EXPLORING BIO-BEHAVIORAL SIGNAL TRAJECTORIES OF STATE ANXIETY DURING PUBLIC SPEAKING

Ehsanul Haque Nirjhar [1]   *Amir Behzadan* [2]   *Theodora Chaspari* [1]

[1] HUman Bio-Behavioral Signals (HUBBS) Lab, Department of Computer Science & Engineering
[2] Construction Informatics and Built Environment Research (CIBER) Lab, Construction Science
Texas A&M University, College Station, Texas, 77843, USA

ABSTRACT

Public speaking anxiety (PSA) is among the top social phobias in the world. Quantifying PSA in a reliable and unobtrusive manner can lay the foundation toward personalized and inexpensive technology-based interventions. Existing work for quantifying PSA often relies on self-reported measures and statistical aggregates of bio-behavioral indices, such as physiology and speech. Such aggregated bio-behavioral indices are not able to capture time-based trajectories of PSA variation, that can be very useful for better understanding and reliably predicting moments of anxiety. We tackle this problem by introducing temporal parametric models to quantify bio-behavioral trajectories of PSA throughout a public speaking encounter. Using data from 55 participants in a real-life public speaking task, the parameters of the proposed models are found to be significantly correlated with individuals' trait characteristics of general and communication-based anxiety, outperforming aggregate mean bio-behavioral measures.

***Index Terms*—** Public speaking anxiety, physiology, speech, transient oscillation, quadratic polynomial

1. INTRODUCTION

Public speaking is an important aspect of human communication, since it enables individuals to present their ideas, persuade others, and make tangible impact [1]. Yet, recent statistics indicate a large portion of people rank public speaking as one of their top fears [2, 3]. This communication-based phobia, often referred to as "public speaking anxiety" (PSA), is known to elicit physiological arousal (e.g., increased heart rate) and negative affect (e.g., feelings of disappointment) [4, 5]. Designing novel technologies for detecting, predicting, and mitigating PSA might help people in various stages of their academic and professional careers, potentially contributing to life-long learning and re-skilling. Also, the detection of such anxiety in less obtrusive manner is necessary to help people overcome this communication disorder.

PSA is often viewed as "state" (or situational) anxiety, experienced in-the-moment given a specific stimuli [6]. Individuals with higher "trait" anxiety, who are inherently prone to stress and negative emotions, may suffer more from state PSA than their low trait anxiety counterparts [7]. Beyond the conventional self-reported indices, PSA can present itself as a change in bio-behavioral signals, such as voice intonation, sweat secretion, and heart activity [8, 9, 10]. PSA further depicts temporal variations throughout the public speaking task, which might include decreasing (habituation) or increasing (escalation) patterns over time, as well as increasing patterns followed by a decrease (sensitization) [5]. However, existing work relies heavily on aggregated measures of bio-behavioral signals for modeling PSA, overlooking temporal variations, which can be indicative of the various individual traits [5]. For example, speakers with low trait anxiety depict reduced escalation in their physiological patterns, compared to their high trait anxiety counterparts [11]. Individuals with higher sensitivity to public speaking depict strong sensitization patterns [12]. These indicate the presence of high inter-individual variability with respect to temporal PSA patterns. Quantifying such differences can potentially help us better predict momentary anxiety during public speaking, afford us novel insights into the ways anxiety evolves throughout the public encounter, and help us design in-the-moment interventions with appropriate feedback mechanisms.

We quantify temporal trajectories of bio-behavioral signals within a public speaking task through two types of parametric models. A transient oscillation model mimics alternating habituation and sensitization trajectories. A quadratic polynomial model represents non-oscillatory escalation or habituation patterns. The considered bio-behavioral indices include acoustic and physiological measures, commonly used for quantifying trait anxiety [13, 14, 10]. The parameters of the proposed transient oscillation and quadratic polynomial models are highly interpretable and serve as quantitative descriptors of PSA trajectories, examined in association to individuals' trait-based general and communication anxiety. Our results obtained in 55 participants indicate that the proposed time-based representations of bio-behavioral measures can more reliably estimate participants' characteristics compared to the corresponding aggregate mean scores.

The authors would like to acknowledge the Engineering Information Foundation (EiF) for funding this work (grant EiF 18.02).

