Proceedings of the 2020 Design of Medical Devices Conference
DMD2020
April 6, 7-9, 2020, Minneapolis, MN, USA

DMD2020-9062

ASSESSING INDUCED EMOTIONS IN EMPLOYEES IN A WORKPLACE SETTING USING WEARABLE DEVICES

Emma Fortune, PhD, Yaqoub Yusuf, Sarah Zornes, Jorge Loyo Lopez, Renaldo Blocker, PhD
Kern Center for the Science of Health Care Delivery, Mayo Clinic
Rochester, MN, USA

ABSTRACT

A working environment which elicits positive emotions in employees is vital for employee retention, engagement and productivity. Wearable sensors provide the means to objectively measure the emotional responses of employees in the workplace in real-time. The study aim was to perform a preliminary investigation into the validity of two multimodal systems to classify employee's emotional responses to positive, neutral or negative video stimuli: (1) using wearable electroencephalography (EEG) in combination with video-based facial expression analysis (FEA), and (2) using a wearable galvanic skin response (GSR) device in combination with video-based FEA. Five office employees each watched three short video clips at three time points during their regular work shifts while wearing EEG sensors on the forehead and GSR sensors on the middle and index fingers of their non-dominant hand with their face in view of a webcam. Russel's circumplex model of affect was used to determine participant's emotional responses to the video clips. The GSR device showed greater accuracy than the EEG device at detecting arousal responses to the video stimuli, with agreement, precision, and recall values of 87%, 100% and 80%, respectively, compared to 53%, 62%, and 80% for the EEG device. The FEA/EEG and FEA/GSR circumplex models were both able to accurately detect positive emotions elicited from video stimuli with levels of agreement and recall greater than 73%. Precision for the FEA/EEG model to detect positive stimuli was lower due to misclassification of 40% of both negative and neutral stimuli as positive. Precision values for both circumplex models were very low for detecting negative emotions. The results suggest that the EEG and GSR devices may be capable of detecting arousal when used alone, and detecting positive emotions when used in combination with video-based FEA in real-time in the workplace.

Keywords: Emotional affect, wearable sensors, electroencephalography, galvanic skin response, facial expression analysis.

1. INTRODUCTION

A working environment which induces positive emotions in employees is fundamental to reducing employee turnover and absenteeism, and to producing optimal engagement and productivity. In the healthcare industry, literature indicates it can cost an organization up to $1 million per physician turnover [1]. Workplace emotion is typically assessed using post-hoc tools such as validated questionnaires [2] which are subjective and not timely in identifying signs of positive or negative emotion. Recent advances in wearable sensors have led to the assessment of emotions in real-time to analyze user experience in marketing [3, 4] and gamification [5] studies, for psychology research studies [6, 7], as well as in the workplace for construction workers [8]. Limited research has been performed on healthcare professionals despite the high prevalence of associated burnout [9]; however, wearable devices could similarly be used to assess the emotions of healthcare employees in the workplace.

A growing body of literature exists on the use of wearable devices, measuring physiological signals such as electroencephalographic (EEG), galvanic skin response (GSR), electrocardiographic (ECG), photoplethysmography (PPG), and facial electromyographic (EMG) signals, and video-based facial expression analysis (FEA) to detect emotional responses and arousal, with EEG being the most direct and accurate measurement of arousal. The highest accuracy in emotion and arousal detection has been observed in studies on the fusion of different device modalities [10, 11, 12]. Positive and negative affect in the workplace can both be used as a metric to predict organizational outcomes such as employee turnover and work dissatisfaction [13]; thus, there's a significant opportunity to use novel technology to measure positive emotion in real-time by wearable devices.

The aim of this preliminary study was to determine the accuracy of two combinations of data fusion to assess the emotional responses of desk-based healthcare employees to video stimuli: (1) using wearable EEG combined with FEA from video data, and (2) using a wearable GSR device combined with FEA from video data. Future work will assess the use of these tools to assess emotions of healthcare employees across many roles throughout the work day.

2. METHODS

2.1 Participants

Five adult employees, recruited as convenience sample, were included in this study. All participants were right handed. Exclusion criteria included any cognitive or physical impairment, traumatic brain injury in the last 12 months, color-blindness, left handedness.

2.2 Procedure

Participants were individually tested three times during their regularly scheduled work shifts: at the beginning (morning), middle (lunch), and end (late afternoon/early evening). Video-based FEA, EEG, and GSR data were simultaneously collected through the iMotions biometric research platform version 7 (iMotions, Boston, MA) as participants were presented with one randomized stimulus at each one of the three observations. The stimuli were videos chosen by research staff: one to elicit negative emotion, one to elicit neutral emotion, and one to elicit positive emotion.

