Longitudinal Observational Evidence of the Impact of Emotion Regulation Strategies on Affective Expression

Daniel McDuff, *Member, IEEE,* Eunice Jun, Kael Rowan and Mary Czerwinski

Abstract—The ability to regulate our emotions plays an important role in our psychological and physical health. Regulating emotions influences how and when emotions are expressed. We performed a large scale, longitudinal observational study to investigate the effect of emotion regulation ability on expressed affect. We found that expression of negative affect increased throughout the day. For people who suppress emotion this increase is slower that for those who do not. For those with stronger cognitive reappraisal abilities, though not significant, there was a trend for higher positive affect and negative affect increased significantly less steeply, suggesting that they might experience more positive and less negative affect. These results reflect some of the first results based on large scale, continuous tracking of behavioral expression of emotion longitudinally. Our results demonstrate the need to carefully consider the time of day and emotion regulation ability, in addition to gender and age, when attempting to automatically infer affective states for facial behavior.

Index Terms—Emotion regulation, longitudinal, large scale, facial expression, aging, sex, gender.

✦

1 INTRODUCTION

EMOTIONS influence our attention [1], memory [2], [3], and decision-making [4]. Philosophers in ancient Greece recognized the control people can exert on their own emotional experiences [5]. Humans have many means by which they can control, experience, and express emotion. This ability to regulate emotions is an important variable in our psychological, social, and physical health [6]. The inability or difficulty to regulate emotions can have a significant impact on one's quality of life and ability to thrive. Specifically, poor emotion regulation is linked to a range of mood and anxiety disorders [7].

Gross summarizes emotion regulation as "the activation of a goal to up- or down- regulate the magnitude or duration of an emotional response" [5]. Given Gross' arguments around the importance of these strategies in daily life, being able to understand the impact of emotion regulation on one's expressed affect is important for building computer systems that sense, interpret and adapt to human emotion. This type of intelligence is arguably one of the main goals of affective computing [8]. For example, understanding if an individual typically suppresses affect versus not would impact how a machine might interpret their facial expressions. A smile, or expression of anger, from an individual that typically suppresses affect, may need to be given greater weight than one from an individual who does not when trying to computationally infer felt emotion. Similarly, an expression of negative affect from someone who appraises an emotion and then uses cognitive practices to down-regulate it might be interpreted differently from a negative

expression from someone who does not have that ability or practice it as often. The first step in addressing this problem would be to ask whether automated facial analysis can be used to identify distinct patterns in expressed affect in people who use different emotion regulation strategies.

One of the most well-established and commonly employed models of emotion regulation is the *process model* [9]. This model distinguishes between *antecedent-focused* and *response-focused* strategies. Two of these strategies are *cognitive reappraisal* and *emotional suppression*, respectively. Cognitive reappraisal is an antecedent-focused strategy that involves employing a mental intervention before the emotional response tendencies have been fully generated, and thus can lead to a change in the ultimate emotion experienced. Emotion suppression is a later, response-focused strategy that involves modifying the behavioral aspects of the emotional response tendencies [9]. Research has found that interacting with people who are using emotion suppression as a strategy to regulate emotion is more stressful than interacting with someone who is using reappraisal as a strategy [10]. Other research has found that people who habitually use emotion suppression are more likely to be avoided [11]. Both findings could explain why suppressors typically have smaller social networks. Ultimately, those who use cognitive reappraisal over emotional suppression are more likely to be able to form and maintain close personal relationships [12]. We will expand on this further in our literature review below but wanted to raise emotion regulation here as a potential issue when considering health and well being at work and in collaborative teams.

Studies have shown that cognitive reappraisal can be used to successfully elevate expressions of positive affect (PA) and decrease those of negative affect (NA). Whereas, suppression, by its nature, leads to decreases in both PA and NA. Most, if not all, of this evidence comes from studies

---

- D. McDuff, K. Rowan and M. Czerwinski are with Microsoft Research, Redmond, WA, USA.
  E-mail: damcduff@microsoft.com
- E. Jun is with the University of Washington, Seattle, WA, USA.

of individuals in laboratory settings. In many cases measurement of emotion has relied on qualitative, subjective ratings of emotions experienced. While traditional methods for collecting observational emotion data offer researchers a great deal of control over the experimental context, they suffer from limited scalability and real-world ecological validity [13]. Typically in these studies, participants are brought into lab settings where experimenters can record information about their emotional behavior. Inferring the generalization of these observations to real-world contexts is not a small leap. However, sensor- and computer-based tracking of expressed emotion in the real world can enable researchers to quantify the magnitude and duration of an emotional response in a scalable, consistent, and repeatable fashion. Some might argue that doing so in-situ offers increased ecological validity, at very least computer-based tracking in-situ complements more traditional research methods nicely.

Large scale observational studies of affective expression are rare. Recently, however, some studies have leveraged the ability to use automated coding of facial expressions to perform large scale analyses of expressions in mundane contexts [13], [14], [15], [16]. These types of studies allow researchers to corroborate and extend prior findings that have often been performed in laboratory settings or by using more contrived stimuli. For example, using these new techniques, large scale observational evidence of gender differences in facial behavior have been found, with women displaying more positive affect (PA) than men [17]. Cultural differences have also been revealed with smiling more common in individualistic and heterogeneous populations [13], [15]. Other studies have found differences in expressed affect with time of day and day of the week [14].

Longitudinal large scale observational studies of affective expression are even less common. First, continuous tracking of emotions for an extended period of time is onerous for participants. Most longitudinal analysis studies have leveraged and analyzed wearable sensor data (e.g., [18]). However, these do not allow for easy measurement of expressed affect, due to various confounds around the wearable technology (e.g., interpretation of affective states from physiological signals, individual baseline physiological differences, short battery life and forgetting to wear the devices leading to missing data, noise from various physical activities and the lack of contextual information, etc.). As such, to our knowledge, a longitudinal observational study of the impact of emotion regulation strategies on expressed positive and negative affect has not yet been performed. Second, where cameras have been used [14], these systems have not captured the same individuals for extended periods of time, but rather have focused on the measurement of crowds.

Previous studies in psychology and the health sciences have revealed diurnal rhythms in positive affect (PA), negative affect (NA), and emotional states [19] that are common among people despite individual variation [20]. The nature of these patterns has been linked to diurnal cortisol levels, age [19], personality type and general activity [21], life stress [22], and the duration of wakefulness [23]. Yet, how these patterns are impacted by emotion regulation strategies that people actually employ has been largely unexplored.

We designed a study to collect large scale, longitudinal observational data of expressed facial affect in a workplace setting. Using a software logging tool, we captured facial expression metrics of workers from their offices as they went about their daily tasks. Once the software was installed, unobtrusive measurements were automatically captured continuously at one second intervals. This data collection resulted in almost 2,000 hours of facial expression measurements from 117 individuals over the course of 120 person/weeks. From these data we were are able to recover diurnal patterns in positive and negative affect and reveal how these patterns were influenced by emotion regulation strategies.

Based on prior research, we made the following hypotheses about the relationship between emotion regulation strategies and facially expressed behavior. In this context, base rate is defined are the proportion of time a measure is observed against the total observation time.

H1. That higher subjectively reported emotion suppression would result in lower base rates of emotional expression (both PA and NA).

H2a. That higher subjectively reported cognitive reappraisal ability would lead to higher expressions of PA.

H2b. That higher subjectively reported cognitive reappraisal ability would lead to lower expressions of NA.

H3. Men would report using more emotion suppression, and this would be correlated with the reduced expression of both PA and NA.

H4. Older people would report using more cognitive reappraisal and less emotion suppression, and this would be correlated with greater expression of both PA and NA.

H5. That higher reported use of emotion suppression would lead to a slower increase in negative affect across the day.

