Facial Self-Touches Are Associated With Mental Stress in Knowledge Work

Fettah Kiran

Department of Computer Science University of Houston Houston, TX, USA

Gayatri Bhatambarekar

Virginia Tech Transportation Institute
Virginia Tech
Blacksburg, VA, USA

MD Tanim Hasan

Department of Computer Science University of Houston Houston, TX, USA

Abhijit Sarkar

Virginia Tech Transportation Institute
Virginia Tech
Blacksburg, VA, USA

Kaushik Ganeshan Virginia Tech Transportation Institute

Virginia Tech

Blacksburg, VA, USA

Ioannis Pavlidis

Department of Computer Science University of Houston Houston, TX, USA

Abstract—We investigate the affective role of spontaneous facial self-touch (sFST) during cognitive tasks. For that, we analyze the open desk-bound knowledge work (DKW) dataset. The dataset was collected through a naturalistic study of 10 academic researchers while working in their university office for four days. It includes behavioral, situational, and dispositional data from \sim 170 hours of continuous recordings. Using a MobileNet CNN, we labeled facial video frames with sFST information. Then we used these categorical variables along with other contextual predictors to explain the sympathetic overactivity of the participants. Sympathetic responses were quantified through facial electrodermal activity (fEDA). The results of the relevant multiple regression model suggest that there is a strong positive association between the frequency of chin-cheek-nose sFST and sympathetic overactivity. Although it has been ignored in the affective computing literature, lower-face self-touch is a solid indicator of sympathetic overactivity, which is a proxy of mental

Index Terms—spontaneous self-touch of the face, sFST, stress, arousal, knowledge work, naturalistic study, affective computing

I. INTRODUCTION

Spontaneous touching of one's own face (sFST) is an ubiquitous behavior that is associated with cognitively and emotionally challenging situations. Self-touch facial gestures have a self-soothing effect [1] with evolutionary roots in primates [2]. They typically involve the non-dominant hand [3]. Recent research demonstrated that sFST also plays a role in improving memory performance. Spille et al. experimented with a haptic working memory task in n=49 participants [4]. They found an underlying grouping of high-touching individuals (HT) versus low-touching individuals (LT), which mirrored a grouping in certain personality traits. Importantly, they found that suppression of sFST in HT individuals had a negative correlation with memory performance. They also showed through EEG analysis that the most significant neurophysiological changes occur when the hand moves toward the face, before any skin contact has occurred [5].

This work was partly funded by the National Science Foundation (NSF).

The sFST feature that attracted the most research attention was frequency. The frequency of self-touch has been shown to be influenced by negative affect and attention distraction and may be involved in the regulation of emotions and working memory functions. Mueller et al. performed a deeper examination of the temporal patterns of sFST through accelerometric and electromyographic (EMG) data [6]. They found that not only frequency, but also point of touch and duration of contact are influenced by cognitive and emotional demands.

Some researchers also consider age to be a potential factor. Kyra et al. performed an experiment in 10 pre-adolescent girls and 10 post-adolescent girls in low and high stress settings. They found three forms of self-touch in post-adolescent girls, two of self-regulatory nature and one of conversational nature. In contrast, no distinct self-touch patterns were found in pre-adolescent girls. The latter suggests that sFST behaviors fully form only after a certain stage of brain development [7].

Grunwald et al. showed that skin areas of the face with a higher number of vellus hairs (and presumably higher innervation density) are touched particularly frequently during sFST. The chin is the most prominent among these facial areas. In fact, the chin, nose, and forehead form the so-called T-zone of the face, which is touched up to 800 times per day [8].

Parallel to the aforementioned psychological studies, there have been efforts in the multimodal and computer vision (CV) communities to develop machine learning (ML) methods for automated recognition of facial self-touch gestures. Bai et al. developed one such method based on signals from smartwatch sensor signals [9]. Beyam et al. trained and tested a Convolutional Neural Network (CNN) on 2 million video frames from small group conversations that were labeled *face-touch* and *no-face-touch* by 16 annotators (Cohen's Kappa=0.89). CNN achieved 83.76% F1 score [10]. Ibrahim et al. introduced a deep learning framework, called FaceTouch, which detects hand-to-face touches in the wild, including video chat channels and bus footage [11]. Other researchers constructed social interaction datasets through laboratory experiments that include much more than facial touches. The central goal of

the said efforts was to support the design and evaluation of ML algorithms that recognize bodily behaviors [12]–[14].