2.3 Facial Expression Analysis (FEA)

Video data, with the participant's face in full view, were recorded at 30 Hz using a portable Logitech C920 HD webcam (Logitech, Lausanne, Switzerland). Video-based FEA is able to estimate how negative or positive an emotion might be (valence) based on the facial expression of an individual. However, it cannot determine arousal: the strength of the response or the level of excitedness or calmness regardless of whether the response is associated with negative or positive emotion. Valence was calculated on a scale from -100 to 100 using FEA performed by iMotions' built-in AFFDEX facial expression engine (Affectiva, Boston, MA), with negative values indicating negative emotion and positive values indicating positive emotion. Positive valence is derived from two facial landmark changes (cheek raise, lip corner pull) to give an integrated valence score from 0 to 100. Negative valence is derived from eight facial landmark changes (brow furrow, chin raise, inner brow raise, lip corner depression, lip press, lip suck, nose wrinkle, upper lip raise) to give an integrated valence score from −100 to 0.

2.4 Electroencephalography (EEG)

EEG signals were collected at 500 Hz from the frontal lobes using the wireless Enobio 8 EEG system (Neuroelectrics, Barcelona, Spain). The EEG signals were recorded from eight channels using dry electrodes secured with a prefrontal headband. Frontal alpha asymmetry, used as a metric of arousal, is defined as the relative difference in alpha power (8-12 Hz) between the frontal left and right hemispheres for the duration of the stimulus exposure. This was calculated from the EEG data and exported using iMotions. Arousal values derived from frontal alpha asymmetry range from negative to positive with negative values indicating emotions associated with avoidance and positive values indicating emotions associated with engagement.

2.5 Galvanic Skin Response (GSR)

Participants' GSR data were acquired at 128 Hz using SHIMMER GSR+ units (SHIMMER research TM, Dublin, Ireland). The SHIMMER unit was secured to participant's non-dominant wrist with a strap, with a pair of bipolar electrodes securely attached to the index and middle finger. Peaks per minute were calculated from the phasic component of the GSR data in iMotions and used as the metric of arousal. Arousal values derived from GSR data range from zero to positive with more positive values indicating stronger intensities of the emotional response.

2.6 Russell's Circumplex Model

Russell's circumplex model of affect was used to determine participant's responses to the video stimuli from their valence and arousal data. The circumplex model was applied for two scenarios: (1) using FEA-derived valence (x-axis) with EEG-derived arousal (y-axis) and (2) using FEA-derived valence (x-axis) with GSR-derived arousal (y-axis). The model using EEG-derived arousal demonstrates the first (upper right side) and fourth quadrants (lower right side) as the positive emotion quadrants, and the second (upper left side) and third quadrants (lower left side) as the negative emotion quadrants. The model using GSR-derived arousal presents the right side as the positive emotion side and the left side as the negative emotion side.

2.7 Data Analysis

First, the accuracy of the EEG and GSR devices alone to detect a response (i.e. absolute arousal values were greater than 0) to the video clips was evaluated using agreement (percentage of correct classifications vs number of total classifications), precision and recall.

The emotional affect classification of a trial using the FEA/EEG model was determined as follows:

(1) Positive emotion - if the valence and absolute arousal values are greater than 0 (i.e. fall in the upper or lower right side quadrants),

(2) Negative emotion - if the valence is less than 0 and absolute arousal value is greater than 0 (i.e. fall in the upper or lower left side quadrants).

(3) Neutral emotion - if the valence or arousal has a value of 0.

The classification of a trial using the FEA/GSR model was determined as follows:

(1) Positive emotion - if the valence and arousal values are greater than 0 (i.e. fall on the right side),

(2) Negative emotion - if the valence is less than 0 and arousal value is greater than 0 (i.e. fall on the left side).

(3) Neutral emotion - if the valence or arousal has a value of 0.

The accuracies of both circumplex models were evaluated using agreement, precision and recall.

3. RESULTS

3.1 EEG

Arousal was detected using EEG data for 100% of the positive stimuli, and 60% of the negative stimuli (Fig. 2). Arousal was additionally detected for 100% of the neutral stimuli. The overall value of agreement was 53% with precision and recall values of 62% and 80%, respectively.

Using the EEG-derived arousal data in the circumplex model, agreement, precision, and recall for positive emotion detection were 73%, 56%, and 100%, respectively. 100% of participants responded with positive emotion to the positive stimuli (Fig. 1). Furthermore, 60% of the participants responded in the first quadrant and 40% responded in the fourth quadrant. Examples of emotions that can be exhibited in the first quadrant are: excited, astonished, happiness, and joy. Examples of emotions that are exhibited in the fourth quadrant are: pleased, content, and relaxed. Agreement, precision, and recall for negative emotion detection were 60%, 0%, and 0%, respectively. Only one of the participants responded with negative valence to the negative stimuli, however, their arousal value was 0. In addition, 40% of participants responded with positive emotion and the other 40% responded with neutral affect to the negative stimuli. Agreement, precision and recall for neutral emotion detection were 60%, 40%, and 40%, respectively, with one misclassified as negative and two misclassified as positive. However, the participants' responses to the neutral stimuli all fell on or very close to the neutral midline.

