Emotional Footprints of Email Interruptions

Christopher Blank[*]
University of Houston
Houston, Texas
blankc13@gator.uhd.edu

Shaila Zaman
University of Houston
Houston, Texas
szaman4@uh.edu

Amanveer Wesley
University of Houston
Houston, Texas
awesley4@uh.edu

Panagiotis Tsiamyrtzis
Politecnico di Milano
Milano, Italy
panagiotis.tsiamyrtzis@polimi.it

Dennis R. Da Cunha Silva
Texas A&M University
College Station, Texas
silva.dennis@tamu.edu

Ricardo Gutierrez-Osuna
Texas A&M University
College Station, Texas
rgutier@cse.tamu.edu

Gloria Mark
University of California
Irvine, California
gmark@uci.edu

Ioannis Pavlidis[†]
University of Houston
Houston, Texas
ipavlidis@uh.edu

ABSTRACT
Working in an environment with constant interruptions is known to affect stress, but how do interruptions affect emotional expression? Emotional expression can have significant impact on interactions among coworkers. We analyzed the video of 26 participants who performed an essay task in a laboratory while receiving either continual email interruptions or receiving a single batch of email. Facial videos of the participants were run through a convolutional neural network to determine the emotional mix via decoding of facial expressions. Using a novel co-occurrence matrix analysis, we showed that with batched email, a neutral emotional state is dominant with sadness being a distant second, and with continual interruptions, this pattern is reversed, and sadness is mixed with fear. We discuss the implications of these results for how interruptions can impact employees' well-being and organizational climate.

Author Keywords
Email interruptions; emotions; facial expressions; convolutional neural network; co-occurence matrix.

CCS Concepts
•**Human-centered computing** → **Laboratory experiments;** *User studies;*

[*]Blank, Zaman, and Wesley contributed equally as first authors.
[†]Pavlidis and Mark contributed equally as senior authors.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
https://doi.org/10.1145/3313831.3376282

INTRODUCTION
Emotions are inherent in organizational life. Organizations are social settings and the display of emotions can affect coworkers. For example, the positive display of emotion in the workplace is associated with greater interpersonal interaction [47], whereas the negative display of emotion can lead to less cooperative behavior [20]. Emotions have been studied extensively as reactions to significant organizational events, such as reorganizations. There has been less attention, however, to how emotions manifest in everyday organizational life [3].

For people who work with information, communication, and technology (ICT), a large portion of their day has been characterized as switching attention among different applications and devices, i.e., they engage in multitasking. Task switching is triggered by interruptions, which can be from external sources (e.g., notifications) [18], or initiated by oneself (e.g., remembering to do something) [21]. Interruptions and multitasking have received a fair amount of attention in the human-computer interaction (HCI) community, because they are pervasive and are associated with stress [13, 9, 34]. Email in particular, has received attention in HCI as a significant source of interruptions and distractions in the workplace [5, 36, 17, 14, 51].

But how does working in an environment with constant interruptions affect emotional expression? Stress has been measured in the workplace context physiologically and subjectively (see [2] for a review), but how are emotions manifest when people are multitasking? Because face-to-face or remote interaction is commonplace among information workers, especially with those who are tightly coupled in work, emotional expression can have significant impact on coworkers. Research on emotional contagion demonstrates how easily emotions can be transferred and picked up by others in a group [7].

The goal of this paper is to investigate what emotional expressions people convey when they are working in an environment with constant interruptions. While the effect of interruptions on stress is well-known, and while the act of multitasking has been shown to be associated with negative mood [35, 32], it has not been investigated how working in an environment with continual interruptions is manifest in people's emotional expression. To examine this, we used automated methods to detect and categorize the emotional expression of participants who were subjected to either high or low rates of multitasking. The results show that high multitasking is associated with a higher display of sadness compared to low multitasking. This research is part of a larger project examining the effects of email and multitasking [1]. The contributions of this paper are:

- Automated detection of emotions during multitasking has been done for the first time.
- Quantification of displayed emotions has been done via co-occurrence matrices - a novel method that properly accounts for mixed vs. single emotions.
- Study results show that participants who multitask more exhibit more sadness, compared to those who multitask less.

RELATED WORK

We consider multitasking as the rapid shift of attention among different activities. In the workplace, it has been documented as quite prevalent [13, 21]. Recent research shows that for information workers, attention duration on a computer screen averaged 47 seconds [36], and for software developers, it was 50.4 seconds [40]. With low standard deviations, these measures indicate that typically people's attention shifts continuously among different computer screens throughout the workday. In fact, it is not just a single distraction that can lead people's attention away from a task at-hand, but rather there can be chains of distraction [27]. Attention shifting has also been documented across devices, and not just while working on the laptop/desktop [29].

Email is an integral part of the information workplace and has shown to be a significant source of interruptions. A diary study showed that email accounted for 24% of the daily tasks of information workers [13]. Research shows that people check their email quite often daily, either triggered through email notifications or self-checking. It was found that 84% of users keep their email client in the background at all times [45], which provides ample opportunity for interruptions. Using computer logging data, people have been shown to check their email on average 77 times a day [36]. The Radicati group reported in 2015 that about 88 emails were received per day and 33 sent [23], and earlier Fisher et al. [17] reported that people received an average of 87 emails per day.

Multitasking, interruptions, and stress

Stress is a negative emotional experience associated with a complex array of emotions such as fear, dread, and sadness [42]. Research has shown that multitasking and interruptions are associated with stress [9, 34]. In examining four types of interruptions, it was found that stress was associated with all types, based on self-reports [19].

Email has been shown to be a component of workplace stress [5, 14]. Studies show there is causal attribution for email to induce stress: people experience less stress when they check their email less frequently [30] and when email is cut off [37]. Thus, research shows that multitasking and interruptions (particularly email), which are commonplace in information work, are associated with stress.

Emotions in the workplace

Whereas it is not possible to fully understand felt emotions based on the display of emotions, felt and displayed emotions are closely intertwined [47, 52]. But irrespective of the felt emotion, the display of emotions in the workplace can have significant effects. Emotional contagion can spread in a group or workplace through the influence of conscious or unconscious processes involving emotional states or physiological responses [8, 43]. Emotional displays are a strong antecedent of social influence [3]. In an experimental study, Barsade [7] found evidence for emotional contagion for both positive and negative emotions, concluding that people are "walking mood inductors."

Positive and negative affect can influence different behaviors. Barsade [7] found that positive emotional contagion led to more cooperative group behaviors. Negative mood, on the other hand is associated with less prosocial behaviors in a group [20].

Some organizations are now paying attention to the emotional culture of the workplace based on the notion that not only do felt emotions shape employee satisfaction and team performance but also that emotional display can foster a particular culture and set norms, e.g., of anger or positivity [6, 44]. Thus, an examination of the emotions that are manifest when people experience interruptions, which comprises a fair amount of information work, is a first step towards understanding what shapes the emotional culture of a workplace.