In summary, previous work on sFST focused on small experimental studies by psychologists [4], [5] or the development of ML classification methods on images often captured outside of real-world situations [10]. We believe that sFST is a complex affective behavior that is best examined in a naturalistic study with appropriate models that draw on multiple channels of information for context. Our work does exactly this, thus filling a gap in the literature. In more detail, we chose to analyze sFST in the context of an open multimodal dataset from a naturalistic study on desk-bound knowledge work by academic researchers [15]. The dataset will be referred to as DKW. Using behavioral, situational, and dispositional channels, the authors of the dataset reported that sympathetic activation of the participants was strongly associated with the amount of reading and writing they did, the extent of use of smartphones, and the frequency of physical breaks they took. An information channel that was not used in that study was the sFST channel. Since facial videos of the participants were recorded throughout the study, sFST extraction is possible. Hence, we developed a method to do just that and remodel the stress levels of participants by adding sFST as an extra predictor.

Our research addresses the following question: **RQ:** Are facial self-touches associated with stress in knowledge work? By stress here we refer to mental stress as opposed to physical stress that occurs during purposeful physical activity. The results of the study provide valuable information on sFST behaviors during knowledge work, a first in the literature. Importantly, our work introduces a new channel to measure mental stress. In contrast to conventional stress channels such as EDA, the sFST channel not only tracks stress but also quantifies a countermeasure to manage it, due to the known soothing effect of sFST [8].

In the remainder of the paper, in Section II we summarize the DKW study [15] that produced the open dataset that we use. In Section III, we detail the ML method that labels the facial frames of participants with sFST information. In Section IV, we present the results of the data analysis, highlighting the association of sFST with mental stress. The paper ends with a discussion and conclusion section.

II. DATASET DESCRIPTION

We use the open dataset DKW associated with the study of knowledge work reported by Hasan et al. [15]. The study procedures were approved by the relevant Institutional Review Board (IRB) of the University of Houston. The dataset includes data from 10 academic researchers (6 men / 4 women) who were monitored in their university office for four days (D1, D2, D3, D4). The state and activities of the participants were captured through the following sensors: a) a facial thermal camera (FLIR Tau 640) to extract perinasal perspiration signals; These signals constitute a facial EDA measure that tracks sympathetic activation [16]; b) a visual facial camera (Logitech HD Pro - C920) to analyze displayed emotions; and

c) a visual ceiling camera (Logitech Brio) to assist in the classification of the activities of the participants (e.g., playing with the smartphone). Participants had to complete certain biographic and psychometric trait questionnaires. They also had to fill out psychometric state questionnaires every morning and night. The dataset features 602737 rows of multimodal information - one for every second, which amounts to nearly 170 hours of observation.

The only controlled process in the study was the daily baseline session. Before starting work every morning, the participants had to relax for four minutes in their chair by imagining a landscape of nature, while the thermal and visual cameras were recording.

A. Description of Variables

Hasan et al. investigated how behavioral, situational, and dispositional factors relate to sympathetic overactivity in desk-bound knowledge work, producing the DKW dataset [15]. Prolonged sympathetic overactivity is not desirable because it is associated with performance degradation [17] and can undermine wellness [18]. Hence, identifying the signature signs of such overactivity provides useful behavioral insights and can inform future interventions. However, Hasan et al. left out sFST, a potentially valuable behavioral factor. In this paper, we introduce this factor and reanalyze the DKW dataset. The following is a description of all the variables we use in the expanded multiple regression model.

Response Variable - Proxy for Sympathetic Activation

In the DKW dataset, sympathetic activation is quantified through perinasal perspiration signals extracted from thermal facial videos according to the method reported by Shastri et al. [16]. Perinansal perspiration (PP), also known as facial electrodermal activity (fEDA), has been shown to correlate with palmar EDA [16]. To reduce interindividual variability, the PP signals are normalized by subtracting their corresponding mean baseline signals PP_{BL} . The latter were collected daily for each participant. Hence, the normalized sympathetic activation of participant P_i at time t of day D_j , complete with logarithmic correction to preserve normality, is:

$$\Delta P P_{ij}(t) = \ln P P_{ij}(t) - \overline{\ln P P}_{BL_{ij}}.$$
 (1)

Behavioral Factors

1) Observed sFST Gestures: In the enhanced DKW dataset that we analyze in this paper, we added labels for sFST gestures as follows:

Che: Participants touched the cheek.

CheChi: Participants touched the cheek and chin. CheChiN: Participants touched cheek, chin, & nose.

Chi: Participants touched the chin.

Others: Participants touched other parts of the face. Based on these labels, the dataset calculates the relative frequency F_{Che} , F_{CheChi} , $F_{CheChiN}$, F_{Chi} , F_{Others} a participant spends each day in Che, CheChiN, Chi, and Others sFST gestures, respectively.

