Capturing Laughter and Smiles under Genuine Amusement vs. Negative Emotion

Cleo Forman[†]
Pomona College
Claremont, California
cafa2016@mymail.pomona.edu

Pablo Thiel[†]
Haverford College
Haverford, Pennsylvania
pthiel@haverford.edu

Raymond Ptucha, Ph.D
Rochester Institute of Technology
Rochester, New York
rwpeec@rit.edu

Miguel Dominguez
Rochester Institute of Technology
Rochester, New York
mad6384@rit.edu

Cecilia O. Alm, Ph.D
Rochester Institute of Technology
Rochester, New York
coagla@rit.edu

***Abstract*—Smiling and laughter are typically associated with amusement. If they occur under negative emotions, systems responding naively may confuse an uncomfortable smile or laugh with an amused state. We present a passive text and video elicitation task and collect spontaneous laughter and smiles in reaction to amusing and negative experiences, using standard, ubiquitous sensors (webcam and microphone), along with participant self-ratings. While we rely on a state-of-the-art smile recognizer, for laughter recognition our transfer learning architecture enhanced on modest data outperforms other models with up to 85% accuracy (F1 = 0.86), suggesting this technique as promising for improving affect models. Subsequently, we analyze and automatically predict laughter as amused vs. negative. However, contrasting with prior findings for acted data, for this spontaneously elicited dataset classifying laughter by emotional valence is not satisfactory.**

***Index Terms*—smiles and laughter under amused vs. negative emotion, spontaneously elicited reactions, affect capture and modeling, human-centered systems**

I. Introduction

Systems need to better understand human behavior and desires in order to provide more considerate user experiences [1]. Currently, commercial affect recognition systems such as FACET and AFFDEX distinguish between anger, contempt, disgust, fear, joy, sadness, and surprise. However, more complex emotional reactions remain less well-studied. For example, smiles are typically associated with amusement, but prior work has noted that people also smile in frustration [2].

With this challenge in mind, this study examines whether laughter and smiles can be elicited under spontaneously-induced amusement and non-amusement, specifically in reaction to negative frustration or discomfort. We investigate three primary research questions:

1) Can laughter and smiles be elicited passively under genuine amusement vs. negative emotions (frustration/discomfort)?
2) Can transfer learning improve laughter detection?
3) What differences, if any, can characterize people's laughter and smiles when they experience genuine amusement vs. negative frustration/discomfort?

TABLE I
Participant emotional self-report form.

Q1: Which face best describes your reaction to the previous item?
 1 2 3 4 5
Q2: Please rate the degree to which (1 = not at all and 5 = extremely) the previous item made you feel: Inspired, Frustrated, Amused, Uncomfortable

Riesberg et al. define frustration as an increase in negative arousal when something uncontrollable impedes the individual's progress toward a goal [3]. While past studies have focused on eliciting frustration by creating games or surveys which deliberately obstruct users' completion of an established goal, we develop a methodology to elicit negative emotions passively (only requiring a browser, a standard webcam, and a microphone for data capture).

To sum up, the contributions include:

- Experimentation with a data collection task engaging observers in viewing text and video content to elicit laughter and smile reactions under genuinely amused vs. negative emotional conditions;
- Demonstration of transfer learning as a viable method to make affect recognition more nuanced when faced with modest data, with laughter detection as a use case;
- Use of participant self-report data to analyze interpretation of laughter and smile reactions when people are genuinely amused vs. negative.

[†]Both authors contributed equally to this research.

II. Select Prior Work

A. Active vs. Passive Emotion Elicitation

Data in affect recognition research often take three forms: acted/posed, induced, and naturalistic [4]. Acted data are acquired by asking individuals to act out a particular affective state; while straightforward to obtain, acted data are the least ecologically valid [2]. Naturalistic data are obtained from real, experienced emotions. These data are reliable but difficult to retrieve [4]. Induced emotional data are produced using a controlled task using active or passive elicitation.