H6. That higher reported use of cognitive appraisal would lead to a increase in negative affect across the day.

In this paper, we present a first effort towards the automated measurement of longitudinal emotion in situ in the workplace with an eye toward validating specific fundamental emotion regulation theory concepts. As such, we introduce a large scale, longitudinal study using automated facial expression analysis to track expressed affect over time in a workplace setting (see Figure 1 for an example). We show how these data reveal different patterns of affect depending on the emotion regulation strategies that people report employing throughout their working days.

2 RELATED WORK

2.1 Longitudinal Tracking of Emotion

In the past, self-report measurements of emotion were the de facto standard for capturing affect. Experience Sampling Methods (ESM), Ecological Momentary Assessment (ESA) [24], diary keeping [25], and the Day Reconstruction Method (DRM) [26] are all approaches for capturing affective measurements longitudinally. However, due to the laborious nature of self reporting and additional disadvantages, including recall biases, experimenter demand effect and cognitive bias, longitudinal studies of affect have
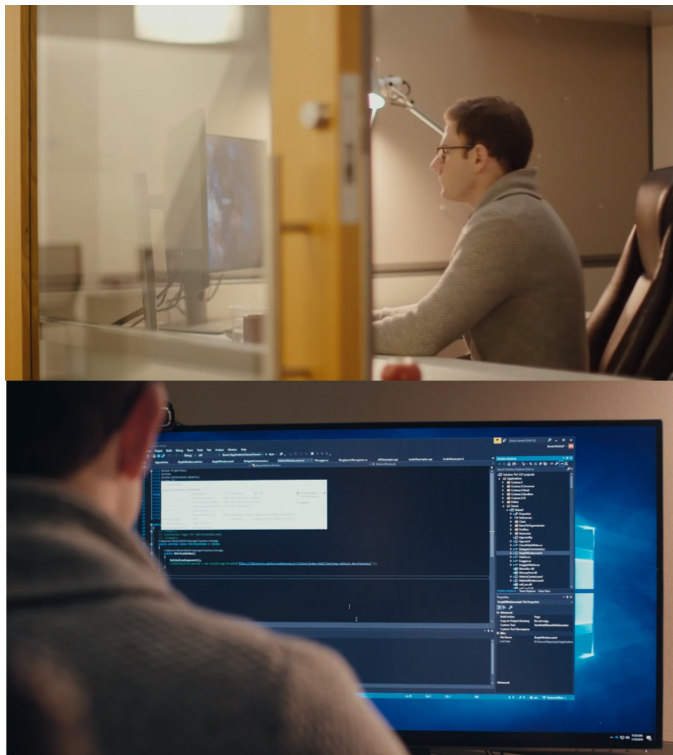
Fig. 1: An example of an office and workstation. While the software was running participants could continue working as normal without interruption. Facial expression metrics were logged using their webcam and a software logging tool.

been limited. The growth of research in the field of affective computing has significantly advanced the state-of-the-art in observational measurement and tracking methods. Researchers are now able to build and deploy powerful software tools, including algorithms for audio, visual, and textual analysis of affect that complement traditional tools, such as ESA and DRM, and offer new possibilities for scientific study.

A leading example of longitudinally tracking affect was the Affective Diary [18]. The diary was designed as an exploratory system that facilitated reflection. The diary enabled users to incorporate bodily and mobile memorabilia in their entries. The system consisted of a mobile phone with a camera to capture images and record mobile usage, sensors in an arm bracelet to collect pedometer (for movement) and electrodermal activity (for arousal) measurements, and a Tablet PC to upload and connect these to a novel user interface for visualizing the data.

Another system for reflection, AffectAura used visual, auditory, and wearable sensor data to track emotional valence and arousal for a period of one week [27]. This system combined multiple sensors (including an Affectiva Q sensor, web cam, microphone and XBOX Kinect) to collect information about the affective state of the user, and PC logging to capture contextual information about what the user was doing. These data sources were combined in order to provide an asynchronous reflection tool about which activities tended to lead toward certain emotional states. Because AffectAura surfaced patterns of affect and behavior,

users found the tool useful for thinking about behavior change.

Focusing on collective rather than personal reflection, Mood Meter [14] was an interactive installation that encouraged, recognized, and counted smiles automatically in public spaces over the course of several weeks. Mood Meter displayed a variety of visualizations, such as a heat map laid over a campus map to display locations with high or low smile counts. Examination periods were correlated with reduced smile count, and graduation with increased smile count. Smiles were most common at weekends and least common on Tuesdays.

De Choudhury and Counts [28] analyzed the sentiment of social media messages exchanged in a workplace. The authors found that the expression of PA decreased dramatically as the day progressed into the evening, suggesting the impact of long work hours on employee affect. The ratio of negative to positive affect also increased. De Choudhury and Counts also found patterns in PA and NA among cross-continental collaborators and between people in different job levels.

All-in-all, there have been a number of attempts to create platforms for longitudinal studies of affect using different types of sensors and algorithms. We build on this tradition and extend prior work by focusing on what affective sensing can inform us about psychological traits and emotional strategies automatically, over time and in situ.

2.2 Emotion Regulation

Researchers have proposed several models of emotion regulation including the process model [9] and the expression-feeling relations model [29].

Perhaps the most widely adopted is the process model of emotion regulation that distinguishes between antecedent-focused (e.g., cognitive reappraisal) and response-focused (e.g., suppression) strategies. Our work was inspired by the work of Gross and John [9] which suggested that there are strong individual differences in cognitive reappraisal and suppression. Reappraisers use strategies early in the emotion generation process in order to modify what they feel and also, therefore, what they express in behavior [9]. Reappraisers negotiate stressful situations with optimism and try to avoid negative feelings by reframing situations [30]. Over the course of several studies and summary articles, Gross and his colleagues have found that reappraisers experienced more positive feelings and therefore felt freer to express these emotions with others deepening their social network and improving their overall well being [31], [32], [33], [34]. On the other hand, Gross et al. found that emotion suppressors used regulation strategies that occurred later in the emotion generation phase and then could only modify what they expressed behaviorally. In other words, suppressors react to strong emotional responses rather than work throughout the emotional experience to mentally prepare and change their thinking. This was found to come at a significant cost to the individual, as they experienced fewer positive as well as negative emotions, tended to share their emotions less frequently socially, and saw themselves as somewhat inauthentic, since they were cognitively aware that they were not expressing socially what they actually felt [31].
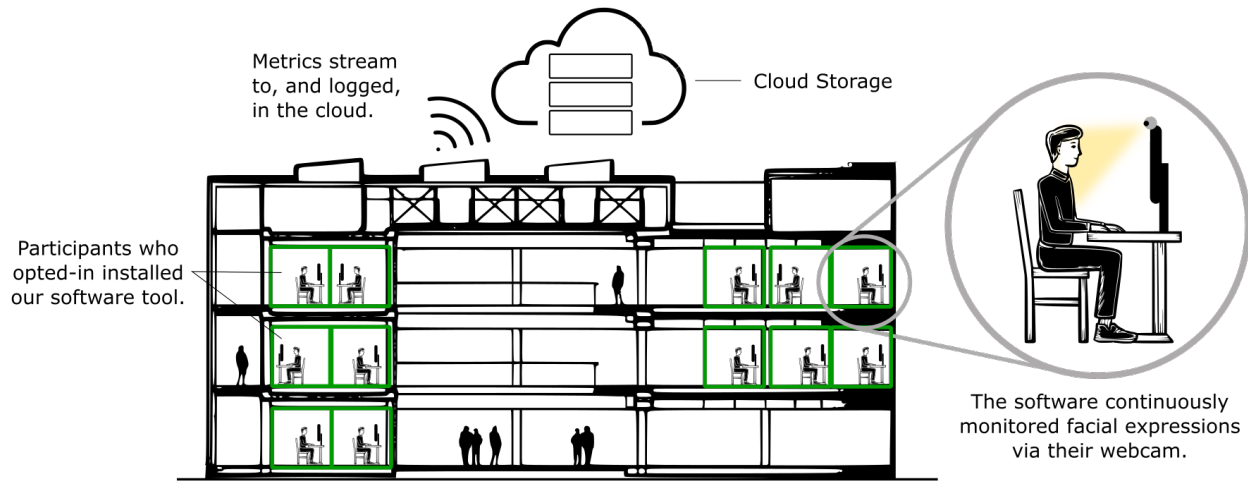
Fig. 2: We deployed software in a workplace setting that continuously monitored the facial expressions of participants that opted-in via their web cameras. The video data were processed locally on the device to preserve the privacy of the participants and the anonymized facial expression metrics were streamed to a secure cloud server.

