Detecting Divisions of the Autonomic Nervous System Using Wearables

Annamalai Natarajan*[1], Kevin S. Xu*[2], *Member, IEEE*, and Brian Eriksson[3]

***Abstract*— The ability to assess a user's emotional reaction from biometrics has applications in personalization, recommendation, and enhancing user experiences, among other areas. Unfortunately, understanding the connection between biometric signals and user reactions has previously focused on black box techniques that are opaque to the underlying physiology of the user. In this paper, we explore a novel user study connecting biometric reaction to external stimuli and changes in the user's autonomic nervous system. Specifically, we focus on two competing responses, namely the sympathetic and parasympathetic nervous system, and how differing activations are related to different user responses. Our experiments demonstrate how prior psychophysiological research distinguishing this activation can be replicated using biometric data collected from wearables. The insights from this work have applications in better understanding emotional state from biometric sensors.**

I. Introduction

The need to assess emotional reactions from biometric signals (*e.g.,* heart rate or electrodermal activity) has applications in content analysis [1] and stress analysis [2]—to name only a few use cases. While prior work has detected stress [3] and attention [4], these typically treat biometrics as a "black box", ignoring the underlying physiology. This approach fails to reveal *why* biometrics are successful in distinguishing some tasks but not others.

Humans rely on the autonomic nervous system to express emotional reactions. The autonomic nervous system (ANS) is divided into two components, the sympathetic (SNS) and parasympathetic (PNS) nervous systems. The modulation of these two components, referred to as the SNS-PNS activation profile, is dependent on the reaction to the specific external stimulus. For certain tasks, these activation profiles have been previously studied but in isolation only. The novelty of our work is to show that tasks with disparate SNS-PNS activation profiles can be distinguished using biometric signals. Additionally, we find that this can be done using biometrics from wearable wristband devices.

In this paper, we choose to study two tasks, namely fear conditioning and cognitive load, which exhibit different SNS-PNS activation profiles, while gathering data from wristband sensors. We find that we can distinguish between two tasks almost perfectly *e.g.,* fear conditioning vs. cognitive load but only moderately when distinguishing within each task *e.g.,* low vs. high cognitive load. More importantly, we observe that these features agree with the underlying theory proposed in the psychophysiological literature. These results provide some insight as to which types of tasks can be detected and reveal a deeper understanding of user reactions from biometric signals. This paper demonstrates that studying the autonomic nervous system is worthwhile as it opens up the possibility of mapping individuals and more complex tasks onto a common SNS-PNS activation profile space. This work is the first step in that direction.

TABLE I: Functional plane of SNS and PNS (adapted from Berntson et al., [5]). The task listed in a cell has been shown to elicit the corresponding SNS and PNS responses (increase: ↗, no change: →, and decrease: ↘).

		PNS Response		
		↗	→	↘
SNS Response	↗	Fear conditioning		Cognitive load
	→		Rest (baseline)	
	↘			

II. Background and Related Work

The ANS controls organs like the heart, lungs, pupils, intestines, and bladder with the purpose of responding to external stimululi. It is a control system that operates involuntarily and reflexively, making it one of the best systems to study without interference.

The ANS is broadly divided into the SNS and PNS divisions. The SNS is generally regarded as the "fight or flight" system that is activated under acute stress situations. High-level SNS functions include turning off the digestive system and increasing heart rate to improve blood flow to the limbs to take off under adverse situations. The PNS, on the other hand, is generally regarded as the "rest and digest" system that is recruited by the ANS under stress-free situations, often to counter balance the activities of the SNS system. The PNS system decreases heart rate and turns on saliva secretion to aid digestion.

Critical to understanding the effects of the ANS is to understand how SNS and PNS interact. Berntson et al., in their seminal paper, proposed that the SNS and PNS operated along a 2-D functional plane [5]. Specifically, Berntson et al., proposed that there are three states in which the underlying divisions of the ANS can operate: increase, no change and decrease. Thus this leads to a 3×3 matrix with nine possible states in which the two divisions can function, as shown in Table I.

One of the major advantages of studying the ANS is that this underlying division of SNS and PNS can be measured

*This work was done while these authors were with Technicolor.