Active elicitation methods ask study participants to perform activities [4], while passive elicitation methods can ask participants to view stimuli. Passive methods tend to include emotional images and emotional film clips, with film clips eliciting stronger emotional reactions than images [4]. Schaefer et al. [5] and Gross et al. [6] presented a database of film clips and films, respectively, that effectively elicited a range of emotions in participants. These resources motivated us to use a visual task. We used videos reported to be highly amusing [7], and selected frustrating and uncomfortable videos to complement the stimuli by consensus between two researchers.

B. Emotion Self-report

Accurately reporting on emotional reactions is another challenge. Third party emotion labeling often does not align with participant self-reported labeling [8]. Often, self-report questionnaires are used to guarantee accurate emotional labels [9]. It is important to record the intensity of emotion and to measure against a baseline neutral affect [10]. Individuals who are high in social desirability may be less willing or capable of reporting negative emotional states [11], [12]. Furthermore, waiting too long before gathering emotional response data leads to systematic biases in respondent recall [13].

C. Automated Smile Detection

While spontaneous smiles are characteristic of amusement, they are also induced by pain or frustration [2], [14]. Prior research has begun to examine the distinguishing characteristics between smiles elicited under different emotional contexts. For example, Hoque et al. found that people smile in 90% of frustrating interactions [2]. McDuff et al. investigated whether features extracted from a time series produced by a smile detector could predict whether a user would want to rewatch an online video [15].

D. Automated Laughter Detection

Previous work has shown that laughter detection can be achieved with high accuracy using deep neural networks. Ryokai et al. trained a 3-layer neural network to detect instance of laughter with 88% per-frame accuracy [16]. Szameitat et al. demonstrated that laughter can be an expression of many emotions beyond just joy [17], additionally finding that laughter from four specific emotions—joy, schadenfreude, tickling, and taunt—can be differentiated based on 12 acoustic parameters, achieving a mean classification accuracy of 84%. In contrast, humans were able to distinguish between these four laughter types only 44% of the time [18]. Furthermore, laughter was always found to be correlated with positive valence [7].

III. Methods

We conducted an experiment on a university campus in the Eastern US with 30 subjects including 12 male, 17 female, and 1 non-binary, aged 18-29. According to our pre-experiment demographic questionnaire, our subjects were 34% White (non-Hispanic), 30% Asian/Pacific Islander, 20% Hispanic or Latino, 13% African-American or Black, and 3% other. Subjects watched a series of 48 videos lasting 18-86 seconds each, with stimuli selected to elicit amusement, frustration, or discomfort. After watching each video, participants filled out a post-clip survey and then completed an affect palate cleanser [7]. The post-clip survey, as seen in Table I, recorded users' valence, inspiration, frustration, amusement, and discomfort on a 5-point scale. We used a condensed and modified short-form version of the Positive and Negative Affect Schedule (PANAS) [19]. We opted to use this condensed instrument over PANAS and SAM [20], in order to reduce the risk of participant fatigue. We recorded data on inspiration to equalize positive and negative survey questions and understand whether we managed to primarily induce the emotions that we were interested in. Each palate cleanser task involved answering a simple multiple choice question about an image (e.g. *How many dinner plates at the table contain spaghetti?*). These quick tasks were intended to bring the subject's emotional state back to neutral. Participants also filled out a pre-experiment survey which collected demographic information.

Videos included intermittent simulated buffering for up to 60% of the video duration in order to explore buffering's impact on eliciting passive frustration. Most videos with buffering were intended to be emotionally neutral to understand the effect of buffering without the influence of strong emotions. Participants were asked whether they suspected that the buffering was simulated.

As a second task, we selected 10 texts that were intended to elicit amusement, frustration, or discomfort. This second task was used to explore whether text stimuli could elicit these emotional experiences. Each text was one to three paragraphs long and was similarly followed by a post-clip survey and a palate cleanser task.

A. Data Collection Experiment

We collected video and audio recordings of participants as they completed the experiment. Video was collected using a Logitech C922x Pro Stream Webcam recording at a resolution of 1080p. Audio was collected using a TASCAM DR-100MKIII recorder with a

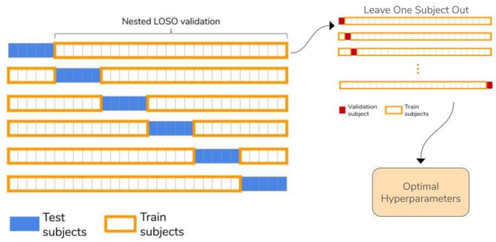

Fig. 1. Model evaluation protocol.