Sensing emotion in situ and in the workplace

Identifying emotion in situ and in the workplace has led to the development of tools for unobtrusive detection. AffectAura [38] was designed to automatically detect emotions and support people in reflecting on their moods. Hernandez et al. [25] used a skin conductance sensor to measure stress at a call center. Social media posts of employees have been used to detect emotions in the workplace [16]. EmotionCheck is a device to help individuals regulate their emotions in situ [12]. Affect has been detected from body posture [48] and location [26].

Recognition of emotions from facial imagery has also been gaining ground the last few years. This is due to its unobtrusive nature, the wealth of valence information it conveys, and significant improvements in recognition accuracy thanks to powerful neural network algorithms and ever expanding training data [4, 22, 11, 39]. Our study capitalizes on these developments, bringing to bear vision-based emotion recognition tools and coupling them with novel analytic methods to address long standing HCI questions.

EXPERIMENTAL DESIGN

The experimental protocol was approved by the institutional review boards of the universities participating in this study. The authors executed this protocol in accordance with the approved guidelines, obtaining informed consent from each participant before conducting the experiments.

This study is part of a larger experimental study examining multitasking behavior [53]. The experiment consisted of five phases, and the main part of the experiment was a parallel group treatment reported here. Twenty six ($n = 26$, 18 females/8 males, age 24.69 ± 10.17) college students who volunteered to participate in the study agreed to have their recorded facial video publicly released. These volunteers were randomly assigned to two groups. Both groups were given 50 min to compose an essay while interrupting themselves as necessary to respond to eight emails - a Dual Task (DT) assignment.

To enforce consistency in the execution of the experiment and the handling of the email interruptions, we developed a custom interface (p-Interface) in Javascript. The p-Interface implemented the experimental protocol, guiding participants step by step through the designed treatments. Specifically, the p-Interface presented to the participants an editor to write their essays. It also featured an email client to deliver the email interruptions and allow participants to send back their responses.

Group treatments differed in the email delivery/response mode. In the Batch (B) group ($n = 13$), all eight emails arrived 10 min after the start of DT, and participants had 5 min to start replying to them. In the Continual (C) group ($n = 13$), individual emails arrived about 4 min after participants sent the previous email, with 10 s as a grace period to start replying to each new email. Hence, the B group had a single long-lasting interruption, while the C group experienced multiple short interruptions.

If participants did not start their reply within the transitional time allotted, the interface shifted into the email page in order to ensure consistency across participants of the same email group. In the B group, all participants answered all eight emails. In the C group, a few participants completed six or seven emails instead of eight. In these cases, participants were slow in composing email responses and due to sequential timing, one or two of the last emails in the delivery queue fell outside the DT's 50 min duration.

The essay topic was on the issue of technological singularity, that is, when machines overtake human intelligence. We chose this topic because is of broad interest to college students. The email set consisted of five emails that asked for opinion/advice and three emails that had scheduling tasks (order randomized). The three scheduling emails asked participants to schedule a meeting among a professor, a student, and an administrator, given calendar constraints. The five opinion/advice email prompts were chosen from a pilot study on MTurk, where an original selection of 30 emails were presented to 270 workers on the platform. Each email was presented to 9 different MTurk workers who were asked to compose a reply as if they worked for an organization. Then, we selected the five emails that generated the highest mean word count in replies; these included: (1) advice on domestic travel (e.g. how early to go to the airport); (2) advice on selection between 4-year and community colleges for an older person; (3) opinion on white lies; (4) advice on summer internship for a tech company; (5) advice on balancing study breaks.

Given the nature of the dual task, to ameliorate confounding factors, all participants had undergraduate education, were native English speakers, and used email regularly for their daily communications. At the end of the DT, participants completed the NASA TLX questionnaire [24] to gauge psychometrically the loading induced by the experiment's treatment. NASA TLX features six sub-scales with a common rating [1=Strongly disagree, 2=Disagree, 3=Somewhat disagree, 4=Neither agree or disagree, 5=Somewhat agree, 6=Agree, 7=Strongly agree]. The sub-scales are:

- Mental Demand: Perceived mental load induced by the complexity of the DT.
- Physical Demand: Perceived physical load induced by the nature of the DT.
- Temporal Demand: Perceived time pressure induced by the pace of the DT.
- Performance: Perceived success in executing the DT.
- Effort: Perceived amount of work expended to achieve the said level of the DT performance.
- Frustration: Perceived level of irritation/annoyance in performing the DT.

The curated experimental data are freely available on the Open Science Framework (OSF) [50]. The OSF repository holds biographic data, quantitative data, and ancillary media; it also contains the email prompts we used in the experiment. The quantitative data feature instantaneous emotional vectors and psychometric scores, while the ancillary media feature annotated facial videos of participants during DT.

Experimental Setup

We carried out the study in an office room. During the experimental session, a visual camera located atop the participant's desktop computer, continuously imaged her/his face. The camera was a Logitech HD Pro - C920 (Logitech, Newark CA) with spatial resolution 1920×1080; we set its image acquisition speed at 10 fps. The distance between the participant's face and the camera was about 1 m; its zoom setting was such that the participant's face fit well into the field of view, providing sufficient resolution for analysis of facial expressions. Participants carried out the assigned tasks on a Dell OptiPlex 7050 desktop computer, featuring an Intel QuadCore i7 - 7700 3.6 GHz processor, 16 GB RAM, and 1 TB solid state disk. The computer was connected to a Dell U2417H - Ultrasharp 24 in display.

METHODS

We used convolutional neural networks (CNN) [33] to obtain each moment a probabilistic estimate of the participant's emotional mix. Specifically, we employed a Keras implementation of CNN by Serengil [46], which was trained and tested on the FER dataset [10]. For participant p, the outcome for a CNN-processed facial frame at time t is a vector $\vec{V}_{p,t} = \{\text{Neutral}, \text{Surprised}, \text{Sad}, \text{Happy}, \text{Afraid}, \text{Disgusted}, \text{Angry}\}$. In this vector, each component $v_{p,t,i}$ represents the probability of the corresponding momentary emotion being manifested on the participant's face; thus, $\sum_{i=1}^{7} v_{p,t,i} = 1.0$.

As at each moment t, manifestations of elemental emotions either dominate or co-exist on participants' faces, co-occurrence matrices become an appealing analytic tool for this study. Co-occurrence matrices are often used in ecology for analyzing distributions of species in an ecosystem [49]. There are several ways of designing a co-occurrence matrix, with new methods reported regularly in the literature. Next, we describe the method we developed for the present application.