[1]University of Massachusetts Amherst, **anataraj@cs.umass.edu**

[2]University of Toledo, **kevin.xu@utoledo.edu**

[3]Technicolor, **brian.eriksson@technicolor.com**

with sensors measuring the heart, lungs and sweat glands. The physiological data of interest from these organs include the heart rate (HR), breathing rate (BR) and electrodermal activity (EDA, also known as galvanic skin response). The HR and BR are influenced by both divisions of the ANS. From the HR one can compute the heart rate variability (HRV) and extract magnitudes of frequency components such as high frequency (HF) and low frequency (LF) components using the area under the power spectral density curve. The HF component serves as a direct index of PNS activity [6]. It is generally agreed upon that the LF component is influenced by both the SNS and PNS divisions [6], [7]. The autonomic balance, defined as LF/HF, is an index of SNS activity [8]. EDA on the other hand, is influenced only by the SNS [9].

To the best of our knowledge, there have not been meta-analysis studies that report tasks performed in each cell in Table I. Instead, we are interested in two specific tasks that are well-studied: fear conditioning and cognitive load. In fear conditioning, a neutral stimulus is paired with an aversive stimulus to evoke expected fear after repeated presentations [10]; Cohen et al. [11] observed that its SNS-PNS activation profile is excitation-excitation, *i.e.* cell (1, 1). In the cognitive load task, Steinhauer et al. in [12] suggested that the SNS-PNS were following an excitation-inhibition activation profile, *i.e.* cell (1, 3).

Wearable computing has mostly focused on tasks within a single cell in Table I (*e.g.,* [3], [13]). These studies detect target tasks from either baseline or varying degrees of subject engagement in the same task. In contrast, this paper focuses on discriminating tasks chosen from *different* cells in Table I in a principled manner with respect to the SNS and PNS divisions of the ANS.

III. Data Collection and Feature Extraction

Our experiments involved 24 volunteers (9 female) with a mean age of 32.29 ± 8.65. This research was conducted in a corporate research environment. Subjects were briefed on the experiment design and participated in the experiments only after giving informed consent. Four additional subjects opted out of the experiment after hearing the experiment design, and all subjects were given the option to end the experiment at any time.

Subjects were seated in a chair wearing noise canceling headphones in a dark room. Subjects wore the Empatica E3 wearable wristband [14] on both hands for the duration of the experiment ($\sim$30 min.). Wearing wrist bands on both hands was a precautionary measure to avoid data loss due to sensor failure. In this study, we analyzed data from the left hand sensor for all but one of the subjects whose left hand sensor crashed (for this subject we used the right sensor data).

A. Video Experiment Setup

A sequence of stimuli were presented to subjects as a single video clip on an iPad. Subjects were instructed to respond verbally to tasks which required subject responses. An experimenter, seated in the back of the room, recorded subjects' responses. Our experiment has three main tasks: fear conditioning, cognitive load, and rest.

1) Fear Conditioning: The purpose of fear conditioning is to train the subject to expect a jarring or fear-inducing event. This is performed by introducing two disparate tones. One "positive tone" often was followed by a fear-inducing sound, while the other "negative tone" was followed by no stimulus. We adopted and modified the fear conditioning paradigm from [15]. Subjects were instructed to listen to the tones passively and uncover the hidden pattern.

In our experiments we chose positive and negative tones as triangle waveforms at 800Hz and 300Hz respectively [16]. The tones were presented for a duration of 5 seconds each accompanied by a blank screen with a fixation cross. The tones were separated by a random jitter of 20 ± 4 seconds. The fear-inducing sound used was the sound of machine gun fire presented for a duration of 3 seconds.

In order to train the user to anticipate the fear-inducing sound, the fear conditioning task has three phases: habituation, acquisition and extinction. In the habituation phase, subjects were presented with positive and negative tones, twice each, in random order. In the acquisition phase, subjects were presented with positive and negative tones, four times each, but the positive tones were immediately followed by the sound of machine gun fire. Finally, in the extinction phase, subjects were presented with positive and negative tones, twice each, in random order. Post experiment discussions revealed that at least half the participants identified the pairing of machine gun fire with the positive tone in the acquisition phase.

2) Cognitive Load: We adopted the cognitive load paradigm from [12]. Subjects were presented with a random integer between 500 to 1000 on the screen. Subjects were then instructed to repeatedly add 2 or subtract 7 for a duration of 30 seconds and verbally respond after each addition or subtraction. Subjects were instructed to add or subtract as many times as possible during the 30 second interval.