Shure SM31FH-TQG microphone worn by the subject. Participants wore headphones to isolate stimulus audio from participant audio. The experiment was carried out in iMotions [21] which recorded survey and cleanser responses. We also used iMotions to extract facial data from raw participant video. We synced the beginning of the first study task with the audio and repeated for the second task.

B. *Model Evaluation Protocol*

Our evaluation protocol is depicted in Figure 1. To evaluate each modeling task, we split the participants into a training and a test set. The size of the training and test sets varies for each modeling task. These sets serve as the outer folds. In the first outer fold, we conducted a computationally expensive leave-one-subject-out cross-validation on the training set in order to tune hyperparameters for the model. The test set of this first fold is a completely held-out test set. We include accuracy on it as a metric of model accuracy on never-before-encountered data. However, to account for possible bias given modest data size, we also include the model's average accuracy when trained (on the same hyperparameters) and tested on all outer folds as a second metric. We randomly re-select our held-out test set for every training condition we run unless otherwise specified.

C. *Laughter Detection Experiment*

Understanding the less canonical nuances of an affective phenomenon often involves exploring modest data. Therefore, we investigated transfer learning (TL), a technique in which a model's weights trained to solve one problem are used to initialize a solution to a different but related problem or domain, to develop an improved laughter detection model. To determine the merits of this approach, we compared laughter detection models produced under three training conditions. The first condition applied a pre-trained three-layer feedforward neural network trained to detect laughter by Ryokai et al. [16], without any further training. In the second condition, we initialized an identical architecture with random weights and trained a model using training data from our modest dataset. Finally, in the third training condition, we used a TL approach in which we began with Ryokai et al.'s [16] pre-trained weights and fine-tuned them using training data from our audio data. In all three training conditions, we trained and evaluated according to the model evaluation protocol described above.

In order to create a dataset of laughter vs. non-laughter instances from our study, we applied Ryokai et al.'s [16] laughter detection model on our raw audio data with a 20% confidence threshold to identify potential laughter instances in our data. This resulted in 1584 potential laughs. Two researchers then listened to these potential laughs and hand-annotated them as either laughter or non-laughter with an agreement rate of over 96%. Adjudication resolved inter-annotator discrepancies. Mel-Frequency Cepstrum Coefficients (MFCCs) were computed for the laughter and non-laughter instances and used as input data for the modeling task. MFCCs are compact feature representations of short frames (i.e. 10-25ms) of audio that are commonly used in speech signal processing.

Using an open-source implementation by Ryokai et al. [16], we obtain 74 consecutive 10ms frames of audio, each represented by a 39-dimensional vector for a total of 2886 features per data instance.

Twenty one participants laughed in the study, so we split participants into groups of 18 for training and 3 for training, for each of the outer folds described in Section III-B. Our final dataset had approximately 87000 feature vectors, evenly balanced between laughter and non-laughter noise (the non-laugher class was undersampled to provide a balanced dataset).

D. *Valenced Laughter Prediction Experiment*

For our second modeling task, we examine the effectiveness of using Ryokai et al.'s [16] architecture initialized with random weights and trained using only our dataset, as well as a TL approach with MFCCs as features. For this task, only laughter instances from the laughter detection task were used, with each laugh labeled as one of two classes that capture emotional valence. Laughs that received a 3 or above for frustration or discomfort were labeled as negative and the rest were labeled as amused (positive).

Ten participants had instances of negative laughter, so we split training and test data into sets of size 8 and 2 participants, respectively. We proceeded to evaluate the model using the strategy described above in Section III-B. For each participant, we undersampled the laughter class that had more vectors.