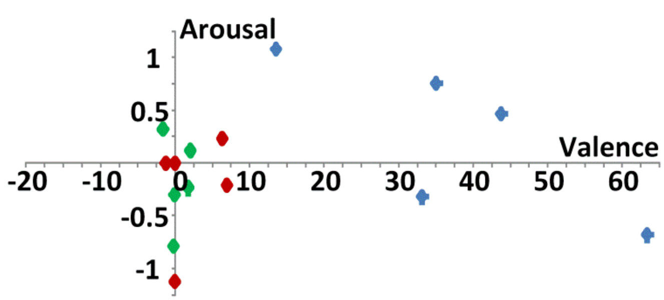

FIGURE 1: RUSSELL'S CIRCUMPLEX MODEL OF AFFECT WITH FACIAL EXPRESSION ANALYSIS (FEA)-DERIVED VALENCE AND ELECTROENCEPHALOGRAPHY (EEG)-DERIVED AROUSAL. THE BLUE DIAMONDS DENOTE THE POSITIVE STIMULI TRIALS, THE GREEN DIAMONDS DENOTE THE NEUTRAL STIMULI TRIALS, AND THE RED DIAMONDS DENOTE THE NEGATIVE STIMULI TRIALS.

3.2 GSR

Arousal was detected using GSR data for 80% of the positive stimuli, and 80% of the negative stimuli (Fig. 2). No arousal was detected during the five neutral stimuli. Overall agreement was 87% with precision and recall values of 100% and 80%, respectively.

Using the EEG-derived arousal data in the circumplex model, agreement, precision, and recall for positive emotion detection were 87%, 80%, and 80%, respectively (Fig. 2). Although 100% of participants responded with positive valence to the positive stimuli, there was no arousal detected for one of the responses. Agreement, precision, and recall for negative emotion detection were 73%, 100%, and 20%, respectively. Only one of the participants responded with negative valence to the negative stimuli. However, no false positives were detected. Agreement, precision and recall for neutral emotion detection were 80%, 63%, and 100%, respectively, No arousal responses were detected for any of the neutral stimuli. However, 60% of negative and 20% of positive stimuli were falsely classified as neutral.

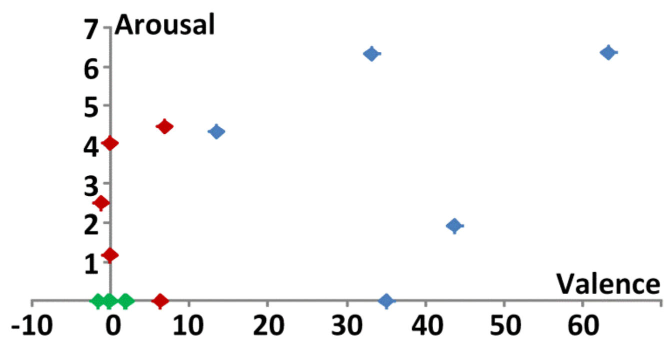

FIGURE 2: RUSSELL'S CIRCUMPLEX MODEL OF AFFECT WITH FACIAL EXPRESSION ANALYSIS (FEA)-DERIVED VALENCE AND GALVANIC SKIN RESPONSE (GSR)-DERIVED AROUSAL. THE BLUE DIAMONDS DENOTE THE POSITIVE STIMULI TRIALS, THE GREEN DIAMONDS DENOTE THE NEUTRAL STIMULI TRIALS, AND THE RED DIAMONDS DENOTE THE NEGATIVE STIMULI TRIALS.

4. INTERPRETATION

The GSR device showed greater accuracy than the EEG device at detecting arousal to the video stimuli due to the EEG detecting arousal for all neutral video trials. The FEA/EEG and FEA/GSR circumplex models were both able to accurately detect positive emotions elicited from video stimuli with levels of agreement and recall greater than 73%. The level of precision for the FEA/EEG model was lower due to 40% of negative and 40% of neutral stimuli being misclassified as positive. This was due to the video-based FEA detecting positive valance as well as the EEG detecting an arousal response. The majority of the negative stimuli resulted in no emotion being detected with either the FEA/EEG or the FEA/GSR models. However, this may be due to a number of limitations with the study including the subjective choosing of the videos to induce emotions by study team members, and small sample size. In addition, an agreed valence threshold value to distinguish between neutral affect and either positive or negative affects does not currently exist in the literature although one study arbitrarily chose an absolute value of 10 [16]. Our threshold value of 0 was similarly chosen arbitrarily. Using an absolute valence threshold of 10 in our study would mean that all negative and neutral trials were classified as neutral. However, our results from the positive trials would remain unchanged. As GSR-derived arousal is only non-zero if a response peak occurs in the signal, the choice of an arousal threshold of greater than 0 seemed appropriate.