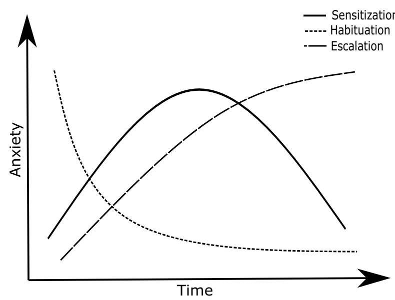


Fig. 1. Theoretically-grounded time trajectories of public speaking anxiety.

2. RELATED WORK

Spielberger [15] presented a clear distinction between "state" and "trait" anxiety. The former corresponds to experiencing stress at a particular scenario and time, while the later refers to a general tendency to experience anxiety across situations and time. Most works view PSA as state anxiety [10, 5]. Batrinca *et al.* [16] incorporated acoustic measures (intensity, fundamental frequency, number of pauses) and gestures (gaze, limb movement) for modeling PSA. Chen *et al.* [17] proposed a multimodal approach to predict audience evaluation scores from speech intonation, pitch, and other acoustic measures. Yadav *et al.* [10] incorporated demographics with a multimodal feature set to predict state anxiety.

While the aforementioned work has modeled PSA through a single aggregated measure, other work indicates that the time-based variability in "state" anxiety can be better explained when incorporating measures of "trait" anxiety, nervousness, and demographics [18]. Bodie presented distinct patterns of habituation and sensitization over time during a public speaking encounter [5] (Fig. 1). Habituation refers to higher stress anticipation at start. Sensitization corresponds to experiencing a rise in anxiety at first, followed by a stable state or a decrease. Escalation is associated with constantly increasing anxiety. These findings indicate that the aggregation of bio-behavioral indices over time might fail to reveal important information about the PSA. Motivated by these, this paper quantifies temporal patterns of bio-behavioral signals to model time trajectories of PSA within a public speaking task. The contributions are the following: (1) We propose to model temporal trajectories of PSA based on parametric transient oscillation and quadratic polynomial models of bio-behavioral signals. The oscillatory model captures alternating habituation and sensitization trajectories, while the polynomial model represents non-oscillatory habituation, escalation, or sensitization; (2) We examine the association of the parameters of the aforementioned models to individuals' psychological characteristics, in an effort to obtain new insights into the ways different people experience the public speaking task.

3. DATA DESCRIPTION

Our data includes 55 participants (23 female, 21 years average age) performing a public speaking task in front of a real audience. Participants were given 10 minutes to prepare their speech based on a randomly assigned news article from various topics (e.g. history, business, well-being/healthcare). Following that, participants had to present the article in front of the audience for 5 minutes. More details on the procedures can be found in Yadav *et al.* [10].

For the acquisition of physiological and speech signals, participants wore a wrist-mounted Empatica E4 watch [19] and a Creative lavalier microphone. The Empatica E4 collects electrodermal activity (EDA), sampled at 4 Hz, and average heart rate (HR), calculated from the photoplethysmography (PPG) sensor and sampled at 1 Hz. Audio signals were obtained at 16 kHz sampling rate and 16-bit encoding. Features from these signals are extracted to model bio-behavioral trajectories during a public speaking encounter with the proposed transient oscillation and quadratic polynomial models.

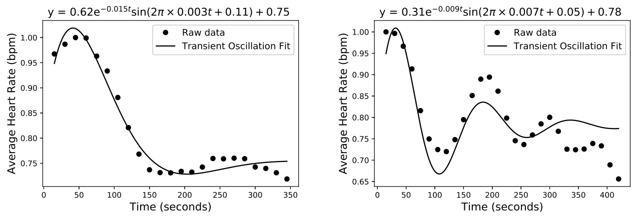
In order to capture individuals' trait characteristics, standardized surveys were administered before the public speaking task, which included the Trait-Scale of the Trait Anxiety Inventory (STAI) [15], Brief Fear of Negative Evaluation (BFNE) survey [20], and the Personal Report of Public Speaking Anxiety (PRPSA)[21]. The STAI and PRPSA measure general and communication-specific trait-anxiety, while the BFNE captures feelings of apprehensions about others' evaluation. These will be used to explore individual differences in terms of bio-behavioral trajectories.

4. PROPOSED METHOD

In this section, we describe the extraction of physiological and acoustic features, which will comprise the bio-behavioral time-series (Section 4.1). We then explain the proposed transient oscillation and quadratic fit models (Sections 4.2, 4.3).