IV. Results

A. *RQ1: Can laughter and smiles be elicited passively under genuine amusement vs. negative emotions (frustration/discomfort)?*

1) Self-reported Amused and Negative Emotional Reactions: Active methods have typically been used to

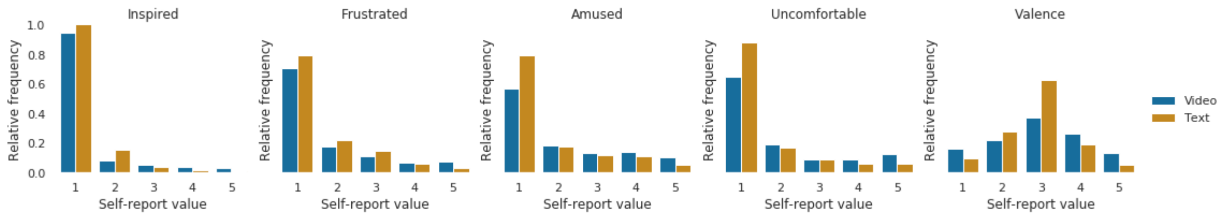

Fig. 2. Comparison of text and video stimuli self-report distributions. Both elicited a range of positive and negative reactions. Non-valence reactions were skewed towards the lower end of the self-report spectrum. Valence appears normally distributed because participants typically remain neutral (n = 1440 for video and n = 300 for text).

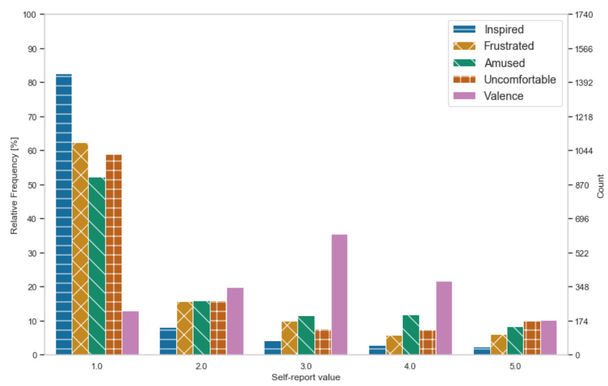

Fig. 3. Aggregated distribution of all participants' self-report ratings on a 5-point scale considering both video and text stimuli (n = 1740), with relative frequency on the left y-axis and frequency on the right y-axis. Stimuli elicit all rating levels for each of the ratings tasks (inspired, frustrated, amused, uncomfortable, and valence), with similar emotional distributions for negative and positive emotions. Valence spans negative to positive and as such is not skewed leftwards.

elicit frustration and discomfort. In contrast, we explore passive methods. The main elicitation strategy was content selection; roughly half of the video stimuli were intended to elicit amusement, and roughly half were intended to elicit frustration or discomfort. Of the videos intended to elicit frustration, some explored adding buffering to relatively emotionally neutral videos.

The results support prior research indicating videos' ability to passively elicit amusement behaviors [7], [22], including smiles and laughter. Using participants' self-report as our assessment standard (Table I) we found that 25% of the video stimuli were rated as 4 or above for amusement, and 39% were rated as 4 or above for valence. In contrast, inspiration (included for rating as a control) scored low, suggesting that we successfully induced the target emotion.

Our results show that the study's stimuli were also effective in passively eliciting frustration and discomfort (see Figure 3). In response to video stimuli, participants self-reported high frustration (responding with 4 or 5 frustration on the post-clip survey) 15% of the time, high discomfort (a 4 or 5 self-report value) 18% of the time, and negative valence (a 1 or 2 valence reaction) 38% of the time.

2) Text Stimuli: Unlike previous studies, we explore whether text stimuli can trigger amusement. We found that 13% of our texts were rated as 4 or above for amusement, 20% were rated 4 or above for valence, whereas only 2% were rated 4 or above for inspiration.

Moreover, in reaction to texts, participants reported high frustration 7% of the time, high discomfort 13% of the time, and negative valence 30% of the time.

Thus, we managed to elicit genuine amusement reactions from both video and text stimuli (Figure 2). Nonetheless, the self-ratings for text stimuli were less intense. This may have been affected by the fact that the text task came after the viewing task.