- 2) Observed Activities: The DKW dataset has labels for each activity the participants performed every second of the observation period. The labeled activities included the core work functions of a researcher, such as reading and writing, and many others. The list of activity labels is as follows.
 - RW: The continuum of reading and writing activities, which represented the knowledge tasks of the participants.
 - SA: Secondary activities, which included eating or listening to music while working or doing something else.
 - SP: Smartphone activities, where participants used their phones for texting, apps, and other reasons.
 - *I*: Participants had interactions with conversational partners physically in the office or virtually.
- Out: Participants walked out of the office taking a break. Based on these labels, the dataset calculates the percentage of time T_{RW} , T_{SA} , T_{SP} , T_{I} , T_{Out} a participant spends each day in RW, SA, SP, I, and Out activities, respectively. The dataset also calculates the daily frequency f_{out} and the mean duration \bar{t}_{out} of breaks per participant.
- 3) Observed Displayed Emotions: The DKW has valence labels that correspond to the emotions displayed on the faces of the participants. The emotion vectors $\overrightarrow{DE}_{i,t}$ of participants are labeled as follows:

$$\begin{split} \mathcal{L}(\overrightarrow{DE}_{i,t}) &= \\ \begin{cases} DE_N, & \text{Neutral} \\ DE_S, & \text{Sad} \\ DE_-, & \text{Surprised} + \text{Afraid} + \text{Disgusted} + \text{Angry} \\ DE_+, & \text{Happy} \end{cases} \end{split}$$

 DE_N indicates a largely neutral facial display. DE_S indicates a facial display dominated by sadness, which in the context of knowledge work is associated with a sober look that people assume when thinking, due to the activation of the corrugator muscle [19], [20]. DE_- indicates a facial display dominated by strong negative emotions, including fear, anger, disgust, and surprise. Surprise is included in the negative emotion category because it is rarely positive before deadlines, which participants were working toward. In any case, it has very low frequency in the dataset and thus is likely inconsequential in the context of statistical modeling. DE_+ indicates a facial display dominated by happiness.

Situational Factors

The DKW dataset has morning (SA_M) and evening (SA_E) scores of the State and Trait Anxiety Inventory (STAI) Form Y1 [21] for each participant. It also has NASA Task Load Index (NASA-TLX) scores, collected from each participant upon leaving the office for the day. These scores correspond to the NASA-TLX six subscales: Mental Demand N_{MD} , Physical Demand N_{PD} , Temporal Demand N_{TD} , Perceived Performance N_P , Effort N_E , and Frustration N_F [22].

Dispositional Factors

The DKW dataset has biographic information on the sex S of the participants, with levels S_M = Men and S_F = Women. It also has information about the academic rank of the participants, with levels R_1 = Doctoral Student, R_2 = Postdoctoral Researcher, R_3 = Junior Faculty, and R_4 = Senior Faculty. Finally, the DKW dataset has scores from the STAI Form Y-2 questionnaire that measures anxiety predisposition TA [21].

III. METHOD FOR ASSIGNING SFST LABELS TO VIDEO FRAMES

We processed each frame of the DKW facial videos using a MobileNet-based CNN architecture [23]–[25] to detect 2D keypoints for the face and hands. These anatomical landmarks were grouped into seven polygonal facial regions: left / right cheeks, eyes, forehead, nose, and chin (see Fig. 1a–c). The hand region was similarly represented as a polygon.

A facial self-touch (sFST) event was recorded when the hand polygon intersected any facial region in the image space. This geometric overlap was used as a proxy for 3D contact or close proximity. Although this simplification is common in image-based behavioral detection, we validated its accuracy through manual inspection of 2,000 randomly sampled frames (200 per participant), where 94% classification accuracy was achieved (Section IV.A.1). Only one instance involved proximity without actual touch, indicating high fidelity.

Each sFST frame was labeled by the region touched — e.g., *Che*, *Chi*, *CheChi*, *CheChiN*, or *Others* — and these labels were aggregated at the daily level to compute relative frequencies for modeling (see coding schema in Section II.A.1). These annotations provided a new temporally aligned behavioral channel to investigate mental stress during knowledge work.

IV. RESULTS

A. Exploratory Analysis

1) Exploratory Analysis of sFST: Figure 2 shows the annotations of sFST performed by MobileNet CNN on a randomly chosen sample of frames from various participants on various days. To assess the accuracy of the sFST labels provided by the MobileNet CNN, we randomly selected a set of 2000 frames (200 frames from each participant). Two members of our team inspected the sample frames and rated the accuracy of CNN classification as 94%. This is an excellent classification rate that can reliably support subsequent statistical modeling.

Figure 3 shows the daily relative frequency of various types of sFST for participant T001. It is evident that the participant (who was right-handed) self-touches his left cheek much more than he touches his right cheek. This bias of sFST towards the non-dominant hand has been documented in the psychological literature [3], offering additional assurances for the goodness of our classification method.