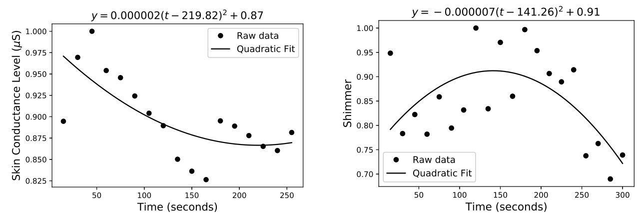
4.1. Feature Extraction

Noise cancellation and outlier removal were performed on the physiological data. Voice activity detection was applied in speech [10]. Bio-behavioral features include the mean skin conductance level (SCL), average heart rate, fundamental frequency (F0), jitter, and shimmer, extracted over 15 second analysis windows and scaled for each participant. In accordance to previous studies [13, 14, 10], we expect that the proposed bio-behavioral features are indicative of trait anxiety, as experienced in-the-moment throughout the public speaking session. Visual inspection of the resulting time-series indicated that HR times-series depicted oscillatory patterns, while the remaining signals followed a smoother quadratic polynomial shape. For this reason, we used the transient oscillation model (Section 4.2) for HR, and the quadratic polynomial model (Section 4.3) for SCL, F0, jitter, and shimmer.



(a) Participant 1 (b) Participant 2

Fig. 2. Transient oscillation model fit of heart rate time-series for two participants. Model parameters are presented in the equation on top of the figures.



(a) Participant 3 (b) Participant 4

Fig. 3. Quadratic polynomial fit depicting (a) habituation patterns of skin conductance levels; and (b) sensitization patterns of shimmer.

4.2. Transient Oscillation Model

The transient oscillation model represents oscillatory time-series through the following damped harmonic oscillation:

$$y(t) = Ae^{-\tau t} \sin(2\pi f t + \phi) + c \tag{1}$$

where $y(t)$ is the HR at time t, and A, τ, f, ϕ, and c correspond to amplitude, decay rate, oscillation frequency, initial phase, and bias. Large values of A and c indicate higher levels of HR and oscillations of larger amplitude. High values of τ depict a fast dampening oscillation. High values of f indicate oscillations of higher frequency, while large ϕ suggests that the first peak of the oscillation is closer to zero.

For each participant's HR time-series, the above parameters are obtained by fitting the data to (1) using a nonlinear least square regression. The approximate optimization process of the non-linear least square regression requires a good starting point for obtaining a good fit. The oscillation frequency f is initialized using the frequency obtained by counting the consecutive peaks and zero crossings. If this yielded an initialization of $f = 0$, we set $f = 0.01$ and $\phi = \frac{\pi}{2}$ to avoid losing the oscillatory part of the model. Phase ϕ is initialized using a small number, e.g., $\phi = 0.01$. In order to find a good initialization point for A_o, τ, and c_o, the given HR time-series is fitted into the exponential function:

$$y(t) = A_o e^{-\tau_o t} + c_o \tag{2}$$

whose parameters hold the same meaning as in 1. Optimization is performed through a non-linear least squares regression. Due to its lower complexity compared to (1) (i.e., 3 parameters to estimate), (2) can yield reliable starting points for A, τ, and c, based on which (1) will be optimized.

Five transient oscillation features are obtained per participant. Illustration of the proposed model using data from two participants is presented in Figure 2. Participant 2 depicts small value of τ, which indicates a slower dampening of the HR compared to Participant 1.

4.3. Quadratic Polynomial Model

A quadratic model is fitted into the bio-behavioral features which do not depict oscillatory patterns, i.e., mean SCL, F0, jitter, and shimmer:

$$y(t) = a(t - t_o)^2 + b \tag{3}$$

where $y(t)$ is the bio-behavioral time-series at time t, a refers is the rate of temporal change, t_o is the time when minima (or maxima) occurs, and b is the bias term. Positive values of a are likely to depict patterns of habituation in terms of the bio-behavioral time-series, while negative a indicates sensitization effects (Section 2, Figure 1). Higher a and b indicate higher fluctuation. On the other hand, t_o dictates how fast or slow bio-behavioral indices will reach a maximum. Lower t_o suggests that the bio-behavioral times series reaches a maximum faster in the case of sensitization. Least squares regression is used to fit each time-series into the model, resulting in 3 parameters per participant for each bio-behavioral index. Examples of bio-behavioral time-series and their quadratic polynomial fit are in Figure 3. Participant 3 depicts high positive values of a, reflecting habituation over time, while Participant 4 has negative a values alluding to sensitization.