The seven element vector $\vec{V}_{p,t}$ conveys integrated emotional context that is difficult to be analyzed statistically. To disentangle this integrated context, we perform exhaustive computation of the joint probabilities between individual vector components:

$$\mathbf{M}_{p,t}(\vec{V}_{p,t}) \equiv \vec{V}_{p,t} \otimes \vec{V}_{p,t} \ , \tag{1}$$

where $\otimes$ is the outer tensor product. The thus formed matrix $\mathbf{M}_{p,t}$ constitutes a 2D probabilistic remapping of the original emotional vector $\vec{V}_{p,t}$. The diagonal elements $(v_{p,t,i} * v_{p,t,i})$ of this matrix represent the stand-alone probabilities of individual emotions, while its off-diagonal elements $(v_{p,t,i} * v_{p,t,j})$ represent the joint probabilities of pairwise mixed emotions. As matrix $\mathbf{M}_{p,t}$ holds all the outcomes of a probability space, the sum of its elements is 1 (Fig. 1):

$$\sum_{i,j} \mathrm{M}_{p,t,i,j} = 1 \ . \tag{2}$$

Consequently, the sum of all momentary co-occurrence matrices $\mathbf{M}_p = \sum_t \mathbf{M}_{p,t}$ is a matrix whose sum of cell values gives the total number of frames analyzed for participant p.

Subsequently, we produce the group co-occurrence matrices $\mathbf{M_B}$ and $\mathbf{M_C}$ (Fig. 2) - one for the participants that received the batch email treatment and one for the participants that received the continual email treatment, respectively. These group matrices are formed through summation of the individual participant matrices: $\mathbf{M_B} = \sum_b \mathbf{M}_{p_b}$ and $\mathbf{M_C} = \sum_c \mathbf{M}_{p_c}$.

Numerical Example

We demonstrate with a numerical example the way the co-occurrence matrix $\mathbf{M}_{p,t}$ of a momentary emotional vector $\vec{S}_{p,t}$ is formed. In the depicted case of Fig. 1, Sad has moderately strong presence (0.6), coexisting with a small amount of Disgusted (0.2) and minuscule amounts of Angry and Neutral (0.1 each).

Taking the outer product of $\vec{S}_{p,t}$ with itself is tantamount to multiplying each vector component $s_{p,t,i}$ with itself as well as all the other vector components $s_{p,t,j}$. The thus formed products are laid out in a matrix. Please note that because this matrix is symmetrical, for notational convenience we keep only the upper triangle, by adding to its elements the corresponding elements of the lower triangle (Fig. 1). The resulting co-occurrence matrix $\mathbf{M}_{p,t}$ displays the associative weights of the emotional mix specified in $\vec{S}_{p,t}$. For example, the diagonal value 0.36 ($= 0.6 \times 0.6$), which is the largest number in the matrix (Fig. 1), reflects the relatively dominating role of Sad. The value 0.24 ($= 0.2 \times 0.6 + 0.2 \times 0.6$) in the second row of the matrix (Fig. 1), reflects the associative weight of co-occurrence between Disgusted and Sad. This is the largest among all off-diagonal values, as it is the product of the two strongest probabilistic estimates coexisting in vector $\vec{S}_{p,t}$.

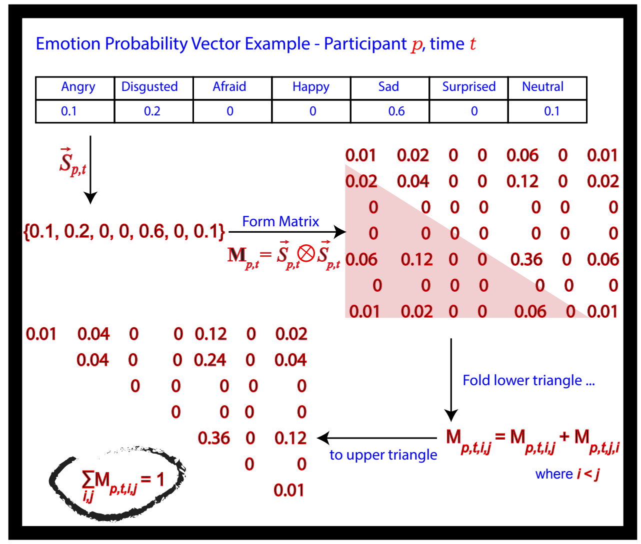

Figure 1: **Numerical example for the computation of a momentary co-occurrence matrix.** The outer product of the emotion probability vector with itself forms the co-occurrence matrix. In essence, the cells of the said matrix hold all the outcomes of joint probabilities between elemental emotions; for this reason, they sum up to 1. The lower triangle of the matrix is added to the upper triangle for notational convenience.

This numerical example lays bear the value of the co-occurrence method in quantifying displayed emotions. Up to now researchers extracted the raw component of interest $s_{p,t,i}$ from the emotion vector (e.g., 0.6 for Sad) for analytical use. The co-occurrence method begs to differ in this case by providing $s_{p,t,i} * s_{p,t,i}$ (e.g., 0.36 for Sad) - a much lesser value than the original. This reflects the fact that the $s_{p,t,i}$ component does not exist in isolation, but in the context of a vector; thus, its stand-alone probability is dictated by the law of joint probabilities. Expressed differently, analysts need to 'pay a penalty' if they want to analyze individual emotions; this penalty goes to cover the joint probabilities of the said emotion coexisting with other emotions in the instantaneous vector.

Hence, it is apparent that outside the co-occurrence framework analyzing mixed emotions is difficult, while analyzing isolated emotions leads to overestimation and biased results.