In practice, people use a variety of emotion regulation strategies in different contexts. Analysis of social context and goals has found that suppression is used more frequently when other people are present [35]. In contrast, cognitive reappraisal is used to help regulate personal feelings, such as making oneself feel better, in the near term. This is another reason why people who use reappraisal may be more likely to have closer, more positive relationships with others [9], [12]; they simply feel more positive and therefore, have a more positive outlook socially.

Rottweiler et al. [36] also found that it is important to not only consider emotion regulation strategies as effective or not, but also to consider the context within which the emotion is experienced, which can have differentiating effects (e.g., before an exam, where suppression actually improved mood in exam-related anxiety). Bonanno and Burton [37] and Hardy [38] relay the importance of context when considering the effectiveness of emotion regulation strategies to help respond to life stressors. Considering situational context can show that changes in emotional response might represent emotional complexity and increased control rather than ability and may lead to better mind and body outcomes. An important dimension of situational context is time.

Emotion regulation has been shown to vary across one's age. Individual differences in abilities and strategies for regulating emotions are learned from childhood and continue to mature long into adulthood. These skills can influence social networks, as well as mental and physical health [39], [40], [41].

It is important that we examine emotion regulation longitudinally in order to understand individual differences and the contexts that influence one's ability to positively cope in a particular situation or not. This study is our first foray into longitudinal examination of emotion regulation while our participants carried out their normal work in situ.

3 EMOTION SENSING PLATFORM

We developed facial expression recognition software to collect data about our participants' expressed affect in situ. The software used standard, computer-attached web cameras and the facial expression recognition software to log facial expression metrics in real-time whenever our participants were present in their office or cubicle (see Figure 1 and Figure 2). The software was designed to capture these signals even when someone was on a video call or using the camera in other applications.

The video signals from the web camera were sampled at 15 frames-per-second (FPS). We used the Microsoft Face API (running locally)[1] to detect the faces in each of the video frames and apply a landmark detector to identify the eyes, nose, and mouth. We logged the number of faces in the frame of the camera and the location of these faces (both bounding box and 15 landmark points). If multiple faces were present in a single frame, the faces are given ID's based on their location. For the purposes of this study we analyzed the expression of the highest face ID within the frame and excluded additional faces. For a large majority of the video frames only one face was detected, and from this we conclude that people were generally alone in their office.

3.1 Facial Expression Recognition

The facial expression recognition engine was a convolutional neural network (CNN) based on the VGG-13 architecture [42]. The classifier was trained to recognize seven expressions: anger, disgust, fear, happiness, sadness, surprise, and contempt and neutral. The first six are the most commonly used categories of expressions [43], with contempt also included in some categorizations [44].

Training. For training, 166,000 images of unique individuals were manually coded. The distribution of men and women is 60% men and 40% women, based on a random sample of 500 images that were manually coded for gender.

---

1. https://azure.microsoft.com/en-us/services/cognitive-services/face/

TABLE 1: Summary statistics. The sample size and mean (and standard deviation) of the emotion regulation scores, hours of face tracking data and base rates of PA and NA.

	Sample (N)	Emotion Regulation		Longitudinal Data (Hrs)		Facial Affect (Base Rate %)	
		Reappraisal	Suppression	Logging	Face Present	PA	NA
Full Set	117	28.4 (5.88)	16.3 (4.91)	178.5	16.6	3.27	0.53
By Gender							
Male	73	28.1 (5.68)	17.3 (4.41)	-	18.4	3.03	0.56
Female	44	28.8 (6.24)	14.8 (5.33)	-	13.2	3.87	0.46
By Age							
35 and Under	75	28.6 (5.99)	16.9 (4.70)	-	17.4	3.15	0.46
Over 35	42	27.9 (5.72)	15.4 (5.16)	-	14.9	3.53	0.69

Human judges labeled the images for the presence of the basic emotions. Ten judges labeled each image and the majority label was taken as the ground-truth for each image. The data were partitioned into independent training and testing sets. A Convolutional Neural Network (CNN) model was trained using the VGG13 architecture [45]. This architecture has 10 convolutional layers and two densely connected layers, for more details see [42]. The input images were scaled to 64×64 pixels. The expression probabilities detected every second were stored on a cloud server. The software also captured context about the participants' work, including the applications they used, calendar appointments, and email communications.

Validation. We validated the facial expression recognition engine on two independent public benchmark data sets, the CK+ dataset [46] and the FER data set testing set [47] (using the FER+ labels [42]). These data sets are comprised of 326 and 3,573 labeled images of the same basic emotion categories as our classifier, respectively. To characterize performance we report the true positive rate, false positive rate and accuracy for the task of categorizing the facial expression images. For CK+ the accuracy was 81.0%, the false positive rate was 3.04% and the true positive rate was 72.5%. For FER the accuracy was 82.0%, the false positive rate was 2.94% and the true positive rate was 76.1%. These results reflect that the model is able to detect expressions and gives relatively consistent performance across different data sets. To further simplify the task of facial affect recognition we averaged the classifier outputs for negative valence expressions (anger, disgust, fear, contempt) to form the NA metric and positive valence expressions (happiness) to form the PA metric. We excluded sadness as we found this was commonly triggered while participants were in focused states (such as concentration) and therefore did not reliably reflect NA.

While facial expression recognition systems still face many challenges, longitudinal tracking enables us to reveal signals even with noisy observations as noise signals are less structured, and more likely to be random, than the signals of interest. Furthermore, by only relying on PA and NA signals, we simplify the recognition task. In the limitations section we discuss some of the challenges of using automated facial expression recognition in this way and also highlight how observational tracking of this kind would otherwise have been impossible (manual coding would be unfeasible).

4 THE STUDY

Collecting longitudinal, observational measurements of affective expressions via cameras is challenging, in part, because it requires participants to be within view of a camera for large parts of the day. In the workplace, information workers are frequently in a fixed location (at their desk) for long periods of time on a regular basis. We designed a study that was conducted on workplace employees at a large multinational technical company. Participants were only recruited in one locale in order to minimize the variability introduced by the effects of weather and environment on the participants' affective states. One hundred and seventeen people (44 females) were recruited for the study via email. Participants who volunteered for the study elected to complete an introductory survey and then have our software tool installed on their computer. If they did not already have a web camera attached to their workstation, they were provided with a camera. The participants were given a link to the software. During the installation, the participants were presented with a consent form and an introductory survey. Upon completing the installation and surveys, the participants could continue with their work as normal, making it very simple for people to participate. The experimental protocol was approved by the institutional ethics review board of Microsoft Research.

The average age was 34 years (*SD* = 9.4). The job roles of the individuals included data scientist/researcher (39%), software developer (33%), program/product manager (7%), administrative assistant (8%), designer (1%), and other (13%). The participants completed the "Big-Five" personality profile [48], the Positive and Negative Affect Scale (PANAS) [49], and the Emotion Regulation Questionnaire [9].