4.4. Evaluation

We examine potential relationship between the proposed models and participants' trait characteristics by computing the Pearson's correlation between self-reports (i.e., STAI, PRPSA, BFNE; Section 3) and the model parameters (i.e., A, τ, f, ϕ, and c for the transient oscillation; a, b, and t_o for the quadratic polynomial; Sections 4.2, 4.3). We further trained a linear regression with a 10-fold cross-validation to estimate participants' trait scores using the model parameters as features. The total experiment is repeated 20 times, to decrease randomness by the 10-fold split. Since each participant corresponded to one sample, there is no contamination between the train and test set. We reported the Spearman's correlation between the actual and estimated trait characteristics, since the latter take integer values. As baseline, we employed a linear regression model whose features included the mean value of the bio-behavioral indices, computed over the public speaking session.

5. RESULTS

5.1. Relation between Individuals' Bio-Behavioral Time Trajectories and Trait Characteristics

Table 1 presents the Pearson's correlation coefficient (r) between the proposed model parameters and individuals' trait characteristics. For the transient oscillation model, the amplitude A and bias c are positively correlated with the trait

Table 1. Pearson's correlation coefficient between the bio-behavioral time-trajectory model parameters and individuals' trait characteristics (State-Trait Anxiety Inventory-STAI, Personal Report of Public Speaking Anxiety-PRPSA, Brief Fear of Negative Evaluation-BFNE).

Modality	Feature	Trait	Pearson's r
Heart rate	c	BFNE	0.31^{*}
		PRPSA	0.22
		STAI	0.28^{*}
	ϕ	BFNE	-0.23
		PRPSA	-0.27^{*}
		STAI	-0.23
	A	BFNE	0.12
		PRPSA	0.15
		STAI	0.22
Skin conductance level	a	BFNE	0.35^{**}
		PRPSA	0.30^{*}
		STAI	0.24
Fundamental frequency	b	BFNE	0.29^{*}
		PRPSA	0.21
		STAI	0.20
Jitter	t_o	BFNE	-0.19
		PRPSA	-0.26
		STAI	0.01

*: p <0.05, **: p <0.01

scores, indicating that individuals with higher trait anxiety depict higher amplitude oscillations in their HR trajectories compared to their low trait anxiety counterparts. Similarly, HR for individuals with high PRPSA reaches a maximum faster (i.e., lower ϕ values) compared to the ones with low PRPSA. For the quadratic polynomial model, the a parameter based on the mean SCL shows significant correlation with all trait scores. Large positive values of a indicate narrower polynomial curve, which might suggest that it takes a shorter time for bio-behavioral signals of individuals with higher trait anxiety to reach a minimum. The t_o parameter of the jitter time-series exhibits negative correlation PRPSA, which indicates that the maximum of the corresponding time-series is reached faster for individuals with higher trait communication anxiety.

5.2. Estimation of Individuals' Trait Characteristics

Since the physiological and acoustic trajectory measures exhibit significant correlations with some of the self-reported trait anxiety scores, we proceed to use them as features to estimate individuals' trait characteristics (Table 2). The proposed feature sets, which model time-trajectories within a public speaking session, tend to outperform the corresponding average measures of bio-behavioral features, used as our baseline. Using both physiological and acoustic features, our model achieves good predictive ability for the BFNE ($\rho = 0.33$, $p = 0.01$), while physiological features only provided an even better performance ($\rho = 0.35$, $p < 0.01$). Features of the transient oscillation model appear to perform very well, obtaining the highest Spearman's correlation for the BFNE estimation ($\rho = 0.37, p < 0.01$). On the other hand, the baseline model displays non-significant correlations when using separate groups of features, and low correlations approaching significance with both modalities ($\rho = 0.18$, $p = 0.2$).

Table 2. Spearman's correlation coefficient between actual and predicted trait scores using different bio-behavioral modalities, for proposed transient oscillatory and quadratic polynomial features, and the baseline average measures.