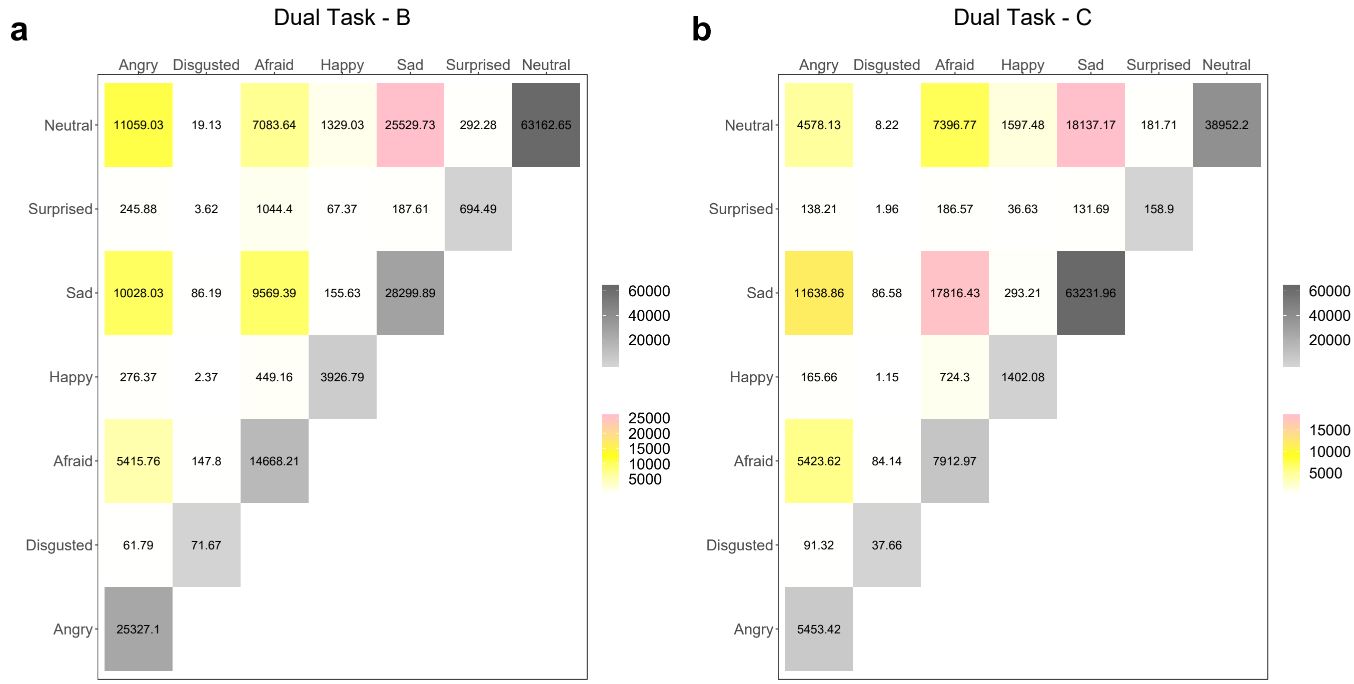

Figure 2: **Emotional co-occurrence matrices for the batch (B) and continual (C) email groups.** To facilitate visualization of the highly imbalanced data within each matrix, we employ two colormaps, one for the diagonals and one for the off-diagonal elements. In the diagonal elements, which represent the extent of single emotions, of interest is the inversion of the Neutral vs. Sad ratio between the two treatment groups. In the off-diagonal elements, which represent the extent of co-occurred emotions, of interest is the more extensive co-occurrence of Sad with Fear in the C group with respect to the B group.

The R code that implements the co-occurrence matrix method described in this paper resides in GitHub (Zaman, S., Tsiamyrtzis, P., & Pavlidis, I. CHI20-Displayed-Emotions-Methods. GitHub `https://github.com/UH-CPL/CHI20-Displayed-Emotions-Methods`).

RESULTS

In both the B and C groups, single (i.e., dominating) emotions are far more widespread with respect to mixed (i.e., co-occurred) emotions - see the diagonal vs. off-diagonal matrix cells in Fig. 2. Consequently, due to the different scales of the phenomena, our analysis will examine separately how the two cohorts differ in terms of dominating and co-occurred emotions, respectively.

Let us represent the dominating emotions (diagonal elements of co-occurrence matrices) with $\mathbf{Y}$. We are interested to test whether the factors “Emotions” (with 7 levels) and “Group” (with 2 levels) are significant in determining the response variable $\mathbf{Y}$. For this reason, we run a two-way ANOVA with interaction terms. To satisfy the standard ANOVA assumptions, we transform the response variable $\mathbf{Y}$ using the natural logarithm, i.e., $\ln(y_i+1)$ (we added the same small positive constant to all y_i to avoid singularities when $y_i = 0$).

The model results are shown in Table 1; it is evident that both the main effects “Emotions” and “Group” are highly significant as the respective p-values are both < 0.01. The interaction plot (Fig. 3) indicates that for very low values (i.e., Disgust, Surprised, and Happy) there is no group effect. For the high value items, there is an interaction in Sad and Neutral, where as we move from group B to group C these two emotions reverse the value they receive. This is the key result of this investigation, that is, participants become ‘sadder’ when they have to service emails every so often amidst a cognitive task. Interestingly, participants who are interrupted just once, having to service all the emails in a long session (i.e., batch), trend angrier (Fig. 3). Dropping completely the interaction term and running the ANOVA again will provide significance for both “Emotions” and “Group”, as Table 2 indicates.

Following the same procedure we adopted for testing differences of dominating emotions between the B and C groups, we also tested differences of co-occurrences of Sad with the other six emotions, taking the corresponding off-diagonal elements. We found that the mix of (Sad, Afraid) trends higher ($p = 0.09$) in the continual group with respect to the batch group. It appears that when participants are continually interrupted with emails, clearly become ‘sadder’ but also trend more fearful. This observational result is also supported psychometrically. Indeed, the NASA-TLX mean summative score in the C group was found to be significantly higher than the B group ($p = 0.03$), suggesting that participants who received the continual email treatment felt overloaded. In particular, participants in the C group reported significantly higher Mental Load ($p = 0.006$) and Effort ($p = 0.011$) compared to the B group. These reports are consistent with the displayed higher sadness in the C group.

Table 1: **ANOVA table with interaction term.** Levels of significance were set at $\alpha = 0.05$ designated by * or $\alpha = 0.01$ designated by ** or $\alpha = 0.001$ designated by ***.

	DOF	F-value	p-value
Emotions (E)	6	35.148	0.000 ***
Group (G)	1	8.575	0.004 **
E*G	6	0.622	0.712

Table 2: **ANOVA table without interaction term.** Levels of significance were set at $\alpha = 0.05$ designated by * or $\alpha = 0.01$ designated by ** or $\alpha = 0.001$ designated by ***.

	DOF	F-value	p-value
Emotions (E)	6	35.611	0.000 ***
Group (G)	1	8.688	0.004 **

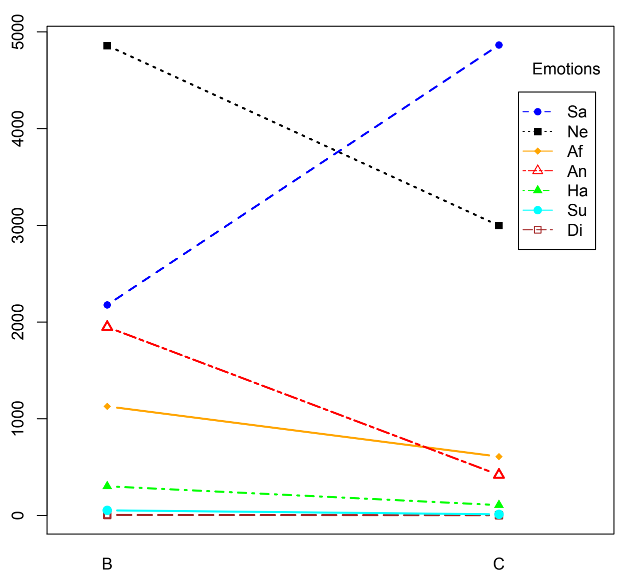

Figure 3: **Interaction plot for dominating emotions (i.e., diagonal elements of co-occurrence matrices).** Sa, Ne, Af, An, Ha, Su, Di stand for Sad, Neutral, Afraid, Angry, Happy, Surprised, Disgusted, respectively.