3) Impact of Buffering: Results show that buffering is not indicative of increased frustration or discomfort. This may be because buffering was primarily added to neutral videos.

4) Presence of Smiles under Negative Emotion: Our results show that smiles were present across all valence ratings, and for highly amusing, highly frustrating, and highly uncomfortable content. Figure 4 contrasts smiles' self-report ratings against valence scores generated by iMotions, equally binned into five levels.

The right panel of Figure 4 shows a clear relationship between iMotions' smile confidence and its outputted valence metric, but this is not supported by our self-report data. Our data show more spread in the confidence of smiles at very high valence (5), with some smiles with confidence above 70 (of 100) at all self-reported valence values. In fact, there is a cluster of smiles with confidence around 75 at a valence of 1. In stark contrast, the iMotions valence metric shows that stimuli viewings with a valence of 4 or 5 as reported by iMotions never have a mean smile confidence below 20 or 50, respectively, and stimuli with an iMotions valence of 1 or 2 never have a mean smile confidence above 30. This suggests that iMotions is unduly relying on smiles as a direct indicator of positive emotion, which does not align with the users' self-reported valence.

To study this further, we observed how iMotions detects smiles with reported confidence levels ranging from 0 to 100 by participants' self-reported rating category. Figure 5 shows that while there is a more substantial spread of high confidence smiles detected for amusing stimuli, there are also plenty of smiles detected with

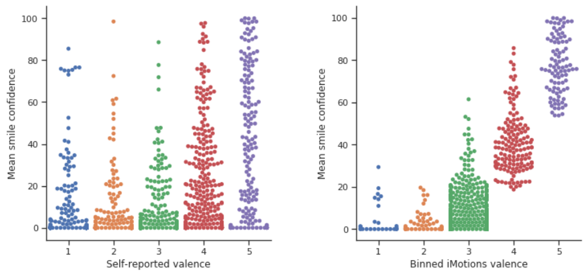

Fig. 4. These graphs compare mean smile confidence output from iMotions to two different valence metrics. The left panel has as its x-axis self-reported valence by each participant for each stimulus, whereas the right panel's x-axis is valence as estimated by iMotions. Each point represents a user engaging with one stimulus (n = 1740). The iMotions software outputs a valence between -100 and 100, so scores are binned equally into five bins for the comparison.

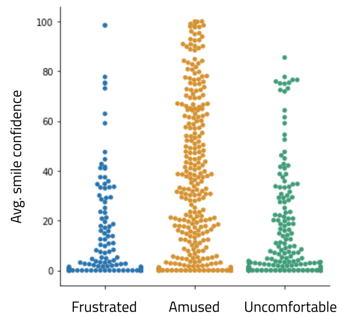

Fig. 5. Average iMotions smile confidence during stimulus viewings self-reported to be highly frustrating, amusing, or uncomfortable (rated 4 or 5) (n = 206 for frustrated, n = 352 for amused, and n = 302 for uncomfortable). Other viewings were below the 4-5 range.

confidence above 50 for highly negative content.

5) Presence of Laughter under Negative Emotion: Laughter occurred when participants self-reported both negative and positive emotions. We identified 207 total occurrences of laughter and extracted these via hand-annotation. Most instances occurred when participants reported high valence and high amusement: 74% of laughter reactions occurred in stimuli where subjects rated their experience high (4 or 5) for amusement, and 82% occurred during similarly highly rated valence reactions. Only 16% occurred when subjects reported high inspiration. In contrast to previous studies, we found participants also laughed during negative emotion: 5% occurred when highly uncomfortable, 6% when feeling highly frustrated, and 6% when reporting a 1 or 2 on the valence rating scale.

B. RQ2: Can transfer learning improve laughter detection?

1) Laughter Detection using Transfer Learning: We computed performance results on two test conditions for laughter detection models trained under the three conditions described in Section III-C.

Table II shows the results (and Table III shows the confusion matrix for one test condition). The pre-trained laughter detection model [16] reached 61% accuracy on

TABLE II
PERFORMANCE ACCURACY FOR THE LAUGHTER DETECTION MODEL TRAINED UNDER THREE TRAINING CONDITIONS.