Future work will involve the use of validated stimuli and the collection of survey data to ascertain participant's emotions prior to testing and self-reported response during each stimulus along with larger participant and stimulus numbers. We will also explore the inclusion of heart rate data measured either using wearable ECG or PPG devices. Heart rate variability is considered as a very sensitive indicator of stress [14] and has been shown to detect emotions such as stress with high accuracy when used in combination with GSR data [15]. Nonetheless, these data suggest that the devices used in this study may be capable of detecting both arousal and positive emotions in real-time in the workplace.

ACKNOWLEDGEMENTS

Funding for this project was provided by the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery.

REFERENCES

1. Perlo, J, Balik, B, Swensen, S, Kabcenell, A, Landsman, J, and Feeley, D. IHI framework for improving joy in work. *Cambridge, MA: Institute for Healthcare Improvement*, 2017, pp. 3-41.
2. Harmon-Jones, C, Bastian, B, and Harmon-Jones, E, "The discrete emotions questionnaire: a new tool for measuring state self-reported emotions", *PloS One*, 2016, 11(8), pp. e0159915.
3. Ural, G, Kacar, F, and Canan, S, "Wavelet Phase Coherence Estimation of EEG Signals for Neuromarketing Studies", *NeuroQuantology*, 2019, 17(2), pp. 112-120.
4. Khushaba, RN, Wise, C, Kodagoda, S, Louviere, J, Kahn, BE and Townsend, C, "Consumer neuroscience: assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking", *Expert Systems with Applications*, 2013, 40(9), pp. 3803-12.
5. Chanel G, Rebetez C, Bétrancourt M, and Pun T. "Emotion assessment from physiological signals for adaptation of game difficulty", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*. 2011, 41, pp. 1052-63.
6. Wacker, J. "Effects of positive emotion, extraversion, and dopamine on cognitive stability-flexibility and frontal EEG asymmetry", *Psychophysiology*, 2018, 55(1), pp. 16.
7. Wu, CT, Dillon, DG, Hsu, HC, Huang, SA, Barrick, E, and Liu, YH. "Depression Detection Using Relative EEG Power Induced by Emotionally Positive Images and a Conformal Kernel Support Vector Machine", *Applied Sciences-Basel*, 2018, 8(8), pp. 18.
8. Hwang, S, Jebelli, H, Choi, B, Choi, M, and Lee, S, "Measuring workers' emotional state during construction tasks using wearable EEG." *Journal of Construction Engineering and Management*, 2018, 144(7), pp. 04018050.
9. Gajewski, PD, Boden, S, Freude, G, Potter, GG, and Falkenstein M, "Burnout is associated with changes in error and feedback processing", *Biol Pyschol*, 2017, 129, pp. 349-58.
10. Zheng, WL, Dong, BN, and Lu, BL. "Multimodal emotion recognition using EEG and eye tracking data," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 5040–43.
11. Soleymani, M, Asghari-Esfeden, S, Pantic, M, and Fu, Y, "Continuous emotion detection using EEG signals and facial expressions," in *2014 IEEE International Conference on Multimedia and Expo*, 2014, pp. 1–6.
12. Tarnowski, P, Kołodziej, M, Majkowski, A, and Rak, RJ, "Combined analysis of GSR and EEG signals for emotion recognition", In *2018 IEEE International Interdisciplinary PhD Workshop (IIPhDW)*, 2018, pp. 137-141.
13. Schutte, L. "Understanding the real costs of recruiting and retaining physicians" *Recruiting Physicians Today*, 2012, 20, pp. 1-3.
14. Sriramprakash, S, Prasanna, VD, Murthy, OVR. "Stress Detection in Working People". *Procedia Computer Science*, 2017, 115, pp. 359–366.
15. De Santos Sierra, A, Ávila, CS, Del Pozo, GB, Casanova, JG. "Stress detection by means of stress physiological template". *In Proceedings of the Third World Congress on Nature and Biologically Inspired Computing*, 2011, pp. 131–136.
16. Brand, R, Ulrich, L. "I can see it in your face. Affective Valuation of Exercise in More or Less Physically Active Individuals". *Frontiers Psychology*, 2019, 10, pp. 2901.