Figure 4 shows the daily relative frequencies of the sFST of the participants as stack plots. To reduce the number of sFST levels, for effective regression modeling,

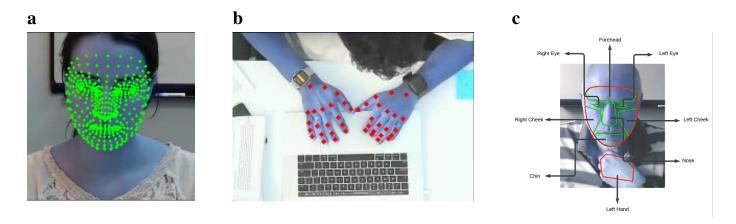


Fig. 1: Visualization of landmarks. [a] Detected facial keypoints. [b] Detected hand keypoints. [c] Seven key face regions $R_{\rm F}$ delineated by our method; they include left eye, forehead, right eye, left cheek, nose, right cheek, and chin. The method also delineated a hand region $R_{\rm H}$; it corresponds to the left hand.

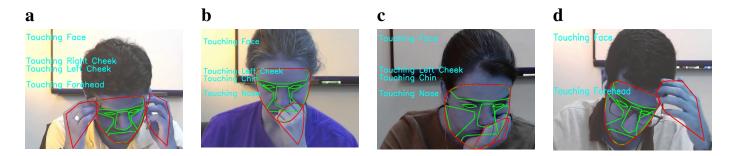


Fig. 2: Examples of sFST from the DKW dataset highlighting the diversity of face-hand interactions. In most cases, participants touch their chin, cheeks, and forehead. The annotations in cyan text are provided by the MobileNet CNN and are correct. The snapshots were randomly chosen from the following recordings: [a] Participant T017 in Day 2. [b] Participant T013 in Day 1. [c] Participant T007 in Day 2. [d] Participant T015 in Day 2.

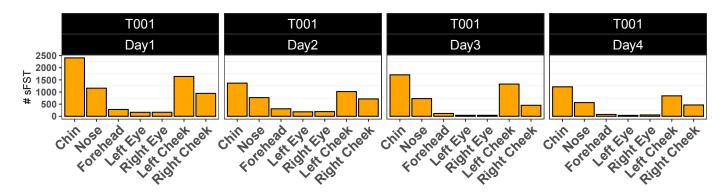


Fig. 3: Distribution of frequencies of sFST for participant T001 during the four days of observation. In all four days, the frequency of touching the left cheek is noticeably bigger than the relative frequency of touching the right cheek. The pattern is representative of all study participants. The said pattern is in agreement with prior reports in the psychological literature, suggesting that sFST are acted predominantly by the non-dominant hand [3]. Consequently, this result serves as a validity check for the goodness of the data and the classification we applied.

we consolidated the left and right sFST. For example, the self-touch of the left cheek and chin LCheChi and the right cheek and chin RCheChi consolidated into $CheChi = LCheChi \cup RCheChi$. Naturally, most of the

time the participants did not touch their face while doing knowledge work. However, some were self-touching a lot more than others throughout the four-day period. Specifically, participants T017 and T019 self-touched their face quite

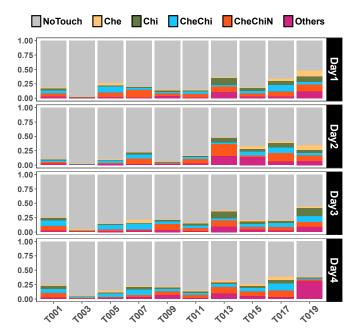


Fig. 4: Stack plots of relative frequency sFST data for each participant and each day. Chi stands for chin self-touch, CheChiN stands for cheek-chin-nose self-touch, Che stands for cheek self-touch, CheChi stands for cheek-chin self-touch, and Others stands for all other observed self-touch combinations. NoTouch stands for the absence of facial self-touch, which is what participants did most of the time.

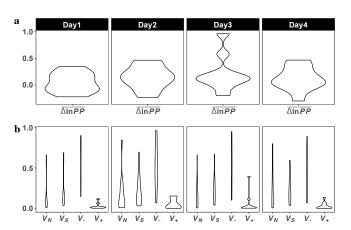


Fig. 5: Descriptive plots of daily sympathetic activation. a. Violin plots of participants' mean log-corrected and normalized perinasal perspiration measurements for each day of the study. The abundance of positive values in these boxplots suggests widespread sympathetic overactivation for the participants across all days b. Violin plots of participants' mean valence probabilities for each day of the study. V_N stands for Neutral facial display. V_S stands for Sad facial display. V_L stands for display of negative emotions as the union of Angry \cup Afraid \cup Surprised \cup Disgusted. This mix of negative emotional displays is dominant in all four days of the study. V_+ stands for Happy facial display.

frequently while participant T003 self-touched her face very sparingly. The two former are high self-touch individuals

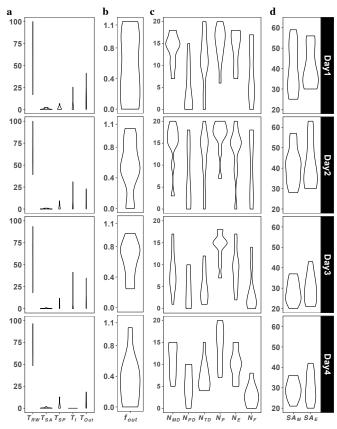


Fig. 6: Descriptive plots of key model predictors. a. Daily violin plots of participants' mean percent time devoted to reading/writing T_{RW} , secondary activities like eating and working T_{SA} , smartphone use T_{SP} , conversations with others T_{I} , and physical breaks away from the desk T_{Out} . b. Daily violin plots of participants' physical break frequency f_{out} . c. Daily violin plots of participants' NASA-TLX subscale scores. d. Daily violin plots of participants' morning SA_{II} and evening SA_{II} anxiety.