On average, participants had the software installed on their desktops or laptops for 7.5 days. Figure 3 shows the number of hours of facial expression data collected during the day. In our analyses we focus on observations between the periods of 7 AM and 7 PM only as data outside these hours were sparse. Meetings, breaks and other activities caused the participants to be away from their machines for parts of the working day. Between the times of 7 AM and 7 PM, participants' faces were tracked on average 18.6% of the time. In total, we collected over 1,900 hours of facial expression data.

Table 1 shows a summary of emotion regulation scores, PA and NA base rates and number of hours of data logged
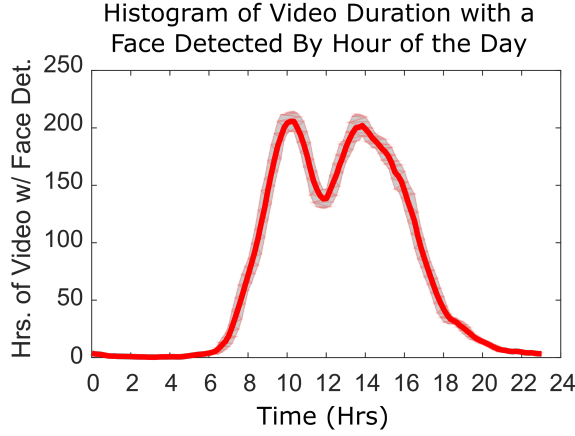
Fig. 3: Histogram showing the number of hours of facial data analyzed per hour of the day for the 117 participants during the course of the study. For our analyses in this paper we only use data between 7AM and 7PM. The band around the line represents the standard error.

TABLE 2: Spearman correlations of the PA and NA base rates, participants' CR and ES scores, age, and time of day (hours since 7AM). (. p < .08, * p < .05, ** p < .01, *** p < .001)

	PA	NA	CR	ES	Age	Time
PA	-	0.086 ***	0.054 **	-0.111 ***	-.036 *	.003
NA		-	.055 **	-.002	.335 ***	.071 ***
CR			-	.035	-0.053	-0.032 .
ES				-	-0.065	0.038 *
Age					-	0
Time						-

per participant. We have broken these measures down by gender and age (35 years old and under, and over 35).

5 RESULTS

To analyze the impact of emotion regulation strategies (i.e., emotion suppression and cognitive reappraisal), gender and age on positive and negative affect base rates, we used a linear mixed effects model:

$$log(BASERATE_i) = \beta_0 + \beta_1(ES_i) + \beta_2(CR_i) +$$
$$\beta_3(Gender_i) + \beta_4(Age_i) + \beta_5(t)$$
$$\beta_6(ES_i * t) + \beta_7(CR_i * t) +$$
$$Z_1(PID_i) + r_i \quad (1)$$

The base rate had to be log transformed as it was log-normally distributed. Here, β_0 is an intercept, β_1 and β_2 are the parameters that estimate the marginal linear effects of the emotion suppression score and cognitive reappraisal score, respectively, on the affect base rate. β_3 and β_4 are the parameters that estimate the marginal linear effect of gender and age on the affect base rate. Gender was a categorical variable and age an original discrete variable (years). β_5 is the parameter estimate for time, the number of hours since 7am. (t) is an ordinal discrete variable (hours since 7am). Z_1 is a parameter describing the variance in the affect base rate that can be explained by the differences among

participants, other than emotion regulation, age, gender and time. Modeling other individual variability as random effects means we do not want to quantify the specific effect of any one participant but rather to account for the overall variability they exert on positive or negative affect. r_i is an error term.

We created two models–one for PA and one for NA. We model PA and NA independently as previous work has shown that these constructs, driven by different bio-behavioral systems, are not two ends of the same continuum [50].

A linear mixed effects model was the most appropriate given that each participant provided multiple hours of data each day over the course of the study. In our models, we aggregate the hours of the day over days of the week for each participant because prior work found similar emotion patterns across the days of the week [51]. In other words, data for a particular hour from a participant is the average of data from that hour over the days the participant contributed data. Because emotion suppression and cognitive reappraisal are two distinct emotion regulation strategies ($\rho = .096$) that are measured as separate constructs in the Emotion Regulation Questionnaire, we include both as independent factors in the linear mixed model. The predicted base rates are log transforms of the raw data observed.

To select the best models for our analyses, we conducted hierarchical regression for positive and negative affect expression. Hierarchical regression focuses on model comparison and is distinct from hierarchical linear modeling, which is a statistical test accounting for nested data (e.g., participants coming from the same team).

For positive and negative affect, the base models were linear regression models that included suppression and reappraisal scores as the only two factors. In successive models, we added gender, then age, then time, and finally the interactions between time and the emotion regulation variables. We added gender before age or time because prior literature has found stronger and more consistent gender differences in expression related to these variables [16], [17], [52].

Tables 3 and 4 summarize the models and include the total variance (R^2) explained by each model. Fig. 4 shows the regression estimates for the positive and negative affect models (Model 5 in Tables 3 and 4 respectively). In the end, we chose to include the final model for both PA and NA for consistency. However, as seen in Table 3, a model with just suppression and reappraisal scores and gender (model 2) would be sufficient for explaining the data. Significantly more of the variance in the data is explained by adding gender to the base model ($R^2 = .042$), but the other factors do not account for significantly more variance than model 2. Therefore, it seems that gender is the most important predictor of PA.

Note that the significance of a R^2 value for a model means that the increase in R^2 for the model from the previous model was statistically significant.

Table 2 shows the correlations between the main variables in our analysis: positive and negative affect base rates, cognitive reappraisal scores, emotion suppression scores, age, and time. Besides the moderate correlation between negative affect and age, the variables have no, or very weak,

TABLE 3: **Positive Affect.** Parameter estimates for linear mixed effects models. Model 2 explains more of the variance (R^2 = .042) in the data than Model 1. None of the subsequent models are significantly better than Model 2. (. p < .08, * p < .05, ** p < .01, *** p < .001)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-4.133 ***	-3.709 ***	-3.785 ***	-3.822 ***	-3.826 ***
Suppression	0.018	0.100	0.106	-0.105	0.137
Reappraisal	0.165	0.134	0.140	.140	0.196
Gender		-0.656 **	-0.683 **	-0.684 **	-0.685 **
Age			0.006	0.006	0.006
Time (hours since 7am)				0.006	0.007
Time*Suppression					-0.006
Time*Cognitive Reappraisal					-0.009
R^2	0.010	0.042 **	0.043	0.043	0.044

TABLE 4: **Negative Affect.** Parameter estimates for linear mixed effects models (. p < .08, * p < .05, ** p < .01, *** p < .001)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-4.762 ***	-4.848 ***	-5.236 ***	-5.432 ***	-5.434 ***
Suppression	-0.070	-0.087	-0.060	-0.065	0.036
Reappraisal	0.075	0.081	0.107	0.110	0.196 *
Gender		0.133	-0.003	-0.009	-0.010
Age			0.029 **	0.029 **	0.029 **
Time (hours since 7am)				0.033 ***	0.033 ***
Time*Suppression					-0.018 ***
Time*Cognitive Reappraisal					-0.014 **
R^2	0.006	0.007	0.049 **	0.055 ***	0.059 ***

TABLE 5: Mixed Linear Effects Model for PA. ES = Emotional Suppression. CR = Cognitive Reappraisal. ES and CR scores are centered on the mean. Time refers to hours since 7am. (. p <= .1, * p < .05, ** p < .01, *** p < .001)