Figure 4a shows the evolving composition of emotional vectors over the DT timeline along with the co-occurrence matrix for participant T016. This participant exemplifies the summative pattern displayed in Fig. 2 for the entire Batch cohort. As it is evident from the time series of stacked plots, the participant underwent a bout of anger in the beginning of the experimental session. This gave way to a brief period of neutrality until the batch email period started. At that point a more intense bout of anger took hold, lasting half way through the batch email task. Thereafter, neutrality appears to dominate most of the time with a couple of brief exceptions, the most notable of which is the return of anger the last few minutes of the experimental session. This temporal information percolates into the co-occurrence matrix, resulting in high diagonal values for Angry and Neutral, the two emotions that alternate in dominating the participant's state. Facial snapshots of participants from the B group, complete with CNN annotation, are depicted in Fig. 5a.

Figure 4b shows the evolving composition of emotional vectors over the DT timeline along with the co-occurrence matrix for participant T064. This participant exemplifies the summative pattern displayed in Fig. 2 for the entire Continual cohort. As it is evident from the time series of stack plots, the participant undergoes semi-periodic fluctuations alternating between Sad dominance and co-occurrence of Sad, Afraid, and Angry emotions. The semi-periodicity of the emotional mix accompanies the semi-periodicity of the email interruptions. These temporal patterns of dominance and coexistence percolate into the co-occurrence matrix, resulting into high diagonal value for Sad and high off-diagonal values for (Sad, Afraid) and (Sad, Angry). Facial snapshots of participants from the C group, complete with CNN annotation, are depicted in Fig. 5b.

DISCUSSION

Our goal was to examine how emotions might be displayed when people multitask. In this paper, we report a novel method, based on co-occurrence matrices, for analyzing emotions from facial expressions. The method preserves the associative relationships between the probabilistic components of the CNN-derived emotional vectors, facilitating analysis at multiple levels, that is, dominance, absence, and co-occurrence of specific emotions. For this reason, we believe the method carries value and will have impact in HCI studies.

We used this method to analyze the emotional footprints of two groups of participants. One group was subjected to a batch email interruption, while the other group was subjected to continual email interruptions amidst an essay writing task. To minimize confounding factors, we carefully controlled the selection of participants and the precision of the experimental protocol.

The key result is that sadness appears to dominate participants who are interrupted frequently during an essay writing task. In contrast, participants who experience a single long interruption trend neutral for the most part. Frequent context switching may require intensification of cognitive function, which manifests facially via corrugator muscle action [31] - something that may contribute to the facial display of sadness.

Interestingly, sadness is occasionally accompanied with a touch of fear in those in the C condition – the frequently interrupted participants. This may have to do with the onset of anticipation as new (and unknown) items keep coming that have to be dealt with in real-time. The anticipation of interruptions could also be related to fear of not being able to finish the main task (essay). In an experimental study where people were continually interrupted, it was discovered that they worked faster, presumably because they knew that inter-

