cognitive processes leading to each basic or achievement emotion in Betty's Brain can help shed more light on finer-grained relations between the two types of emotions during complex learning tasks.
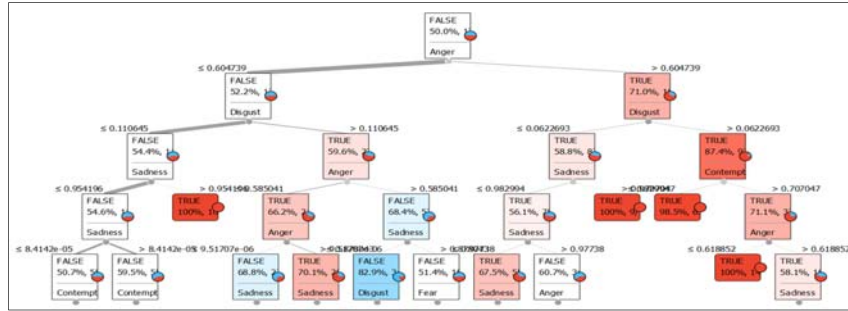

Fig. 1. Decision tree model to predict confusion from basic emotions

Fig. 2 presents the decision tree for predicting frustration. While disgust appears to be the most informative feature at the root node, higher predictive recall (cf., strength of colors in the tree nodes in Fig. 2) is obtained at lower levels of the tree, where disgust is combined with other states such as fear, sadness or contempt. For example, in the right half of the tree (where $L_{Disgust} \geq 0.014$), we find a 92.9% recall for frustration=TRUE at the decision node $L_{Disgust} \geq 0.014$ & $L_{Fear} \leq 1.4 * 10^{-7}$ at $depth = 3$. In the left half of the tree ($L_{Disgust} \leq 0.014$), we find higher recall (=94.5%) for confusion prediction when the low disgust likelihood is combined with high sadness and low contempt (refer to Fig. 2 at depth=4 where $L_{Disgust} \leq 0.014$ & $L_{Sadness} \geq 4.9 * 10^{-5}$ & $L_{Contempt} \leq 0.002$).

From the above, frustration seems to closely map to two complex states: (1) a state of high disgust + low fear, and (2) a state of low disgust + possibility of sadness + low contempt. Affect dynamics models in learning [12] show two affect transition scenarios leading to a state of frustration: (1) *confusion→frustration*, where disequilibrium associated with confusion remains unresolved and leads to failure/blocked goals. Such a transition into frustration may be related to the *[high disgust + low fear]* state noted above, which is associated with low trust/acceptance and high annoyance (cf., [32], where disgust is complementary to trust/acceptance & submissive states like fear are

complementary to aggressive states like annoyance & anger). A prolonged non-acceptance of agent feedback or quiz results due to conflict with expectations may translate into a state of frustration for the learner. (Negative feedback from a tutor has been previously established as possible antecedent of frustration [10]); (2) *boredom→frustration*, where having to endure a learning session despite disengagement may translate into frustration. This state may be related to the *[low disgust + sadness + low contempt]* state we find from the decision tree model, suggesting an affective state that is still negative in valence, but with lower activation than the [high disgust + low fear] state
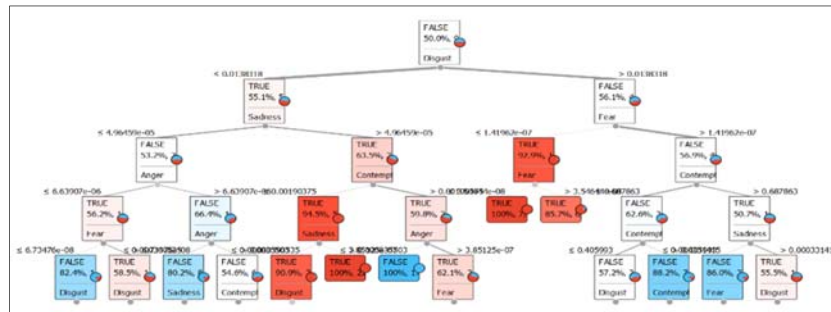

Fig. 2. Decision tree model to predict frustration from basic emotions

Predicting engaged concentration from basic emotion likelihoods. In the decision tree that predicts engaged concentration, *joy* is the most informative feature, and a very high joy likelihood ($L_{Joy} \geq 0.84$) is associated with a prediction of *engaged concentration = FALSE* with a recall of 72.7%. This implies that *engaged concentration* was not the dominant emotion here; so the dominant emotion could be any of the other achievement emotion states, including *delight*, an achievement emotion state whose definition closely matches that of joy but which could not be modeled here due to insufficient data instances. Indeed, *engaged concentration* is often seen as having neutral activation whereas joy and *delight* are often seen as high activation [3].