Variable	Estimate	SE	t-score	Pr (> |t|)	Sig.
Intercept	-3.826	0.264	-14.502	<0.01	***
ES	0.137	0.122	1.124	0.263	
CR	0.196	0.120	1.629	0.106	
Gender	-0.685	0.248	-2.757	0.007	**
Age	0.006	0.013	0.440	0.661	
Time	0.007	0.007	0.952	0.341	
ES*Time	-0.006	0.007	-0.839	0.402	
CR*Time	-0.009	0.007	-1.335	0.182	

TABLE 6: Mixed Linear Effects Model for NA. ES = Emotional Suppression. CR = Cognitive Reappraisal. ES and CR scores are centered on the mean. Time refers to hours since 7am. (. p < .1, * p < .05, ** p < .01, *** p < .001)

Variable	Estimate	SE	T-score	Pr (> |t|)	Sig.
Intercept	-5.434	0.218	-24.88	<0.01	***
ES	0.036	0.099	0.361	0.719	
CR	0.196	0.098	2.007	0.047	*
Gender	-0.010	0.206	-0.049	0.961	
Age	0.029	0.011	2.660	0.009	**
Time	0.033	0.005	7.141	<0.01	***
ES*Time	-0.018	0.005	-3.846	0.000	***
CR*Time	-0.014	0.005	-3.122	0.002	**

correlations with one another. This supports that PA and NA are separate constructs and so are ES and CR. The strongest correlations were a weak negative relationship between positive affect and ES and a positive correlation between age and negative affect. Higher suppression was related to lower expressed positive affect. Aging was related to more frequent expressions of negative affect. The latter of these effects may be due to a bias in the classifier, as discussed in the limitations section.

Table 5 shows the model estimates, standard errors, t-scores and p values for the positive affect model. Table 6 shows these values for the negative affect model. In the supplementary material (Tables 3 and 4) we show the model estimates and significance indicators.

5.1 Emotion Suppression

We found no evidence to support **H1**. Emotion suppression had no significant main effect on positive or negative affect.

5.2 Cognitive Reappraisal

We hypothesized in **H2a** that participants with higher reported cognitive reappraisal scores would generally express more positive facial affect than those with lower reported cognitive reappraisal. Although cognitive reappraisal had no statistically significant main effect on the PA model, there is still a trend in the direction we hypothesized (p = .106), based on the positive parameter estimate for cognitive reappraisal (in Table 5), such that PA was higher for those that used cognitive reappraisal more.

Differing from prior work [9] and what was expected in **H2b**, cognitive reappraisal had a main effect on negative affect (p=.047). Participants who self-reported as having higher cognitive reappraisal scores expressed more negative affect on average.

5.3 Gender

There was a significant gender effect for self-reported suppression scores and expressed positive affect, partially supporting **H3**. Males scored approximately 17% higher on reported emotional suppression than females (male mean= 17.3; female mean = 14.8; t(15) = 2.606, p = 0.01, significant). To compare expression of PA and NA between men and women, we used two-sample Kolmogorov-Smirnov (KS) hypothesis tests because the base rates were not normally
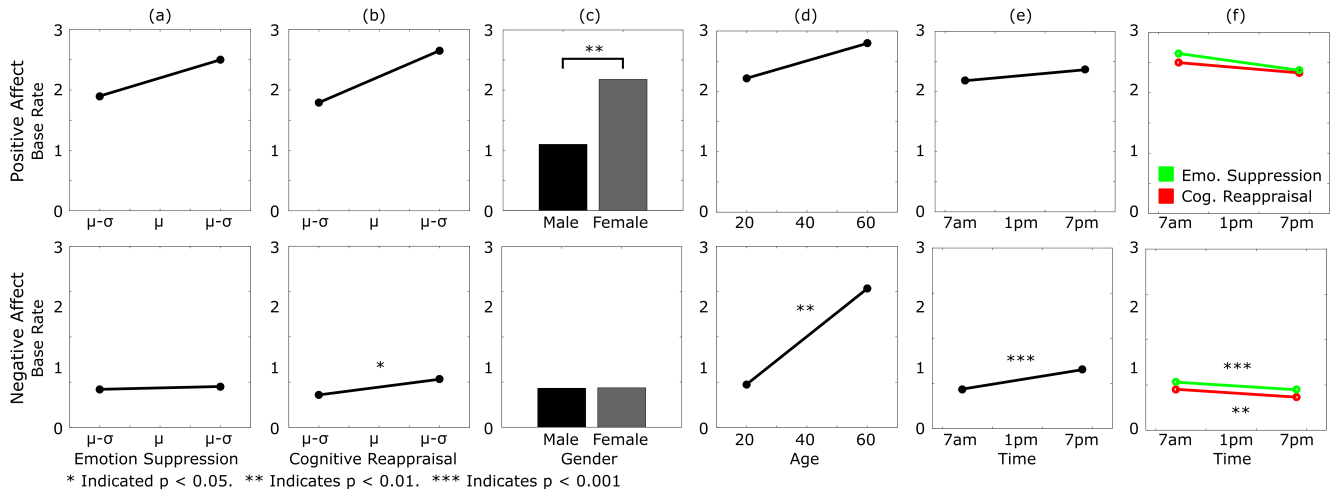
Fig. 4: Regression estimates for the (a) emotion suppression, (b) cognitive reappraisal, (c) gender, (d) age and (e) time main effects. Estimates for the (f) emotion suppression by time and cognitive reappraisal by time interactions. For the emotion regulation effects, estimates for the mean score and one standard deviation above and below the mean score score are shown.

distributed. Females displayed 27.7% more positive affect than males (male mean = 3.03; female mean = 3.87; D(117) = 0.297, $p = 0.027$). There was no significant difference in negative affect expressed by males or females.

5.4 Age

In **H4** we expected age differences in cognitive reappraisal and suppression scores and increased expression of PA and NA. The findings partially confirmed our hypothesis. There were no age differences in emotion regulation scores, but age had a significant effect on displayed affect. People over 35 years old displayed a trend toward more positive affect (12.1%) than people 35 years old or younger (35 years or under mean = 3.15; over 35 years mean = 3.53; D(117) = 0.256, $p < 0.09$). People over 35 years old also displayed significantly (50.0%) more negative affect than people 35 years old or younger (35 years old or under mean = 0.46; over 35 years mean = 0.69; D(117) = 0.423, $p < 0.01$). However, in our temporal models only the NA result remained significant. It should noted that the strong effect on NA with respect to age may be due to classifier bias with old people having more false positives for NA expressions, which we discuss further in the limitations section.

5.5 Temporality of Affect

Overall, people expressed more negative affect over the course of the day, regardless of their emotion regulation strategy, gender, and age.

Confirming **H5**, we observed that, for participants who scored higher on emotion suppression, the rate at which they expressed negative affect over the course of the day was slower (p<.001). There was no significant interaction effect of emotion suppression and time for positive affect. In line with **H6**, the rate at which participants who reported more cognitive reappraisal expressed negative affect grew more slowly than those who reported less cognitive reappraisal (p=.002).

6 DISCUSSION

Our results support prior work and extend it by introducing a temporal analysis of the impact of emotion regulation strategies on expressions of positive and negative affect at work, in situ.

First, we found that males reported a higher use of suppression as an emotion regulation strategy than females [9]. Second, we find that females show positive affect significantly more often than males. This supports a large body of prior work [52]. Third, we do not find that females show negative affect less often than males when accounting for age and emotion regulation strategies. While numerous studies of gender differences have found that negative emotions are less common in females, these differences are dependent on the particular type of negative affect studied, the way it was measured, and the time of day the study was conducted. Given our study's longitudinal nature and our focus on NA as a whole, it is not surprising that there would not be an overall significant gender difference in expressed NA in our sample.