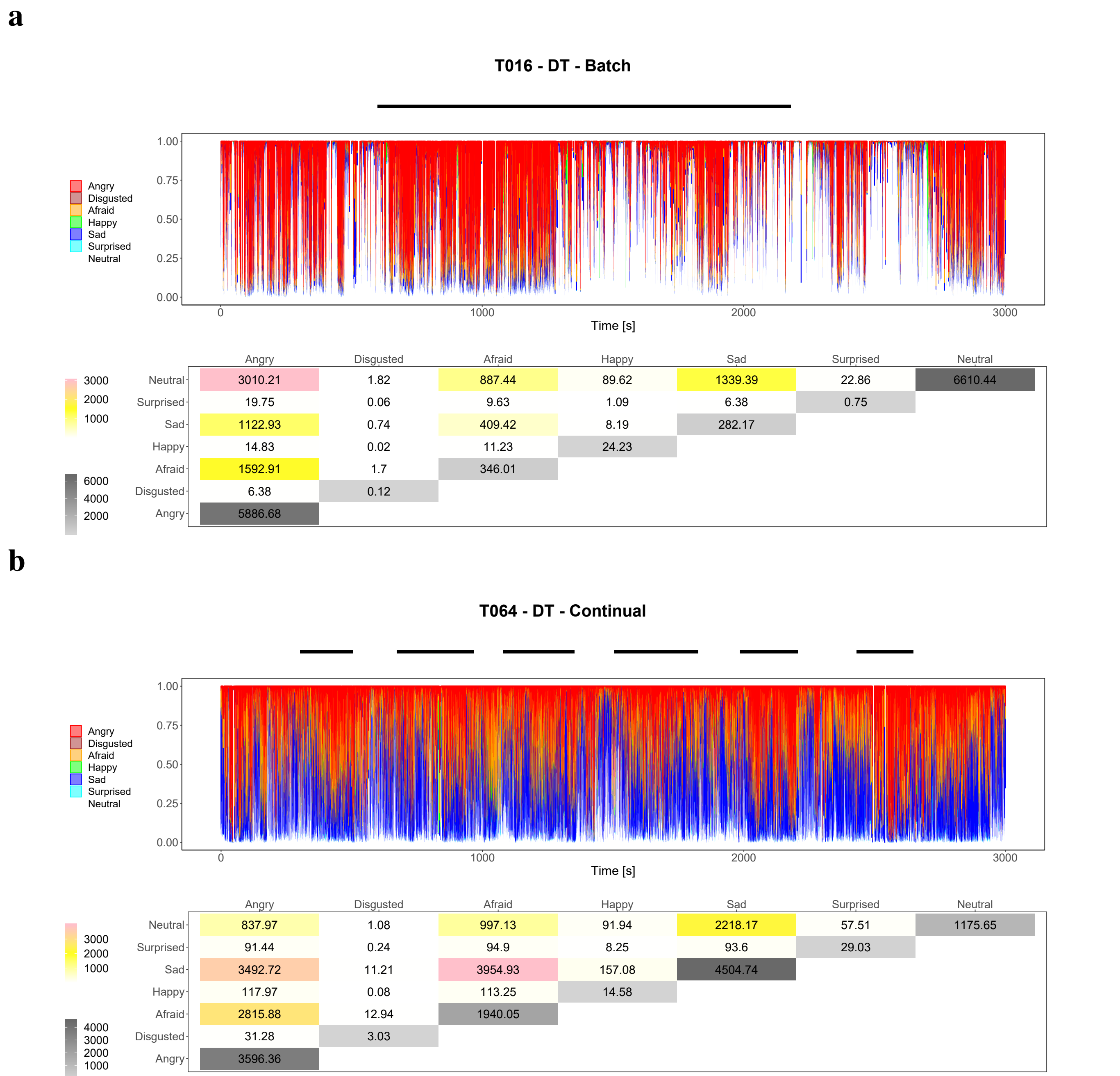

Figure 4: **Representative examples of batch and continual cases given through two pairs of stack plot/co-occurrence matrix.** Stack plot time series of the emotion probability vector shows the evolution of the participant's state during DT. In each momentary stack plot, probabilities of emotions are arranged from bottom to top in the order shown in the plot's color legend. The accompanying co-occurrence matrix presents the summative associative distribution of emotions for the entire DT session. There are different colormaps for the diagonal and off-diagonal elements of the matrix to facilitate visualization. **a. Batch example - participant T016.** The black bar at the top of the plot indicates the period the participant was responding to emails in batch mode. Neutral (white) and Angry (red) alternate in their domination of stack plots as time progresses. In the co-occurrence matrix, the high values in the Neutral and Angry diagonal cells summarize the evidence from the time series. **b. Continual example - participant T064.** The black bars at the top of the plot indicate the periods the participant interrupted his essay writing to respond to the arrived email item. Sad (blue) fluctuates in its dominance throughout DT in a semi-periodic fashion. When Sad subsides, Afraid (orange) in combination with Angry (red) appear to fill the void.

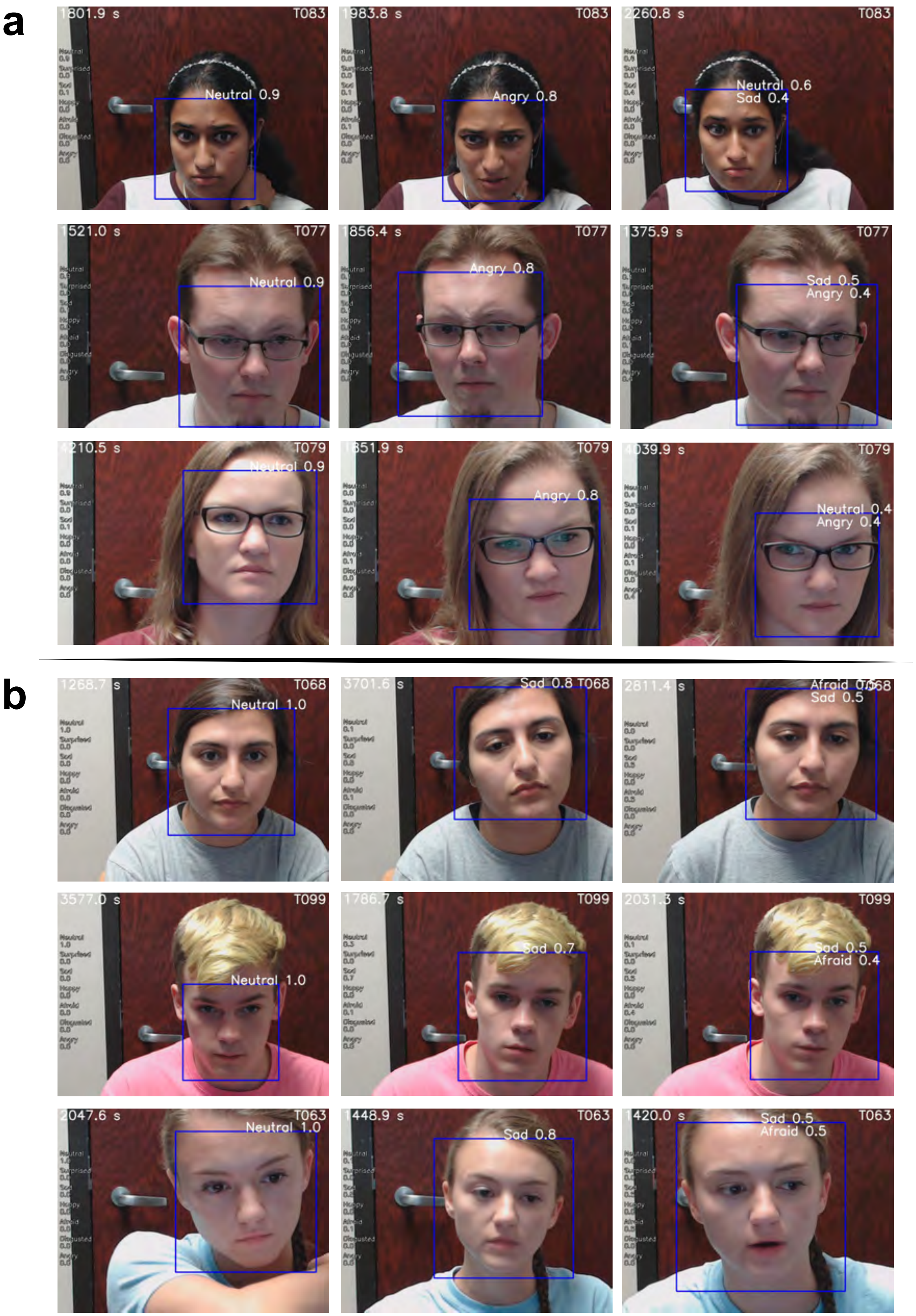

Figure 5: **Snapshots of participant faces taken during DT.** The images are annotated with the CNN output. Each row holds characteristic snapshots of a specific participant. **a. Batch (B) group examples.** The left column holds Neutral faces of participants. The middle column holds Angry faces of participants. The right column holds faces of reported emotional co-occurrences. **b. Continual (C) group examples.** The left column holds Neutral faces of participants. The middle column holds Sad faces of participants. The right column holds faces of reported (Sad, Afraid) co-occurrence.

ruptions detracted from the time needed to complete the main task [34]. Thus, the awareness of opportunity costs could lead to such an emotional expression.

Sadness and fear are known to trigger 'fight or flight responses' [15], which are probably low grade in the context of email interruptions. But in the context of continual interruptions and multitasking over longer periods, as is shown to occur in information work [13, 21], such responses can be cumulative. They are thus likely to be unhealthy in the long run [28], as they occur on a daily basis in people bent on servicing their incoming emails in real-time.

A secondary result is that although neutrality dominates the state of participants in the low multitasking condition (Batch), there is an element of anger. Participants may get angry when they realize the amount of work needed to process all the emails in one session, before returning to the main task. This could potentially be addressed if the email batch is processed at a later time, when responding to emails is the only consideration (i.e., single task). Unfortunately, due to office pressures, such neat arrangements are not always possible. At the very least, the present research makes people aware of hidden emotional effects that different email strategies could introduce.

A legitimate question is if the CNN algorithm we used provides accurate emotion recognition results. CNN algorithms represent the state of the art when it comes to quantification of displayed emotions. The performance of these algorithms depends on the datasets used for benchmarking; e.g. CNN algorithms tend to score between 60% and 70% in the FER dataset [10], while the same algorithms score above 90% in the CK+ dataset [41]. The reason is that the FER dataset is much more challenging than the CK+ dataset, as it features low resolution images and occluded faces. By design, our application's dataset is analogous to the CK+ dataset, featuring high resolution images and non-occluded faces, and so is the expected performance of the CNN algorithm.