Laughter Detector	Single Acc.	Rotating Acc.
Pre-Trained	61%	67% ($\sigma = 3.3$)
Newly Trained	48%	70% ($\sigma = 17.1$)
Transfer Learning	83%	85% ($\sigma = 5.6$)

TABLE III
CONFUSION MATRIX FOR SINGLE HELD-OUT TEST SET FOR LAUGHTER DETECTION UNDER TL CONDITION ($n = 17489$). 83% ACCURACY, AND 0.9 RECALL FOR THE LAUGHTER CLASS.

	Pred. Non-laugh.	Pred. Laughter
True Non-laugh.	6737	2066
True Laughter	864	7822

the held-out test set and 67% accuracy on the rotating test set, with a standard deviation of 3.3. Ryokai et al.'s [16] architecture trained only on our data reached 48% accuracy on the held-out test set and 70% accuracy on the rotating test set, however with a large standard deviation of 17.1. We also noted that single test set accuracies varied widely.

Using TL to fine-tune Ryokai et al.'s [16] pre-trained network weights using training data from our experiment yielded 83% accuracy on the held-out test set and 85% accuracy on the rotating test set, with a standard deviation of 5.6. These results indicate that TL is an effective strategy when only modest data is available for affect recognition.

C. RQ3: What differences, if any, can characterize people's laughter and smiles when they experience genuine amusement vs. negative frustration/discomfort?

We treat amusement and frustration/discomfort as two distinct classes, under the assumption that the associated valence is a viable proxy. This decision is supported by our findings that self-reported frustration and discomfort were positively correlated (Spearman's correlation: 0.82) and were each negatively correlated with amusement (Spearman's correlations: -0.81, -0.86). Furthermore, we also found that frustration and discomfort were negatively correlated with valence (Spearman's correlations: -0.58, -0.63), while amusement was positively correlated with valence (Spearman's correlation: 0.79).

Classifying Laughter as Genuinely Amused vs. Negative: As shown in Table IV neither the newly-trained nor the TL conditions effectively classified laughter as elicited under amusement vs. negative emotion (confusion matrices are in Table V). The held-out test sets for this experiment are randomly selected using the

TABLE IV
AMUSED VS. NEG. LAUGHTER CLASSIFICATION.

Classifier	Single Acc.	Rotating Acc.
Newly Trained	43%	54% ($\sigma = 16.2$)
Transfer Learning	67%	59% ($\sigma = 10.1$)

TABLE V

LAUGHTER CATEGORIZATION CONFUSION MATRIX FOR A SINGLE HELD-OUT TEST SET IN NEWLY TRAINED CONDITION AND TL CONDITION ON THE LEFT AND RIGHT OF EACH CELL, RESPECTIVELY ($n = 7536$).

	Pred. Negative	**Pred. Amused**
True Negative	2057 | 4455	3404 | 1006
True Amused	881 | 1451	1194 | 624

same random seed to enable better comparison between training conditions.

V. DISCUSSION AND CONCLUSION

We hypothesized that affect recognition systems, like iMotions, can misconstrue smiles and laughter experienced during negative emotions as signals of positive emotion. To begin to test this, we compared participants' self-reported valence and the iMotions' estimated valence to the smile confidence outputted by iMotions. We found that iMotions indeed assumes a strong positive relationship between smile confidence and valence, whereas we find no clear relationship between smile confidence and participants' actual self-reported valence.

We introduce a method to scalably induce these emotional reactions passively. Short videos and text stimuli, coupled with readily available webcams and microphones, can be used to gather human laughter and smile data under both positive and negative emotions.

Additionally, we demonstrate the utility of TL in improving affect modeling on modest data, showing that using a TL strategy could improve the average efficacy of a pre-trained laughter detection model. This can be leveraged to enhance recognition models of affective behaviors in settings where new, more varied data is difficult to obtain.