The expression of positive and negative affect is subject to common diurnal emotion patterns and thus it is important that temporal dynamics are considered when doing research on this topic. Our temporal analysis contextualizes emotion regulation effects in the presence of daily patterns of emotion and work. The process model of emotion regulation predicts that people who habitually use cognitive reappraisal strategies should experience and express more positive affect and less negative affect. The model also predicts that individuals who chronically use emotion suppression strategies should express less negative affect behaviorally. Based on this, we hypothesized that facial expressions of negative affect would be lower in those that use reappraisal and suppression strategies.

Our results partially support findings that people who habitually use cognitive reappraisal strategies display more positive affect over time. We find an interesting relationship between time of day and the use of cognitive reappraisal
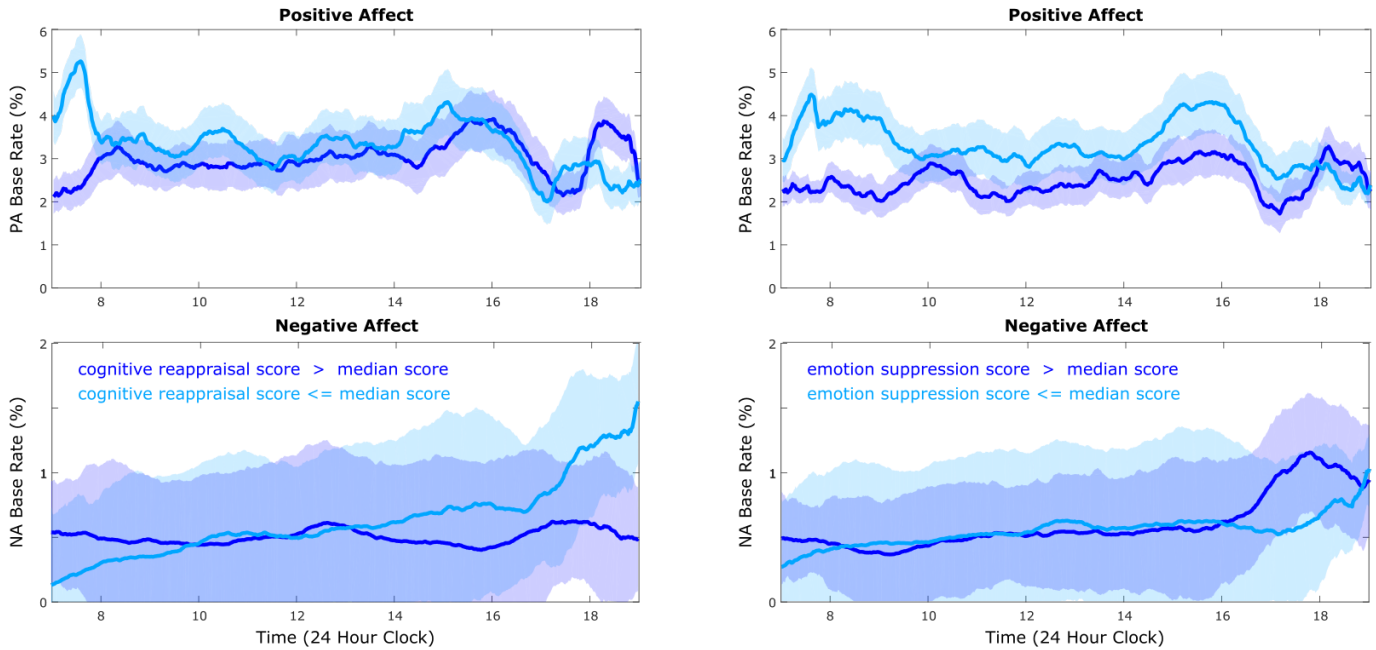
(a) Plots for cognitive reappraisal. Notice how more infrequent use of cognitive reappraisal as a strategy (and thus a lower score) is related to a faster increase in expressed negative affect across the course of the day.

(b) Plots for emotion suppression. Notice how greater use of emotion suppression as a strategy is related to lower expression of positive affect and greater negative affect toward the end of the day.

Fig. 5: Average PA and NA trends between 7am and 7pm for high and low (a) cognitive reappraisal and (b) emotion suppression scores. Scores above the median (dark blue) and scores below the median (light blue). The shaded regions represent 90% confidence intervals.

in the expression of negative affect. In accordance with expected diurnal patterns, as a day progresses, people express more negative affect. Yet, people who report using more cognitive reappraisal skills show slower increases in the expression of negative affect, suggesting that cognitive reappraisal is an effective strategy in managing daily patterns of negative affect. See Figure 5 for average PA and NA trends between 7am and 7pm for high and low cognitive reappraisal and emotion suppression scores. Notice how more infrequent use of cognitive reappraisal as a strategy is related to a faster increase in expressed negative affect across the course of the day in Figure 5(a). We also observed that there was a higher level of negative affect expressed at the beginning of the day by people who reported using cognitive reappraisal more. Our best explanation of this is that it is the result of more consistent behavior overall, with respect to the expression of negative affect. Something that might be expected of those with better emotion regulation skills.

Suppression shows a relationship with decreases in expression of negative affect over the course of a day but, unlike cognitive reappraisal, is not correlated with daily patterns of positive affect. As such, cognitive reappraisal appears to be a more effective strategy for regulating healthier daily emotion. Qualitatively, plots of the average patterns of positive affect for those employing more suppression strategies, show reduced dynamics and a flatter profile. The top plot of Figure 5(b) shows that those with a high suppression score seemed to have a damped profile of expressed positive affect.

Overall, our results highlight that time and emotion regulation strategies are important factors to consider when studying affective expressions. The advances in affective computing and facial expression analysis have enabled new ways to measure affect longitudinally and we are excited about the opportunities this presents for future research.

7　IMPLICATIONS

Our longitudinal empirical findings have implications for the study of affect, the development of emotionally aware technology, and the application of such technology.

Time is important to consider when collecting and interpreting behavioral, affective signals. Our results, in line with previous research, show that the base rate frequency of expressions of PA and NA change significantly over the course of the average day, especially for NA. Furthermore, these changes interact strongly with the types of emotion regulation strategies that people employ. Therefore, researchers interested in running affect studies should take into consideration the times of day their studies occur when scheduling them and interpreting the results that are collected. The same participant might be more likely to express more negative affect in the evening than in the morning. This difference may be further exaggerated by the participant's emotion regulation strategy (i.e., cognitive reappraisal or suppression), gender and age. We are not aware of many studies of facial affect that control for the time of day or emotion regulation strategies when analyzing behaviors.

Our findings on individual differences in emotion regulation and expression over the course of a day can also be applied to the design of emotionally aware, artificial agents. Computational models controlling agents should integrate knowledge of people's emotion regulation strategies into their decision processes. For instance, someone who suppresses their emotion may appear to be less angry or frustrated than someone who reappraises their emotion. Artificial agents should also be sensitive to the local time when interacting with humans and, among healthy adults, expect increased negative expression as the day progresses. Artificial agents should anticipate emotion regulation and temporal differences and learn to detect and address people differently based on their emotion regulation habits.

Embodied agents who seek to mimic humans may need to artificially suppress emotion with more neutral "faces" when interacting with people who employ suppression [53]. With our empirical work, we expand the design space of emotional agents to include emotion regulation strategies and time. Future work is needed to understand when and how avatars should express or suppress emotions to follow or complement human norms. This paper uncovers the importance of temporality and that agents may need to control their emotion by employing as diverse a set of strategies as humans do. The impact of agents employing these strategies on their perceived sociability is an open area of research.

Based on the present and prior work, training courses in workplaces and schools should focus on cognitive reappraisal and be sensitive to gender differences in emotional expression. Significant emotional shifts occur over the course of the day, and strategies to cope, reframe, and communicate emotional experiences are important for individual and group well-being.