Looking also at the macroscopic performance of the algorithm, it appears to behave soundly; for example, it correctly gives minuscule weight to Disgusted (Fig. 2), which is an emotion that by study design is out of the question in our experiment. Importantly, in extensive visual inspection of the CNN results, we found the algorithm to correctly annotate participants' facial frames in the overwhelming majority of cases. A small amount of this visual evidence is presented in Fig. 5 - the entire annotated facial dataset is publicly available on OSF [50].

Emotional Display, Multitasking, and Information Work

Our results, which showed that multitasking leads people to display negative emotions, particularly sadness, builds on work which found that multitasking is associated with felt stress [9, 34]. Our results suggest that not only do people experience stress with multitasking, but their faces may also express unpleasant emotions, and as discussed, emotional expression can have consequences in groups. Negative displayed emotions in particular can work against prosocial behaviors [20]. The results suggest that multitasking may not only affect the individual, but may have consequences for colleagues, teams, and the organization, particularly if multitasking is a common practice in the organization.

Our results suggest that organizations should pay more attention to the multitasking practices among individuals and teams in their own environments in order to examine ways to ameliorate it. As we discuss, emotional display can have contagion effects [7, 8, 43]. Multitasking in information work may have less consequences when people work in closed offices, but in open office settings, people may be more vulnerable to contagion effects. An example of workplace design is to introduce batching of email, as we tested, a strategy that could reduce attention shifts to email [35], yet this is also not free from problems since we discovered it to be associated with some emotional expression of anger. The ripple effects of emotional expression during multitasking in situ in the information workplace is a topic that warrants further study. Further research is also needed to examine how emotional expression is manifest with multitasking longitudinally.

Limitations

One interesting line of investigation in future studies would be the role of ethnicity/culture in emotional display during multitasking. In the present study, this factor could not be examined as all participants were born in, lived in, and studied in the U.S.

The relatively small number of participants ($n = 26$) is another limitation of the work reported here. However, one should not underestimate the significant amount of evidence present in the long observational horizon - 50 min per subject, totaling about 200,000 emotional vectors per treatment group. In this respect, we feel that the main result of this research, which is the dominance of sadness in continual email interruptions vs. neutrality in batch email interruptions, is likely to scale up in larger studies.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the National Science Foundation (NSF) through the medium collaborative award "Managing Stress in the Workplace: Unobtrusive Monitoring and Adaptive Interventions" (grants # 1704682, # 1704889, and # 1704636). Professor Pavlidis acknowledges additional support from the NSF REU award # 1659755 and the Greek Diaspora Fellowship Program (GDFP).

REFERENCES

[1] Fatema Akbar, Ayse Elvan Bayraktaroglu, Pradeep Buddharaju, Dennis Rodrigo Da Cunha Silva, Ge Gao, Ted Grover, Ricardo Gutierrez-Osuna, Nathan Cooper Jones, Gloria Mark, Ioannis Pavlidis, and et al. 2019. Email makes you sweat: Examining email interruptions and stress using thermal imaging. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 668, 14 pages.

[2] Ane Alberdi, Asier Aztiria, and Adrian Basarab. 2016. Towards an automatic early stress recognition system for office environments based on multimodal measurements:

A review. *Journal of Biomedical Informatics* 59 (2016), 49–75.

[3] Neal M Ashkanasy, Charmine EJ Härtel, and Catherine S Daus. 2002. Diversity and emotion: The new frontiers in organizational behavior research. *Journal of Management* 28, 3 (2002), 307–338.

[4] Tadas Baltrušaitis, Daniel McDuff, Ntombikayise Banda, Marwa Mahmoud, Rana El Kaliouby, Peter Robinson, and Rosalind Picard. 2011. Real-time inference of mental states from facial expressions and upper body gestures. In *Face and Gesture 2011*. IEEE, Piscataway, NJ, USA, 909–914.

[5] Stephen R Barley, Debra E Meyerson, and Stine Grodal. 2011. E-mail as a source and symbol of stress. *Organization Science* 22, 4 (2011), 887–906.

[6] Sigal Barsade and Olivia A O'Neill. 2016. Manage your emotional culture. *Harvard Business Review* 94, 1 (2016), 14.

[7] Sigal G Barsade. 2002. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly* 47, 4 (2002), 644–675.

[8] Sigal G Barsade, Constantinos GV Coutifaris, and Julianna Pillemer. 2018. Emotional contagion in organizational life. *Research in Organizational Behavior* 38 (2018), 137–151.

[9] Duncan P Brumby, Helena Du Toit, Harry J Griffin, Ana Tajadura-Jiménez, and Anna L Cox. 2014. Working with the television on: An investigation into media multitasking. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1807–1812.

[10] Pierre-Luc Carrier and Aaron Courville. 2013. Facial Expression Recognition (FER) Dataset. `https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data`. (2013).

[11] Ciprian Adrian Corneanu, Marc Oliu Simón, Jeffrey F Cohn, and Sergio Escalera Guerrero. 2016. Survey on RGB, 3D, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 8 (2016), 1548–1568.

[12] Jean Costa, Alexander T Adams, Malte F Jung, François Guimbretière, and Tanzeem Choudhury. 2016. EmotionCheck: Leveraging bodily signals and false feedback to regulate our emotions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 758–769.

[13] Mary Czerwinski, Mary Czerwinski, Eric Horvitz, and Susan Wilhite. 2004. A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 175–182.

[14] Laura A Dabbish and Robert E Kraut. 2006. Email overload at work: An analysis of factors associated with email strain. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*. Association for Computing Machinery, New York, NY, USA, 431–440.

[15] Richard J Davidson, Klaus R Sherer, and H Hill Goldsmith. 2009. *Handbook of Affective Sciences*. Oxford University Press, New York, NY, USA.

[16] Munmun De Choudhury and Scott Counts. 2013. Understanding affect in the workplace via social media. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. Association for Computing Machinery, New York, NY, USA, 303–316.

[17] Danyel Fisher, AJ Brush, Eric Gleave, and Marc A Smith. 2006. Revisiting Whittaker & Sidner's "email overload" ten years later. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*. Association for Computing Machinery, New York, NY, USA, 309–312.

[18] Nicholas Fitz, Kostadin Kushlev, Ranjan Jagannathan, Terrel Lewis, Devang Paliwal, and Dan Ariely. 2019. Batching smartphone notifications can improve well-being. *Computers in Human Behavior* 101 (2019), 84–94.

[19] Keaton A Fletcher, Sean M Potter, and Britany N Telford. 2018. Stress outcomes of four types of perceived interruptions. *Human Factors* 60, 2 (2018), 222–235.