We found it a challenge to classify positive vs. negative laughter, suggesting that exploring features capable of distinguishing spontaneous laughs induced under amused and negative emotion remains an open problem. A possible extension would be to explore other features like those considered in Szameitat et al. [17] with our naturally elicited dataset.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Award No. IIS-1851591. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] M. O. Riedl, "Human-centered artificial intelligence and machine learning," *Human Behavior and Emerging Technologies*, vol. 1, no. 1, pp. 33–36, 2019.

[2] M. E. Hoque, D. J. McDuff, and R. W. Picard, "Exploring temporal patterns in classifying frustrated and delighted smiles," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 323–334, 2012.

[3] J. Riseberg, J. Klein, R. Fernandez, R. W. Picard *et al.*, "Frustrating the user on purpose: Using biosignals in a pilot study to detect the user's emotional state."

[4] J. Kory and S. D'Mello, "Affect elicitation for affective computing," in *The Oxford Handbook of Affective Computing*.

[5] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, "Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers," *Cognition and Emotion*, vol. 24, no. 7, pp. 1153–1172, 2010.

[6] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition & Emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[7] M. Saraf, T. Roberts, R. Ptucha, C. Homan, and C. Ovesdotter Alm, "Multimodal anticipated versus actual perceptual reactions," in *21st ACM International Conference on Multimodal Interaction*, October 2019.

[8] M. Muszynski, L. Tian, C. Lai, J. Moore, T. Kostoulas, P. Lombardo, T. Pun, and G. Chanel, "Recognizing induced emotions of movie audiences from multimodal information," *IEEE Transactions on Affective Computing*, 2019.

[9] L. F. Barrett, "Feelings or words? understanding the content in self-report ratings of experienced emotion." *Journal of Personality and Social Psychology*, vol. 87, no. 2, p. 266, 2004.

[10] J. Rottenberg, R. D. Ray, and J. J. Gross, ser. Series in affective science. New York, NY, US: Oxford University Press, 2007, ch. Emotion elicitation using films., pp. 9–28.

[11] R. D. Lane, G. L. Ahern, G. E. Schwartz, and A. W. Kaszniak, "Is Alexithymia the Emotional Equivalent of Blindsight?" *Biological Psychiatry*, vol. 42, no. 9, pp. 834–844, Nov. 1997. [Online]. Available: https://doi.org/10.1016/S0006-3223(97)00050-4

[12] D. L. Paulhus and O. P. John, "Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives." *Journal of Personality*, vol. 66, no. 6, pp. 1025–1060, 1998.

[13] T. R. Mitchell, L. Thompson, E. Peterson, and R. Cronk, "Temporal adjustments in the evaluation of events: The "rosy view"," *Journal of Experimental Social Psychology*, vol. 33, no. 4, pp. 421–448, 1997.

[14] M. Kunz, K. Prkachin, and S. Lautenbacher, "Smiling in pain: Explorations of its social motives," *Pain Research and Treatment*, vol. 2013, p. 8, 2013. [Online]. Available: http://dx.doi.org/10.1155/2013/128093

[15] D. McDuff, R. el Kaliouby, D. Demirdjian, and R. Picard, "Predicting online media effectiveness based on smile responses gathered over the internet," in *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, April 2013, pp. 1–7.

[16] K. Ryokai, E. Durán López, N. Howell, J. Gillick, and D. Bamman, "Capturing, representing, and interacting with laughter," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 358.

[17] D. P. Szameitat, K. Alter, A. J. Szameitat, D. Wildgruber, A. Sterr, and C. J. Darwin, "Acoustic profiles of distinct emotional expressions in laughter," *The Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 354–366, 2009.

[18] C. Brück, B. Kreifelts, and D. Wildgruber, "Emotional voices in context: a neurobiological model of multimodal affective information processing," *Physics of Life Reviews*, vol. 8, no. 4, pp. 383–403, 2011.

[19] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the panas scales." *Journal of Personality and Social Psychology*, vol. 54, no. 6, p. 1063, 1988.

[20] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[21] Affectiva, "imotions," 2019. [Online]. Available: https://imotions.com/

[22] Y. Diaz, C. O. Alm, I. Nwogu, and R. Bailey, "Towards an affective video recommendation system," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2018, pp. 137–142.