Finally, the ubiquity of everyday video cameras and access to validated computer vision models enabled this study to be conducted in a large technology organization with the employees' consent. The improved development and use of affective sensing technologies will continue to be important for deepening knowledge of human emotion and opening up the design space at the intersection of technology and human emotion.

8 LIMITATIONS AND FUTURE WORK

There are a few important limitations to our study. In line with prior work, we focused on the expression of affect, but our facial analyses do not directly measure the experience of affect. Expressing affect is an important aspect, but not the entirety of, experiencing emotions.

Our reliance on a camera attached to the participant's work stations means that we were not able to measure people outside of working hours. Therefore, we did not capture the full daily diurnal patterns of our participants. Future work with even more complete tracking may find more nuanced effects of emotion regulation on diurnal patterns. A mobile application which collects similar data would help us obtain a complete representation of affect throughout the day. Additionally, although the web camera was unobtrusive, participants were at least partially aware that the camera was on. Although it is unlikely that participant behavior was altered for several days and weeks on end, awareness of the camera and email tracking may have influenced behavior. It is also possible that the participants that opted-in to our study were biased as a result of self-selection.

Future work should extend the present work and examine the impact of different cultural context, job roles (beyond information workers), personality, and age on emotion regulation and affect. We found that older participants expressed more affect and hypothesized this was a result of increased emotion regulation ability. However, the driver of this age effect is unknown. There could be potentially other factors, such as job role or social network size, that influence emotion regulation and expression. These data were collected in a signal geographic location in the Northwest of the United States, and thus our population was predominantly people who have lived in this culture for a significant period. We would need to replicate this work in other US regions and countries to examine cultural effects rigorously. Further work also needs to be carried out in order to understand how generalizable results from this population are to broader domains.

Using automated classifiers is for all intents and purposes the only practical approach to longitudinal analysis of facial behavior. Manual coding would be too labor and time intensive. We analyzed almost 2,000 hours of videos featuring faces in this study. Manual annotation of such data may well have taken 20 times that duration (~40,000 hours) to manually code. This is equivalent to over 10 years of coding 10 hours per day, 365 days per year.

Although automated approaches are necessary for scalability, automated classifiers are not perfect and can be subject to biases. One such bias may be that the NA facial expression classification has more false positives for older people, due to deepening of the nasolabial folds and the furrow lines as people grow older. Therefore, we are careful not to place too much weight on the age specific results. We have tried to interpret our results in light of prior work such that we can be more confident of the findings, especially where they are consistent with previous theories and laboratory experiments. Furthermore, time-based effects across the course of the day are unlikely to be due to demographic biases (such as age) because peoples' appearances tend to be fairly consistent from one hour to the next and one day to the next.

Future work could also collect measures of physiological arousal using non-contact forms of measurement and would allow us to look at the rate of change of expressed affect as a function of time and maintain the unobtrusive nature of our data collection system [54], [55].

9 CONCLUSIONS

Emotion regulation plays an important role in psychological and physical health. We conducted a longitudinal, large scale observational study to investigate the role of emotion regulation on expressed affect, in situ, over the course of the day at work. Corroborating previous findings, we observed that males scored significantly higher on emotional suppression tactical use and displayed significantly less positive affect than females. Extending previous findings, we performed a time-based analysis using linear mixed effects

models. Our models revealed that time is an important factor in the regulation and expression of emotion, which has thus far been overlooked.

Our results reveal the importance of considering gender, age, time, and emotion regulation strategies in the interpretation of affective signals. These considerations will be critical in the useful and usable design of conversational agents that respond respectfully to human users, as well as emotionally sentient robots and other assistive technology.

Future work should look at how aspects of personality and other traits may affect emotion regulation and expression longitudinally.

REFERENCES

[1] P. Vuilleumier, "How brains beware: neural mechanisms of emotional attention," *Trends in cognitive sciences*, vol. 9, no. 12, pp. 585–594, 2005.

[2] E. A. Phelps, "Emotion and cognition: insights from studies of the human amygdala," *Annu. Rev. Psychol.*, vol. 57, pp. 27–53, 2006.

[3] R. Miranda and J. Kihlstrom, "Mood congruence in childhood and recent autobiographical memory," *Cognition & Emotion*, vol. 19, no. 7, pp. 981–998, 2005.

[4] G. Loewenstein and J. S. Lerner, "The role of affect in decision making," *Handbook of affective science*, vol. 619, no. 642, p. 3, 2003.

[5] J. J. Gross, "Emotion regulation: taking stock and moving forward." *Emotion*, vol. 13, no. 3, p. 359, 2013.

[6] ——, *Handbook of emotion regulation*. Guilford publications, 2013.

[7] L. Campbell-Sills and D. H. Barlow, "Incorporating emotion regulation into conceptualizations and treatments of anxiety and mood disorders." 2007.

[8] R. W. Picard, "Affective computing," *The MIT Press, Cambridge (MA)*, vol. 167, p. 170, 1997.

[9] J. J. Gross and O. P. John, "Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being." *Journal of personality and social psychology*, vol. 85, no. 2, p. 348, 2003.

[10] E. A. Butler, B. Egloff, F. H. Wlhelm, N. C. Smith, E. A. Erickson, and J. J. Gross, "The social consequences of expressive suppression." *Emotion*, vol. 3, no. 1, p. 48, 2003.

[11] A. M. Tackman and S. Srivastava, "Social responses to expressive suppression: The role of personality judgments." *Journal of personality and social psychology*, vol. 110, no. 4, p. 574, 2016.

[12] T. English, O. P. John, S. Srivastava, and J. J. Gross, "Emotion regulation and peer-rated social functioning: A 4-year longitudinal study," *Journal of Research in Personality*, vol. 46, no. 6, pp. 780–784, 2012.

[13] J. M. Girard and D. McDuff, "Historical heterogeneity predicts smiling: Evidence from large-scale observational analyses," in *Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on*. IEEE, 2017, pp. 719–726.

[14] J. Hernandez, M. E. Hoque, W. Drevo, and R. W. Picard, "Mood meter: counting smiles in the wild," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 301–310.

[15] D. McDuff, "Smiling from adolescence to old age: A large observational study," in *Affective Computing and Intelligent Interaction (ACII), 2017 Seventh International Conference on*. IEEE, 2017, pp. 98–104.

[16] D. McDuff, J. M. Girard, and R. El Kaliouby, "Large-scale observational evidence of cross-cultural differences in facial behavior," *Journal of Nonverbal Behavior*, vol. 41, no. 1, pp. 1–19, 2017.

[17] D. McDuff, E. Kodra, R. el Kaliouby, and M. LaFrance, "A large-scale analysis of sex differences in facial expressions," *PloS one*, vol. 12, no. 4, p. e0173942, 2017.

[18] M. Lindström, A. Ståhl, K. Höök, P. Sundström, J. Laaksolathi, M. Combetto, A. Taylor, and R. Bresin, "Affective diary: designing for bodily expressiveness and self-reflection," in *CHI'06 extended abstracts on Human factors in computing systems*. ACM, 2006, pp. 1037–1042.

[19] A. A. Stone, J. E. Schwartz, D. Schkade, N. Schwarz, A. Krueger, and D. Kahneman, "A population approach to the study of emotion: diurnal rhythms of a working day examined with the day reconstruction method." *Emotion*, vol. 6, no. 1, p. 139, 2006.

[20] L. M. Kurina, B. Schneider, and L. J. Waite, "Stress, symptoms of depression and anxiety, and cortisol patterns in working parents," *Stress and Health*, vol. 20, no. 2, pp. 53–63, 2004.

[21] T. Bogg and R. B. Slatcher, "Activity mediates conscientiousness relationship to diurnal cortisol slope in a national sample." *Health Psychology*, vol. 34, no. 12, p. 1195, 2015.