The second most predictive feature for engaged concentration is *sadness*. A closer analysis of the model shows that, when joy is low, a greater likelihood of *sadness* is a stronger predictor of *engaged concentration = TRUE*. However, a more acceptable predictive recall is obtained at *depth=4*, where a combination of low to medium *joy*, low to medium *sadness* and a possibility of *fear* predict *engaged concentration = TRUE* with a recall value of 70.7%. This situation is obtained by the decision node $L_{Joy} \leq 0.84$ & $L_{Sadness} \in (3.1 * 10^{-6}, 0.78)$ & $L_{Fear} \geq 3.4 * 10^{-3}$.

Engaged concentration being an affect state of neutral activation and mildly positive valence [3] is not associated with high positive valence and high activation emotions like joy or negative valence, low activation emotions like sadness. The association of engaged concentration with lower joy and a possibility of fear may also be related to the fact that engaged concentration associated with the high competence and high challenge scenario of flow (cf., [3, 29]).

6   Conclusion and Future Scope

This research uses interpretable prediction models built from classroom data to suggest links between fundamental *basic emotions* and complex *achievement emotions* during learning in a CBLE. To summarize, we found that the achievement emotion state of *confusion* seems to map most closely to basic emotion states like [high *anger* + high *disgust*], while *frustration* maps closely to states like [*high disgust + low fear*] or [low *disgust + sadness +* low *contempt*], and *engaged concentration* maps closely to low/moderate levels of joy, sadness and a possibility of fear.

While data collection in a classroom setting suggests that our findings have the potential to generalize across natural learning settings, the non-constrained setting adds its own limitations, as discussed in Section 4.1 and in prior research on collecting affect data in real classrooms [5].

Since our research objective was to map basic and achievement emotions, we could use only the subset of our collected data, where both basic and achievement emotion likelihoods were high. Moreover, our emotion logs were likelihood measures that were not direct human-observed emotions but obtained from affect detector models trained on codes obtained from human observations. While these detector models allow for automated detection of affect at scale in a noisy classroom environment, they are likely to be less reliable than human observations. We intend to further validate our findings by replicating our methods with human-coded emotion labels in future classroom studies. Furthermore, since our affect detectors were built off action sequences and performance data in Betty's Brain, it is hard to claim generality for these results.

Despite the data limitations, the research methods and findings reported in this work have implications for shaping future research directions on affect modeling in AIED. First, our approach using interpretable classifiers to model affect in learning accords with prior work [9] that underlines the importance of interpretable ML in AIED. Secondly, this paper presents a scalable and accessible way to identify achievement emotions whose instructional implications have been studied extensively by prior work in the field. Moreover, since commercial software packages for detecting basic emotions are trained on much larger and more varied data than the affect detector models currently used in education, understanding the relations between basic and achievement emotions can help education researchers make use of these commercial software to detect achievement emotions during learning in a computer-enabled classroom.

In future research, beyond addressing limitations noted earlier, we hope to collect affect data from a wider variety of samples and investigate cross-cultural differences in the presentation of affect [19, 21], including achievement emotions [2]. We also intend to conduct further analyses into the cognitive-affective relationships in Betty's Brain, such as how students' socio-cognitive states during agent interactions influence affect.

7　Acknowledgements

References

1. Azevedo, R., Behnagh, R., Duffy, M., Harley, J., and Trevors, G. (2012). Metacognition and self-regulated learning in student-centered leaning environments, Theoretical foundations of student-centered learning 171-197.
2. Baker, R.S., Ogan, A.E., Madaio, M., Walker, E. (2019) Culture in Computer-Based Learning Systems: Challenges and Opportunities. Computer-Based Learning in Context, 1(1), 1-13.
3. Baker, R.S.D., D'Mello, S., Rodrigo, M.M.T., Graesser, A. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68, 4, 223-241.
4. Barnett, M. A., King, L. M., Howard, J. A., & Dino, G. A. (1980). Empathy in young children: Relation to parents' empathy, affection, and emphasis on the feelings of others. Developmental Psychology, 16(3), 243–244.
5. Bosch, N., D'Mello, S., Baker, R.S., Ocumpaugh, J, Shute, V., Ventura, M., Wang, L., Zhao, W. (2016). Detecting Student Emotions in Computer-Enabled Classrooms. IJCAI 2016.
6. Bowlby, J. (1973). Attachment and loss, Vol. 2: Separation. New York: Basic Books.
7. Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary re-view of models, methods, and their applications. IEEE Transactions on affective computing, 1(1), 18-37.
8. Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 16 (2002) 321–357.
9. Conati, C., Porayska-Pomsta, K., and Mavrikis, M. (2018). AI in Education needs interpretable machine learning: Lessons from Open Learner Modelling. 2018 ICML Workshop on Human Interpretability in Machine Learning (WHI 2018), Stockholm, Sweden.
10. D'Mello, S., Craig, S., Sullins, J., and Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue, International Journal of Artificial Intelligence in Education, 16 (1) (2006), pp. 3-28.
11. D'Mello, S., and Graesser, A. (2009). The half-life of cognitive-affective states during complex learning. Cognition and Emotion, Vol 25, Issue 7, pp 1299-1308.
12. D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction Volume 22, Issue 2, April 2012, Pages 145-157.
13. D'Mello, S., Lehman, B., Pekrun, R., and Graesser, A. (2014). Confusion can be beneficial for learning. Learning and Instruction, Vol 29, 153-170.
14. Ekman, P. (1992). Are there basic emotions? Psychological Review, 99(3), 550–553.
15. Ekman, P. (1999). Basic emotions. Handbook of Cognition and Emotion, Ch 3, pp. 45-60.
16. Ekman, P., and Cordaro, D. (2011). What is Meant by Calling Emotions Basic. Emotion Review, Vol 3, Issue 4, 2011.
17. Ekman, P., and Friesen, W.V. (1975). Unmasking the face. Englewood Cliffs, N.J., Prentice-Hall, 1975.
18. Ekman, P. and Friesen, W. (1978). "Facial Action Coding System: A technique for the measurement of facial movements." Consulting Psychologist.