(HT), while the latter is a low self-touch individual (LT). Such cases have been reported in the psychological literature [5]. Hence, we encounter the full spectrum of self-touch profiles in the DKW dataset, which is a sign of representative sampling.

2) Exploratory Analysis of Other Variables: Figure 5a shows daily violin plots of log-corrected and normalized perinasal perspiration values for the DKW dataset. A significant portion of the values are positive ($\overline{\Delta \ln PP} = 0.1 \pm 0.2$), indicating the widespread presence of sympathetic overactivation in the participants throughout the monitoring period. Figure 5b shows daily violin plots of the probabilities of emotions displayed facially for the DKW dataset. The mix of negative emotions stands out with the overall probability $V_- = 0.5 \pm 0.3$, reflecting the challenging nature of continuous cognitive work, and is consistent with the sympathetic overactivation manifested in measurements of perinasal perspiration. Positive emotions are scarcely displayed, having a probability $V_+ = 0.04 \pm 0.07$. Neutral expressions and sadness act

as a counterweight to the negative mix, as they are almost as prevalent when considered together ($V_N = 0.2 \pm 0.2$ and $V_S = 0.2 \pm 0.2$). Sadness here does not appear to be felt sadness, but rather the sober look people assume when thinking hard about something, due to autonomic activation of the corrugator muscle [19], [20].

Figure 6 shows descriptive plots of key study variables at the day level, reflecting values used in our modeling process. In more detail, Fig. 6a shows the boxplots of relative times of the activities the researchers were engaged in. Across days, the distribution of relative time devoted to reading/writing T_{RW} (on average 67.7 ± 21.7 %) far outweighs the relative time distributions of all other activities, which on average are as follows: $T_{SA} = 0.3 \pm 06$ % for secondary activities like eating while working, $T_{SP}=3.0\pm3.9~\%$ for smartphone use; $T_I=$ 2.9 ± 9.0 % for interactions with other people; and $T_{Out} =$ 10.2 ± 10.5 % for physical breaks, away from the desk. These numbers confirm the cognitive nature of the daily work of the participants. Furthermore, Fig. 6b shows the box plots of the frequency of physical breaks, which on average is 0.5 ± 0.4 per hour, that is, researchers go away from their desk about every two hours.

Figure 6c shows the score boxplots of the six NASA-TLX subscales. On average, physical demand scores are quite low $(N_{PD} = 5.3 \pm 5.2)$, as desk-bound research is a sedentary activity. The frustration scores are also low (on average $N_F =$ 6.2 ± 6.1), suggesting the absence of strong negative affect. Temporal demand is moderate (on average, $N_{TD} = 9.0 \pm 5.6$), indicating a significant, but not overwhelming presence of time pressure. Mental demand and effort tend to be high, which is in line with the demanding nature of research work; on average, $N_{MD}=12\pm4.9$ and $N_E=11.6\pm4.9$. Performance scores tend to have the highest values among all subscales (on average $N_P = 14.9 \pm 4.0$, indicating that researchers felt their hard work paid off. Figure 6d shows the boxplots of the morning and evening anxiety scores. On average, the scores are moderate, being located close to the middle of the normal range ($SA_M = 34.6 \pm 10.2$ for morning anxiety and $SA_E =$ 35.8 ± 10.7 for evening anxiety).

B. Multiple Regression Model and Results

We constructed an enhanced multiple linear regression model (MLR), which, unlike the model reported by Hasan et al. [15], includes predictors that manifest sFST behaviors. In more detail, the model response variable is the mean sympathetic activation of the participant $P_i \equiv i$ on day $D_j \equiv j$ while working in the office. The momentary sympathetic activation of participants is proxied by the log-corrected and normalized measurements $\Delta \ln PP(t)$ of their facial EDA see Eq. (1).