Each subject was presented two instances of add 2 and two instances of subtract 7. The task ordering was add 2, subtract 7, add 2, and subtract 7. The random integers and task orderings remained the same across all subjects. Our focus was to analyze changes of user reactions when performing a low cognitive load task (add 2) against a high cognitive load task (subtract 7). This differentiation can initially be seen in the performance of the subjects, as on average they performed 17 ± 6 additions, but only 10 ± 4 subtractions.

3) Rest: Subjects sat still and stared at a blank screen for the last 30 seconds of the experiment. This we use as a rest period where the subjects were not engaged with any task.

B. Wearable Sensor Setup

We associate wearable biometric signals with subject responses to the experiment. Among other data streams, Empatica wrist bands collect EDA at 4 Hz and photoplethysmogram (PPG) data at 64 Hz. EDA is a measure of skin conductance changes in the subject as a result of psychological arousal or excitation events [9]. In order to record high

quality EDA, we placed two adhesive Ag/AgCl electrodes on both palms, with specific placement at the thenar and hypothenar eminences as informed by [9]. These adhesive electrodes were soldered to the wearable wristbands. PPG measures changes in blood volume via the use of LEDs and a photodiode measured at the top of the subject's wrist.

C. Biometric Signal Features

1) Electrodermal Activity Features: We observed significant baseline shifts in the EDA time series data. This can be attributed to the sweat gland activity over time as well as the nature of the task. To mitigate the effects of baseline shifts and to make comparison across subjects valid, we performed a baseline correction by mean subtracting the EDA data. This correction was performed per subject within each of the individual tasks, namely fear conditioning, cognitive load, and rest.

Following baseline correction we extracted EDA features associated to skin conductance levels (SCLs) and skin conductance responses (SCRs). We detected EDA peaks by taking the derivative of baseline-corrected EDA data. Data points whose slopes were below a hard threshold of 1 were eliminated as these peaks were considered noise. We then detected pairs of change points where the slopes changed from negative to positive and positive to negative, respectively. These pairs were identified as the start and peak of SCRs, respectively [17]. From locations of the SCRs we computed inter-arrival times (time duration between adjacent SCRs), rise times and amplitudes. Along with these features we also computed mean EDA over each 30-second window, which takes both SCLs and SCRs into account.

2) Cardiovascular Features: From the PPG data we estimated the HR since PPG also reflects cardiac rhythm. From the heart rates we computed LF and HF components as the area under the power spectral density curves corresponding to frequency ranges $0.04-0.15$ Hz and $0.15-0.4$ Hz respectively [18], [6]. Like in [3], LF and HF were normalized to minimize the impact of difference in total power across tasks and across subjects. We also computed the autonomic balance as ratio of LF to HF. The breathing rate (BR) was computed as the center frequency of the highest peak in the high frequency range $0.15 - 0.4$ Hz [3].

3) Feature Groupings: We split the EDA and PPG measurements into time windows of varying lengths (explained in Section IV) and computed a total of 16 features for each time window. We were specifically interested in three feature groupings: SNS, PNS, and BOTH. We regard mean EDA; SCR count; mean, median, and variance of SCR amplitude; arrival times; and rise times as features influenced by SNS division only. The HF component of HR is the only feature indicative of PNS division. BOTH corresponds to HR, HRV, LF component of HR, and BR, all of which are influenced by both divisions of the ANS. Two other variants include SNS-PNS, which combines SNS and PNS features but not BOTH, and ALL, which includes all features. In total we experimented with five groupings of features.

TABLE II: Mean leave-one-subject-out AUROC over all subjects for classification experiments

	SNS	**PNS**	**SNS-PNS**	**BOTH**	**ALL**
Fear vs. Neutral	0.60 ± .01	0.53 ± .01	0.67 ± .01	0.53 ± .01	0.63 ± .01
Add vs. Subtract	0.58 ± .01	0.43 ± .01	0.60 ± .01	0.60 ± .01	0.62 ± .01
Fear vs. Rest	1.00 ± .00	1.00 ± .00	1.00 ± .00	0.46 ± .02	1.00 ± .00
Subtract vs. Rest	1.00 ± .00	0.96 ± .01	1.00 ± .00	0.96 ± .01	0.98 ± .00
Fear vs. Subtract	1.00 ± .00	1.00 ± .00	1.00 ± .00	0.99 ± .00	1.00 ± .00
Fear vs. Subtract vs. Rest	0.95 ± .00	0.99 ± .00	1.00 ± .00	0.49 ± .01	0.99 ± .00

IV. Results

Our results detail the ability to distinguish between tasks with specific SNS-PNS activation profiles. We use the subject's response from biometric features extracted from wearable wristbands as input into a classification problem (*e.g.,* add vs. subtract tasks).