[20] Jennifer M George. 1990. Personality, affect, and behavior in groups. *Journal of Applied Psychology* 75, 2 (1990), 107.

[21] Victor M González and Gloria Mark. 2004. Constant, constant, multi-tasking craziness: Managing multiple working spheres. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 113–120.

[22] Ian J Goodfellow, Dumitru Erhan, Pierre Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski, Yichuan Tang, David Thaler, Dong-Hyun Lee, and others. 2015. Challenges in representation learning: A report on three machine learning contests. *Neural Networks* 64 (2015), 59–63.

[23] The Radicati Group. 2015. Email Statistics Report 2015-2019. Developer's Online Log. (2015). Retrieved from `http://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf`.

[24] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, Peter A Hancock and Najmedin Meshkati (Eds.). North-Holland, Amsterdam, 139–183.

[25] Javier Hernandez, Rob R Morris, and Rosalind W Picard. 2011. Call center stress recognition with person-specific models. In *International Conference on Affective Computing and Intelligent Interaction*. Springer, Berlin, Germany, 125–134.

[26] Yun Huang, Ying Tang, and Yang Wang. 2015. Emotion map: A location-based mobile social system for improving emotion awareness and regulation. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Association for Computing Machinery, New York, NY, USA, 130–142.

[27] Shamsi T Iqbal and Eric Horvitz. 2007. Disruption and recovery of computing tasks: Field study, analysis, and directions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vol. 7. Association for Computing Machinery, New York, NY, USA, 677–686.

[28] Robert-Paul Juster, Bruce S McEwen, and Sonia J Lupien. 2010. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews* 35, 1 (2010), 2–16.

[29] Amy K Karlson, Shamsi T Iqbal, Brian Meyers, Gonzalo Ramos, Kathy Lee, and John C Tang. 2010. Mobile taskflow in context: A screenshot study of smartphone usage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2009–2018.

[30] Kostadin Kushlev and Elizabeth W Dunn. 2015. Checking email less frequently reduces stress. *Computers in Human Behavior* 43 (2015), 220–228.

[31] James A Levine, Ioannis T Pavlidis, Leslie MacBride, Zhen Zhu, and Panagiotis Tsiamyrtzis. 2009. Description and clinical studies of a device for the instantaneous detection of office-place stress. *Work* 34, 3 (2009), 359–364.

[32] David M Levy, Jacob O Wobbrock, Alfred W Kaszniak, and Marilyn Ostergren. 2012. The effects of mindfulness meditation training on multitasking in a high-stress information environment. In *Proceedings of Graphics Interface 2012*. Canadian Information Processing Society, CAN, 45–52.

[33] André Teixeira Lopes, Edilson de Aguiar, Alberto F De Souza, and Thiago Oliveira-Santos. 2017. Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order. *Pattern Recognition* 61 (2017), 610–628.

[34] Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. The cost of interrupted work: More speed and stress. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 107–110.

[35] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, Paul Johns, and Akane Sano. 2016a. Neurotics can't focus: An in situ study of online multitasking in the workplace. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1739–1744.

[36] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. 2016b. Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1717–1728.

[37] Gloria Mark, Stephen Voida, and Armand Cardello. 2012. A pace not dictated by electrons: An empirical study of work without email. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 555–564.

[38] Daniel McDuff, Amy Karlson, Ashish Kapoor, Asta Roseway, and Mary Czerwinski. 2012. AffectAura: An intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 849–858.

[39] Daniel McDuff, Abdelrahman Mahmoud, Mohammad Mavadati, May Amr, Jay Turcot, and Rana el Kaliouby. 2016. AFFDEX SDK: A cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3723–3726.

[40] Andre N Meyer, Laura E Barton, Gail C Murphy, Thomas Zimmermann, and Thomas Fritz. 2017. The work life of developers: Activities, switches and perceived productivity. *IEEE Transactions on Software Engineering* 43, 12 (2017), 1178–1193.

[41] Shervin Minaee and Amirali Abdolrashidi. 2019. Deep-Emotion: Facial expression recognition using attentional convolutional network. *arXiv e-prints* (02 2019), arXiv:1902.01019.

[42] Stephan J Motowidlo, John S Packard, and Michael R Manning. 1986. Occupational stress: Its causes and consequences for job performance. *Journal of Applied Psychology* 71, 4 (1986), 618.

[43] Roland Neumann and Fritz Strack. 2000. "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology* 79, 2 (2000), 211.

[44] Michael R Parke and Myeong-Gu Seo. 2017. The role of affect climate in organizational effectiveness. *Academy of Management Review* 42, 2 (2017), 334–360.

[45] Karen Renaud, Judith Ramsay, and Mario Hair. 2006. "You've got e-mail!"... shall I deal with it now? Electronic mail from the recipient's perspective. *International Journal of Human-Computer Interaction* 21, 3 (2006), 313–332.

[46] Sefik Ilkin Serengil. 2018. Facial Expression Recognition with Keras. Developer's Online Log. (1 January 2018). Retrieved September 10, 2019 from https://sefiks.com/2018/01/01/facial-expression-recognition-with-keras/.

[47] Barry M Staw, Robert I Sutton, and Lisa H Pelled. 1994. Employee positive emotion and favorable outcomes at the workplace. *Organization Science* 5, 1 (1994), 51–71.

[48] Chiew Seng Sean Tan, Johannes Schöning, Kris Luyten, and Karin Coninx. 2013. Informing intelligent user interfaces by inferring affective states from body postures in ubiquitous computing environments. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*. Association for Computing Machinery, New York, NY, USA, 235–246.

[49] Joseph A Veech. 2013. A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* 22, 2 (2013), 252–260.

[50] Amanveer Wesley, Shaila Zaman, Christopher Blank, and Ioannis Pavlidis. 2020. Displayed Emotions Dataset on Dual Task. *Open Science Framework* https://doi.org/10.17605/osf.io/mhdgt. (2020).

[51] Steve Whittaker and Candace Sidner. 1997. Email overload: Exploring personal information management of email. In *Culture of the Internet*, Sarah Kiesler (Ed.). Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 277–295.

[52] Robert B Zajonc, Sheila T Murphy, and Marita Inglehart. 1989. Feeling and facial efference: Implications of the vascular theory of emotion. *Psychological Review* 96, 3 (1989), 395.

[53] Shaila Zaman, Amanveer Wesley, Dennis Rodrigo Da Cunha Silva, Pradeep Buddharaju, Fatema Akbar, Ge Gao, Gloria Mark, Ricardo Gutierrez-Osuna, and Ioannis Pavlidis. 2019. Stress and productivity patterns of interrupted, synergistic, and antagonistic office activities. *Scientific Data* 6, 1 (2019), 1–18. DOI: http://dx.doi.org/10.1038/s41597-019-0249-5