The model predictors include the behavioral, situational, and dispositional factors described in Section II-A. Because the relative times of activities add to 100%, we drop one factor to avoid cross-correlations; we chose this factor to be the relative time of physical breaks T_{Out} . Similarly, because the probabilities of facially displayed emotions add up to 1,

we drop the neutral expression probability factor DE_N . For the remaining factors, we compute the variance-inflation factor (VIF) to examine if there are any strong collinearities among them. As a result of this examination, we remove from the model the factors N_E , N_{TD} , N_F , and DE_+ , because we find that they correlate strongly with other factors. As our study features a repeat measures design, we take into account participant-centered random effects - see term $(1|P_i)$ in the models. The full model is shown in Eq. (2).

$$\overline{\Delta \ln PP}(i,j) \sim \\
\beta_0 + \beta_1 T_{RW} + \beta_2 T_{SA} + \beta_3 T_{SP} + \beta_4 T_I + \\
\beta_5 f_{out} + \beta_6 \overline{t}_{out} + \beta_7 DE_S + \beta_8 DE_- + \\
\beta_9 D + \beta_{10} T_D + \beta_{11} SA_M + \beta_{12} SA_E + \\
\beta_{13} N_{MD} + \beta_{14} N_{PD} + \beta_{15} N_P + \\
\beta_{16} G + \beta_{17} R + \beta_{18} TA + \\
\beta_{19} F_{Che} + \beta_{20} F_{CheChi} + \beta_{21} F_{CheChiN} + \\
\beta_{22} F_{Chi} + \beta_{23} F_{Others} + (1|P_i).$$
(2)

The first two lines of Eq. (2) have the daily behavioral characteristics of the participants, including the relative time they devote to various types of activities $(T_{RW}, T_{SA} T_{SP})$ and T_I , the frequency and mean length of their breaks (f_{out}) and t_{out} , and their facial emotions (DE_S) and DE_L . The next two lines of Eq. (2) have the participants' daily situational characteristics, including the day of observation D (D_I is reference), the daily time T_D spent in the office, the anxiety of the participants in the morning and evening (SA_M) and SA_E , and their perceived workload (N_{MD}, N_{PD}) , and N_P).

The next line of Eq. (2) has the dispositional characteristics of the participants, including their sex G (G_F is taken as a reference), academic rank R (R_1 is taken as a reference), and trait anxiety (TA). The last two lines of Eq. (2) hold the relative frequencies of the participants in the sFST modes.

Subsequently, we run a backward elimination process, based on the Akaike Information Criterion (AIC), arriving at the reduced model shown in Eq. (3).

$$\overline{\Delta \ln PP}(i,j) \sim \beta_0' + \beta_1' T_{RW} + \beta_2' T_{SP} + \beta_3' f_{out} + \beta_4' N_{PD} + \beta_5' F_{CheChiN} + (1|P_i).$$
(3)

The AIC of the reduced model in Eq. (3) is AIC = 27.874. This is a radical improvement over the full model in Eq. (2) whose AIC = 187.5. The summary results of the reduced model are presented in Table I and detailed graphs are shown in Fig. 7. Keeping in mind that sympathetic activation in the context of knowledge work is tantamount to mental stress, the results show that the relative frequency of touching the cheek, chin and nose $F_{CheChiN}$ is positively correlated with sympathetic activation, which answers the question that motivated this work. Furthermore, the results show that the frequency of physical breaks f_{out} , the relative time spent doing activities related to knowledge T_{RW} , and the use of smartphones T_{SP} are also positively correlated with the sympathetic activation of the researchers.

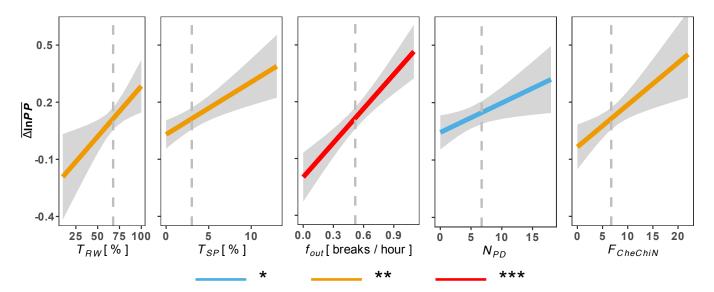


Fig. 7: Main effects of the reduced mixed-effects model (Eq. (3)) for sympathetic activation. Shown are the quantitative associations of sympathetic activation $\overline{\Delta \ln PP}$ with read/write relative time T_{RW} , smartphone use relative time T_{SP} , frequency of physical breaks f_{out} , perceived physical demand N_{PD} , and relative frequency of cheek-chin-nose self-touch $F_{CheChiN}$. Significance levels have been set as follows: *: p < 0.05, **: p < 0.01, ***: p < 0.001.

TABLE I: Results for the sympathetic activation predictors featured in the reduced model shown in Eq. (3). The model's AIC = 27.874 with n = 40 observations (4 days × 10 participants). Levels of significance: *: p < 0.05, **: p < 0.01, ***: p < 0.001.