In our experiments we performed leave-one-subject-out cross validation on our 24 subjects. We used penalized versions of binomial and multinomial logistic regression implemented in the *minFunc* toolbox [19] with penalty parameters optimized by cross validation. We present results using ℓ_2-penalized logistic regression[1]. We report the area under receiver operating characteristics curve (AUROC) for all classification problems. For multiclass problems we report unweighted pairwise AUROC [20].

A. Within-Cell Classification

The fear conditioning task has been reported to cause an increase in SNS and PNS divisions of the ANS. We hypothesize that the ANS response to fear-inducing tones is distinguishable from ANS response to neutral tones. Using the physiological data from the acquisition phase (the phase where machine gun fire is presented), we used the period between 1 to 10 seconds after the stimulus as a window to the subject's fear-induced response. After extracting SNS and PNS features during this 10-second window for both fear-inducing and neutral tones, we report AUROC under row 'Fear vs. Neutral' in Table II. We find that ANS features can weakly discriminate between the responses to the two tones, as indicated by the best AUROCs of 0.67.

Next, we examine ANS divisions for tasks with different cognitive load. In our experiments, subtract 7 was considered a high cognitive load task when compared to the add 2 task. We used physiological data from 30 second intervals to extract features. We trained a classifier to detect windows of high cognitive load from windows of low cognitive load. We report the AUROC under row 'Add vs. Subtract' in Table II. Here, we also observe weak discrimination between the two tasks, with the best AUROC of 0.62. Poor PNS performance suggests that the PNS activation profile may be very similar for the two cognitive load tasks but in conjunction with SNS it leads to different activation profiles for different loads.

[1]We observed comparable results for gradient-boosted trees, decision trees, ridge regression, and AdaBoost classifiers

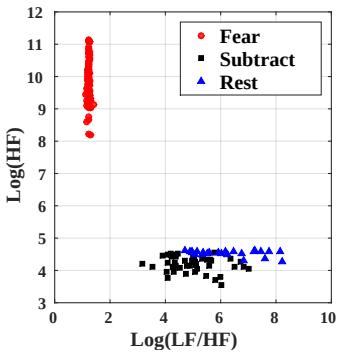

Fig. 1: Scatter plots of subjects' responses using HF component of HR (PNS) and LF/HF ratio (SNS). Each point denotes the response of a single subject to a single task.

B. Between-Cell Classification

While the previous classification results were for tasks within cells, in this section we classify between tasks in *different* cells in the 2-D functional plane. Specifically, we compare the SNS-PNS activation profiles for fear-induced response, high cognitive load, and rest periods. Informed by the prior literature, we hypothesize that the PNS activation profiles should be higher for fear-induced response (PNS increases) and lower for high cognitive load tasks (PNS decreases) when compared to rest period. In addition, the rest task should have different SNS activation compared with subtract and fear tasks.

Using both pairwise and multi-class classification, we report AUROCs in the last four rows of Table II for the four classification tasks, respectively. Unlike the within cell classification tasks, we observe near-perfect classification accuracy for between-cell classification.

The reason for the extremely high classification accuracy between tasks is revealed in Fig. 1. The vertical and horizontal axes of this scatter plot are measures of PNS and SNS activity, respectively. Notice that subjects' responses to fear-inducing tones are easily separated from both their responses to the subtraction task and during rest periods even when considering just these 2 features. Specifically notice that fear is very easily separated from subtraction using only PNS; this agrees with the locations of the two tasks in cells (1,1) and (1,3) in Table I indicating opposition in the PNS states for these tasks. Additionally, notice the separability of fear and subtraction using SNS only. This suggests that, despite both tasks exhibit the same direction component, they are separable along a magnitude component (small increase versus big increase). This insight has never been explored in the prior literature.