19. Elfenbein, H.A., & Ambady, N. (2003). Cultural Similarity's Consequences A Distance Per-spective on Cross-Cultural Differences in Emotion Recognition. Journal of Cross-cultural Psychology, 34(1), 92-110.

20. Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition rela-tions. Psychological Review, 99(3), 561–565.

21. Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. Proceedings of the National Academy of Sci-ences,109(19), 7241-7244.

22. Jiang, Y., Bosch, N., Baker, R.S., Paquette, L., Ocumpaugh, J., Andres, J.M.A., Moore, A., and Biswas, G. (2018). Expert Feature-Engineering vs. Deep Neural Networks: Which Is Better for Sensor-Free Affect Detection? In Proceedings of AIED 2018, 198-211.

23. Kinnebrew, J., Segedy, J.R. & Biswas, G. (2017). Integrating Model-Driven and Data-Driven Techniques for Analyzing Learning Behaviors in Open-Ended Learning Environ-ments. IEEE Transactions on Learning Technologies, 10(2), 140-153.

24. Leelawong, K., and Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. International Journal of Artificial Intelligence in Education, Vol 18(3), 181–208.

25. Lien, J. J., Kanade, T., Cohn, J. F., & Li, C. C. (1998, April). Automated facial ex-pression recognition based on FACS action units. In Proceedings Third IEEE Interna-tional Confer-ence on Automatic Face and Gesture Recognition (pp. 390-395). IEEE.

26. McDuff, D., Amr, M., Mahmoud, A., Turcot, J., Mavadati, M., and Kaliouby, R. (2016). AFFDEX SDK: A Cross-Platform RealTime Multi-Face Expression Recognition Toolkit. CHI'16 Extended Abstracts, San Jose, CA, USA. ACM.

27. Munshi, A., Rajendran, R., Ocumpaugh, J., Biswas, G., Baker, R. S., & Paquette, L. (2018). Modeling Learners Cognitive and Affective States to Scaffold SRL in Open Ended Learning Environments. *In Proceedings of the International Conference on User Modelling, Adapta-tion and Personalization (UMAP),* Singapore, pp. 131-138.

28. Munshi, A., & Biswas, G. (2019). Personalization in OELEs: Developing A Data-Driven Framework to Model and Scaffold SRL Processes. *In Proceedings of the 20th International Conference on Artificial Intelligence in Education (AIED 2019),* Chicago, pp. 354-358.

29. Ocumpaugh, J., Baker, R.S., and Rodrigo, M.M.T. (2015). Baker Rodrigo Ocumpaugh Mon-itoring Protocol (BROMP) 2.0 technical and training manual. Technical Report. New York, NY: Teachers College, Columbia University. Manila, Philippines: Ateneo Laboratory for the Learning Sciences.

30. Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corol-laries, and implications for educational research and practice. Educational Psychology Re-view, 18 (2006), pp. 315-341.

31. Picard, R.W. (1997). Affective computing. M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321.

32. Plutchik., R. (2001). The nature of emotions. American scientist, Vol 89, Issue 4.

33. Senechal, T., McDuff, D., Kaliouby, R. (2015). Facial Action Unit Detection Using Active Learning and an Efficient Non-Linear Kernel Approximation. In Proc. ICCV IEEE `15.

34. Winne, P.H., and Hadwin, A.F. (2008) The weave of motivation and self-regulated learning. In Motivation and Self-Regulated Learning: Theory, Re-search and Applications, eds D. H. Schunk and B. J. Zimmerman (New York, NY: Lawrence Erlbaum Associates), 297–314.