Coefficient	Estimate	Standard Error	Degrees of Freedom	t-value	Pr (> t)
β'_0 for Intercept	-0.858	0.187	34.000	-4.579	< 0.001***
β_1^{γ} for T_{RW}	0.005	0.002	34.000	2.827	0.008**
β_2^{\dagger} for T_{SP}	0.028	0.008	34.000	3.590	0.001**
β_3^7 for f_{out}	0.603	0.110	34.000	5.491	< 0.001***
β_4^{\prime} for N_{PD}	0.015	0.006	34.000	2.399	0.022*
β_5^{\dagger} for $F_{CheChiN}$	0.022	0.007	34.000	3.074	0.004**

When optimizing the model using the elastic net methodology instead of the backward elimination process, two variables survive: the frequency of physical breaks f_{out} and the frequency of self-touch on the cheek-chin-nose $F_{CheChiN}$. Hence, irrespective of the optimization method, sFST remains a significant predictor of mental stress in the DKW dataset. Moreover, the model of Eq. 3 has $R^2=0.52$. This is reduced to $R^2=0.41$, if we remove $F_{CheChiN}$. Hence, the mental stress model loses about 20% of its explainability if we do not consider sFST. This finding becomes even more important if we consider the absence of emotional facial expressions as a significant predictor of mental stress.

V. DISCUSSION

We introduced a new set of predictors in a multiple linear regression model that explains the sympathetic activation of academic researchers included in the KWD dataset. These new predictors are instantaneous facial self-touch (sFST) labels extracted through a MobileNet CNN algorithm. To the best of our knowledge, this is the first time such behavioral predictors have been used in affective computing research. The results of the modeling suggest that the relative frequency of simultaneously touching the cheek, chin, and nose (Fig. 2b) is positively associated with sympathetic activation, that is,

the physiological manifestation of mental stress in knowledge work.

Regarding all other predictors, our results mirror those reported by Hasan et al. [15]. In more detail, we found that reading and writing, the use of smartphones, and the frequency of physical breaks are associated with sympathetic activation. The result of smartphone use complements recent reports in the literature that associate smartphone use with stress levels [26]. The physical break result likely points to a coping mechanism for sympathetic overactivity in desk-bound work. The significance of both sFST and physical breaks in the model suggests that these predictors may act synergistically to manage elevated stress during cognitive work, opening a promising line of future affective research.

VI. CONCLUSION

Our findings suggest that in the context of solitary knowledge work, facial expressions offer limited predictive value, whereas facial self-touch behaviors, often overlooked in affective computing, are strongly related to mental stress. This points to sFST as a potentially valuable observational channel for affective monitoring systems in knowledge-based work environments.

ETHICAL IMPACT STATEMENT

Privacy and protocol adherence. We use the open data set associated with the article by Hasan et al., CHI 2023 [15]. The facial self-touch predictor that we introduced involves facial imagery, which inherently affects participant privacy. In addition to following the public guidelines of the dataset, we consulted its authors to ensure proper handling. Per informed consent, eight participants allowed their facial images to be used in publications, while two (T001, T003) permitted use only for analysis but not for public display. We strictly adhered to these terms.

Generalizability. We used a naturalistic dataset in the context of knowledge work. Thus, our results pertain to individuals engaged in prolonged cognitive tasks. We make no claim about other contexts, such as conversations. Although the participant pool is small, it features 170 hours of second-by-second monitoring over four days, offering depth and breadth. Our model incorporates several covariates to mitigate confounding. Findings such as dominant left-hand touches align with psychological literature, suggesting representativeness and classification fidelity.

Transparency. To ensure reproducibility, we share the code and the relevant data. Annotated videos are released for the eight consenting participants. The code and data are available on GitHub [https://github.com/UH-ACDC/ACII-2025-sFST] and OSF [https://osf.io/pu9as/].

Acknowledgments. This work was partly supported by the NSF under award #1704682, "Managing Stress in the Workplace: Unobtrusive Monitoring and Adaptive Interventions."

REFERENCES

- D. Tsogbe, "Limits of expression: On touch, emotion, and communication," Ph.D. dissertation, Massachusetts Institute of Technology, 2023.
- [2] S. D. Suarez and G. G. Gallup Jr, "Face touching in primates: A closer look," *Neuropsychologia*, vol. 24, no. 4, pp. 597–600, 1986.
- [3] N. Zhang, W. Jia, P. Wang, M.-F. King, P.-T. Chan, and Y. Li, "Most self-touches are with the nondominant hand," *Scientific Reports*, vol. 10, no. 1, p. 10457, 2020.
- [4] J. L. Spille, M. Grunwald, S. Martin, and S. M. Mueller, "The suppression of spontaneous face touch and resulting consequences on memory performance of high and low self-touching individuals," *Scientific Reports*, vol. 12, no. 1, p. 8637, 2022.
- [5] J. L. Spille, S. M. Mueller, S. Martin, and M. Grunwald, "Cognitive and emotional regulation processes of spontaneous facial self-touch are activated in the first milliseconds of touch: Replication of previous EEG findings and further insights," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 22, no. 5, pp. 984–1000, 2022.
- [6] S. M. Mueller, S. Martin, and M. Grunwald, "Self-touch: contact durations and point of touch of spontaneous facial self-touches differ depending on cognitive and emotional load," *PLOS One*, vol. 14, no. 3, p. e0213677, 2019.
- [7] K. Densing, H. Konstantinidis, and M. Seiler, "Effect of stress level on different forms of self-touch in pre-and postadolescent girls," *Journal of Motor Behavior*, vol. 50, no. 5, pp. 475–485, 2018.