C. Conclusions

Our results demonstrate that the measures of PNS and SNS activity computed from data collected using wearables do reveal insights into user reactions during tasks. This is a first step towards better understanding the underlying physiology of the ANS of the user. A limitation of the current study is that it was conducted in a one-time setting in a lab environment, so the expected accuracy is likely to be lower in a real-world application due to day-to-day variations and measurement artifacts. Future work includes analyzing SNS-PNS activations during more complex stimuli (*e.g.,* film clips) and observing how changes in SNS-PNS activation relate to underlying user emotions.

REFERENCES

[1] F. Silveira, B. Eriksson, A. Sheth, and A. Sheppard, "Predicting audience responses to movie content from electro-dermal activity signals," in *Proc. ACM UbiComp*, 2013, pp. 707–716.

[2] H. Lu, D. Frauendorfer, M. Rabbi, M. Mast, G. Chittaranjan, A. Campbell, D. Gatica-Perez, T. Choudhury, "StressSense: Detecting Stress In Unconstrained Acoustic Environments Using Smartphones," in *Proc. ACM UbiComp*, 2012, pp. 351–360.

[3] D. McDuff, S. Gontarek, and R. Picard, "Remote measurement of cognitive stress via heart rate variability," in *36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 2957–2960.

[4] E. Haapalainen, S. Kim, J. F. Forlizzi, and A. K. Dey, "Psycho-physiological measures for assessing cognitive load," in *Proc. ACM UbiComp*, 2010, pp. 301–310.

[5] G. G. Berntson, J. T. Cacioppo, and K. S. Quigley, "Autonomic determinism: the modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint." *Psychol. Rev.*, vol. 98, no. 4, p. 459, 1991.

[6] J. Sztajzel, "Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system," *Swiss Med. Wkly.*, vol. 134, pp. 514–522, 2004.

[7] D. S. Goldstein, O. Bentho, M.-Y. Park, and Y. Sharabi, "Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes," *Exp. Physiol.*, vol. 96, no. 12, pp. 1255–1261, 2011.

[8] C. A. Polanczyk, L. E. Rohde, R. S. Moraes, E. L. Ferlin, C. Leite, and J. P. Ribeiro, "Sympathetic nervous system representation in time and frequency domain indices of heart rate variability," *Eur. J. Appl. Physiol. Occup. Physiol.*, vol. 79, no. 1, pp. 69–73, 1998.

[9] J. T. Cacioppo, L. G. Tassinary, and G. Berntson, *Handbook of psychophysiology*. Cambridge University Press, 2007.

[10] S. Lissek, A. S. Powers, E. B. McClure, E. A. Phelps, G. Woldehawariat, C. Grillon, and D. S. Pine, "Classical fear conditioning in the anxiety disorders: a meta-analysis," *Behav. Res. Ther.*, vol. 43, no. 11, pp. 1391–1424, 2005.

[11] D. Cohen and D. Randall, "Classical conditioning of cardiovascular responses," *Annu. Rev. Physiol.*, vol. 46, no. 1, pp. 187–197, 1984.

[12] S. R. Steinhauer, G. J. Siegle, R. Condray, and M. Pless, "Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing," *Int. J. Psychophysiol.*, vol. 52, no. 1, pp. 77–86, 2004.

[13] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Troster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable eda device," *IEEE Trans. Inf. Tech. Biomed.*, vol. 14, no. 2, pp. 410–417, 2010.

[14] M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti, "Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition," in *Proc. 4th Int. Conf. Wireless Mobile Commun. Healthcare*, 2015, pp. 39–42.

[15] S. T. Pohlack, F. Nees, C. Liebscher, R. Cacciaglia, S. J. Diener, S. Ridder, F. G. Woermann, and H. Flor, "Hippocampal but not amygdalar volume affects contextual fear conditioning in humans," *Hum. Brain Mapp.*, vol. 33, no. 2, pp. 478–488, 2012.

[16] "Online tone generator," http://onlinetonegenerator.com/.

[17] K. H. Kim, S. Bang, and S. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, 2004.

[18] L. A. Lipsitz, J. Mietus, G. B. Moody, and A. L. Goldberger, "Spectral characteristics of heart rate variability before and during postural tilt. relations to aging and risk of syncope." *Circulation*, vol. 81, no. 6, pp. 1803–1810, 1990.

[19] M. Schmidt, "Minfunc toolbox," http://www.di.ens.fr/~mschmidt/Software/minFunc.html.

[20] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.