[22] L. D. Doane, S. Mineka, R. E. Zinbarg, M. Craske, J. W. Griffith, and E. K. Adam, "Are flatter diurnal cortisol rhythms associated with major depression and anxiety disorders in late adolescence? the role of life stress and daily negative emotion," *Development and Psychopathology*, vol. 25, no. 3, pp. 629–642, 2013.

[23] D. B. Boivin, C. A. Czeisler, D.-J. Dijk, J. F. Duffy, S. Folkard, D. S. Minors, P. Totterdell, and J. M. Waterhouse, "Complex interaction of the sleep-wake cycle and circadian phase modulates mood in healthy subjects," *Archives of general psychiatry*, vol. 54, no. 2, pp. 145–152, 1997.

[24] A. A. Stone and S. Shiffman, "Ecological momentary assessment (ema) in behavorial medicine." *Annals of Behavioral Medicine*, 1994.

[25] N. Bolger, A. Davis, and E. Rafaeli, "Diary methods: Capturing life as it is lived," *Annual review of psychology*, vol. 54, no. 1, pp. 579–616, 2003.

[26] D. Kahneman, A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone, "A survey method for characterizing daily life experience: The day reconstruction method," *Science*, vol. 306, no. 5702, pp. 1776–1780, 2004.

[27] D. McDuff, A. Karlson, A. Kapoor, A. Roseway, and M. Czerwinski, "Affectaura: an intelligent system for emotional memory," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 849–858.

[28] M. De Choudhury and S. Counts, "Understanding affect in the workplace via social media," in *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 2013, pp. 303–316.

[29] C. E. Izard and C. Z. Malatesta, "Perspectives on emotional development i: Differential emotions theory of early emotional development." in *The first draft of this paper was based on an invited address to the Eastern Psychological Association, Apr 1, 1983.* John Wiley & Sons, 1987.

[30] J. J. Gross, "Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology." *Journal of personality and social psychology*, vol. 74, no. 1, p. 224, 1998.

[31] ——, "Emotion regulation: Affective, cognitive, and social consequences," *Psychophysiology*, vol. 39, no. 3, pp. 281–291, 2002.

[32] G. Sheppes, S. Scheibe, G. Suri, and J. J. Gross, "Emotion-regulation choice," *Psychological science*, vol. 22, no. 11, pp. 1391–1396, 2011.

[33] D. DeSteno, J. J. Gross, and L. Kubzansky, "Affective science and health: The importance of emotion and emotion regulation." *Health Psychology*, vol. 32, no. 5, p. 474, 2013.

[34] J. J. Gross, "Emotion regulation: Current status and future prospects," *Psychological Inquiry*, vol. 26, no. 1, pp. 1–26, 2015.

[35] T. English, I. A. Lee, O. P. John, and J. J. Gross, "Emotion regulation strategy selection in daily life: The role of social context and goals," *Motivation and emotion*, vol. 41, no. 2, pp. 230–242, 2017.

[36] A.-L. Rottweiler, J. L. Taxer, and U. E. Nett, "Context matters in the effectiveness of emotion regulation strategies," *AERA Open*, vol. 4, no. 2, p. 2332858418778849, 2018.

[37] G. A. Bonanno and C. L. Burton, "Regulatory flexibility: An individual differences perspective on coping and emotion regulation," *Perspectives on Psychological Science*, vol. 8, no. 6, pp. 591–612, 2013.

[38] J. Hardy, "Distinguishing among within-person variability: Affective intra-individual variability, affective psychological flexibility, and health in a national us sample," 2015.

[39] M. L. Cooper, P. R. Shaver, and N. L. Collins, "Attachment styles, emotion regulation, and adjustment in adolescence." *Journal of personality and social psychology*, vol. 74, no. 5, p. 1380, 1998.

[40] R. A. Fabes and N. Eisenberg, "Regulatory control and adults' stress-related responses to daily life events." *Journal of personality and social psychology*, vol. 73, no. 5, p. 1107, 1997.

[41] R. L. Repetti, S. E. Taylor, and T. E. Seeman, "Risky families: family social environments and the mental and physical health of offspring." *Psychological bulletin*, vol. 128, no. 2, p. 330, 2002.

[42] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training deep networks for facial expression recognition with crowd-sourced label distribution," in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 2016, pp. 279–283.

[43] P. Ekman, "Universals and cultural differences in facial expressions of emotion." in *Nebraska symposium on motivation*. University of Nebraska Press, 1971.

[44] P. Ekman and K. G. Heider, "The universality of a contempt expression: A replication," *Motivation and emotion*, vol. 12, no. 3, pp. 303–308, 1988.

[45] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[46] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 94–101.

[47] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee *et al.*, "Challenges in representation learning: A report on three machine learning contests," in *International Conference on Neural Information Processing*. Springer, 2013, pp. 117–124.

[48] O. P. John and S. Srivastava, "The big five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and research*, vol. 2, no. 1999, pp. 102–138, 1999.

[49] 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.

[50] D. Watson, *Mood and temperament*. Guilford Press, 2000.

[51] S. A. Golder and M. W. Macy, "Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures," *Science*, vol. 333, no. 6051, pp. 1878–1881, 2011.

[52] M. LaFrance, M. A. Hecht, and E. L. Paluck, "The contingent smile: A meta-analysis of sex differences in smiling." *Psychological bulletin*, vol. 129, no. 2, p. 305, 2003.

[53] D. McDuff and M. Czerwinski, "Designing emotionally sentient agents," *Communications of the ACM*, vol. 61, no. 12, pp. 74–83, 2018.

[54] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation." *Optics express*, vol. 18, no. 10, pp. 10762–10774, 2010.

[55] W. Chen and D. McDuff, "Deepphys: Video-based physiological measurement using convolutional attention networks," *arXiv preprint arXiv:1805.07888*, 2018.

Kael Rowan is a Principal Research Software Design Engineer in Microsoft Research AI. He joined Microsoft Corporation in 2000 after graduating with a degree in Computer Science from Washington State University. First he was a member of the Windows Embedded group before later moving to Microsoft Research in 2004. Areas of research and development range from foundations of software engineering to human-computer interaction. He is currently a member of the Human Understanding and Empathy group in MSR AI working on human sensing and synthesis technologies.
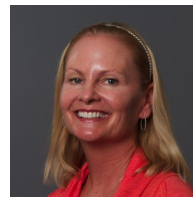
Mary Czerwinski is a Principal Researcher and Research Manager of the Human Understanding and Empathy (HUE) Research Group at Microsoft Research AI. Mary's latest research focuses primarily on emotion tracking and intervention design, health and wellness for individuals and groups and productivity at work. Her research background is in visual attention and multitasking. She holds a Ph.D. in Cognitive Psychology from Indiana University in Bloomington. Mary was awarded the ACM SIGCHI Lifetime Service Award, was inducted into the CHI Academy and received the Distinguished Alumni award from Indiana University's College of Arts and Sciences. Mary is a Fellow of the ACM and the American Psychological Science Association. More information about Dr. Czerwinski can be found at her website: https://www.microsoft.com/en-us/research/people/marycz/.

Daniel McDuff (S'09, M'15) is a researcher at Microsoft Research AI. He received the bachelor's degree, with first-class honors and master's degree in engineering from Cambridge University. He received his PhD from the MIT Media Lab in 2014, working in the Affective Computing group. Prior to joining the Media Lab, he worked for the Defense Science and Technology Laboratory (DSTL) in the United Kingdom. He is interested in computer vision and machine learning to enable the automated recognition of affect. He is also interested in technology for remote measurement of physiology.


Eunice Jun (S'16) is a PhD student in the Paul G. Allen School of Computer Science and Engineering at the University of Washington. In her research, she develops tools to support experts (e.g., data scientists and researchers) and novices design, conduct, and analyze data from large-scale experiments.