- [8] M. Grunwald, W. P. Pasatu, J. Spille, R. Haensel, J. Stieler, M. Holzer, M. Ziemer, K. H. Butz, S. Martin, S. M. Mueller *et al.*, "Number of facial hair corresponds to frequency of spontaneous face-touch in humans," *Advanced Biology*, p. 2400243, 2024
- Advanced Biology, p. 2400243, 2024.
 [9] C. Bai, Y.-P. Chen, A. Wolach, L. Anthony, and M. T. Mardini, "Using smartwatches to detect face touching," Sensors, vol. 21, no. 19, p. 6528, 2021
- [10] C. Beyan, M. Bustreo, M. Shahid, G. L. Bailo, N. Carissimi, and A. Del Bue, "Analysis of face-touching behavior in large scale social interaction dataset," in *Proceedings of the 2020 International Conference* on Multimodal Interaction, 2020, pp. 24–32.
- [11] M. R. Ibrahim and T. Lyons, "Facetouch: Detecting hand-to-face touch with supervised contrastive learning to assist in tracing infectious diseases," *PLOS One*, vol. 19, no. 6, p. e0288670, 2024.
- [12] M. Balazia, P. Müller, Á. L. Tánczos, A. v. Liechtenstein, and F. Bremond, "Bodily behaviors in social interaction: Novel annotations and state-of-the-art evaluation," in *Proceedings of the 30th ACM International Conference on Multimedia*, 2022, pp. 70–79.
- [13] P. Müller, M. X. Huang, and A. Bulling, "Detecting low rapport during natural interactions in small groups from non-verbal behaviour," in Proceedings of the 23rd International Conference on Intelligent User Interfaces, 2018, pp. 153–164.
- [14] K. Li, P. Liu, D. Guo, F. Wang, Z. Wu, H. Fan, and M. Wang, "Mmad: Multi-label micro-action detection in videos," arXiv preprint arXiv:2407.05311, 2024.
- [15] M. T. Hasan, S. Zaman, A. Wesley, P. Tsiamyrtzis, and I. Pavlidis, "Sympathetic activation in deadlines of deskbound research-a study in the wild," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–8.
- [16] D. Shastri, M. Papadakis, P. Tsiamyrtzis, B. Bass, and I. Pavlidis, "Perinasal imaging of physiological stress and its affective potential," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 366–378, 2012.
- [17] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *Journal of Comparative Neurology and Psychology*, vol. 18, no. 5, pp. 459–482, 1908.
- [18] D. S. Goldstein and B. McEwen, "Allostasis, homeostatsis, and the nature of stress," Stress, vol. 5, no. 1, pp. 55–58, 2002.
- [19] C. Blank, S. Zaman, A. Wesley, P. Tsiamyrtzis, D. R. Da Cunha Silva, R. Gutierrez-Osuna, G. Mark, and I. Pavlidis, "Emotional footprints of email interruptions," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–12.
- [20] J. A. Levine, I. T. Pavlidis, L. MacBride, Z. Zhu, and P. Tsiamyrtzis, "Description and clinical studies of a device for the instantaneous detection of office-place stress," Work, vol. 34, no. 3, pp. 359–364, 2009.
- [21] C. D. Spielberger, "State-trait anxiety inventory," in *The Corsini Encyclopedia of Psychology*. Hoboken, NJ: Wiley Online Library, 2010. [Online]. Available: https://doi.org/10.1002/9780470479216.corpsy0943
- [22] R. Hernandez, S. C. Roll, H. Jin, S. Schneider, and E. A. Pyatak, "Validation of the National Aeronautics and Space Administration Task Load Index (NASA-TLX) adapted for the whole day repeated measures context," *Ergonomics*, pp. 1–16, 2021.
- [23] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee et al., "Mediapipe: A framework for building perception pipelines," arXiv preprint arXiv:1906.08172, 2019.
- [24] G. Bradski, A. Kaehler, and V. Pisarevsky, "Learning-based computer vision with Intel's open source computer vision library," *Intel Technology Journal*, vol. 9, no. 2, 2005.
- [25] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond neural face detection on mobile GPUs," arXiv preprint arXiv:1907.05047, 2019.
- [26] Z. Vahedi and A. Saiphoo, "The association between smartphone use, stress, and anxiety: A meta-analytic review," *Stress and Health*, vol. 34, no. 3, pp. 347–358, 2018.