Modality	Trait	Our Method	Baseline
Heart Rate only	BFNE	0.37^{**}	0.04
	PRPSA	0.29^{*}	-0.88^{**}
	STAI	0.31^{*}	-0.13
Physiological	BFNE	0.35^{**}	0.08
	PRPSA	0.15	-0.12
	STAI	0.23	-0.07
Acoustic	BFNE	0.31^{*}	0.10
	PRPSA	0.11	0.12
	STAI	-0.04	-0.22
Physiological & Acoustic	BFNE	0.33^{*}	0.18
	PRPSA	0.16	0.11
	STAI	0.17	-0.23

*: p <0.05, **: p <0.01

6. DISCUSSION

The proposed transient oscillation and quadratic polynomial approaches can provide highly interpretable models to capture escalation, habituation, and sensitization patterns of state-anxiety within a public speaking session. The parameters of the corresponding models can further quantify the degree of such phenomena, yielding valuable insights into the generative physiological processes that occur during the public speaking encounter. Results indicate that our proposed models can capture temporal bio-behavioral patterns, which are significantly correlated with individuals' general and communication-based anxiety. Findings from this study can have valuable implications toward personalized models and time trajectories of state PSA during the public speaking encounter, which can help toward reliable prediction of state anxiety and in-the-moment interventions. Despite the encouraging results, the proposed method only uses data collected from up to 5 minutes of a public speaking task. The model scalability needs to be explored in large data collection settings. Also, the quadratic model can oversimplify existing longitudinal trajectories, while more sophisticated models might be more useful in this context.

7. CONCLUSION

We introduced transient oscillation and quadratic polynomial models to quantify bio-behavioral time trajectories of habituation, sensitization, and escalation during a public speaking task. Significant correlations between the proposed model parameters and individuals' general and communication-based anxiety trait characteristics have been identified. The proposed model parameters outperform a baseline model, whose features include commonly used mean values of the corresponding bio-behavioral signals, for automatically estimating trait characteristics. As part of our future work, we will focus on predicting in-the-moment anxiety using the trajectory features developed in this paper. We will further leverage additional sources of information from electrocardiogram (EEG) signals, and examine data-driven time trajectory models.

8. REFERENCES

[1] Franklin R Schneier, "Social anxiety disorder," *New England Journal of Medicine*, vol. 355, no. 10, pp. 1029–1036, 2006.

[2] Karen Kangas Dwyer and Marlina M Davidson, "Is public speaking really more feared than death?," *Communication Research Reports*, vol. 29, no. 2, pp. 99–107, 2012.

[3] National Collaborating Centre for Mental Health (UK et al., "Social anxiety disorder: recognition, assessment and treatment," British Psychological Society, 2013.

[4] John Augustine Daly, *Avoiding communication: Shyness, reticence, and communication apprehension*, Hampton Pr, 1997.

[5] Graham D Bodie, "A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety," *Communication education*, vol. 59, no. 1, pp. 70–105, 2010.

[6] Michael J Beatty, "Situational and predispositional correlates of public speaking anxiety," *Communication education*, vol. 37, no. 1, pp. 28–39, 1988.

[7] Ulf Dimberg, Mats Fredrikson, and Olle Lundquist, "Autonomic reactions to social and neutral stimuli in subjects high and low in public speaking fear," *Biological Psychology*, vol. 23, no. 3, pp. 223–233, 1986.

[8] Michael J Beatty and Ralph R Behnke, "Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance," *Human Communication Research*, vol. 18, no. 2, pp. 147–176, 1991.

[9] Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T Chittaranjan, Andrew T Campbell, Daniel Gatica-Perez, and Tanzeem Choudhury, "Stresssense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 351–360.

[10] Megha Yadav, Md Nazmus Sakib, Kexin Feng, Theodora Chaspari, and Amir Behzadan, "Virtual reality interfaces and population-specific models to mitigate public speaking anxiety," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019, pp. 1–7.

[11] Steven Booth-Butterfield and Malloy Gould, "The communication anxiety inventory: Validation of state-and context-communication apprehension," *Communication Quarterly*, vol. 34, no. 2, pp. 194–205, 1986.

[12] Paul L Witt, Kennaria C Brown, James B Roberts, Jessica Weisel, Chris R Sawyer, and Ralph R Behnke, "Somatic anxiety patterns before, during, and after giving a public speech," *Southern communication journal*, vol. 71, no. 1, pp. 87–100, 2006.

[13] X. Li, J. Tao, M.T. Johnson, J. Soltis, A. Savage, K.M. Leong, and J.D. Newman, "Stress and emotion classification using jitter and shimmer features," in *IEEE ICASSP*. IEEE, 2007, vol. 4, pp. IV–1081.

[14] T. Chaspari, S.C. Han, D. Bone, A.C. Timmons, L. Perrone, G. Margolin, and S.S. Narayanan, "Quantifying regulation mechanisms in dating couples through a dynamical systems model of acoustic and physiological arousal," *ICASSP*, 2017.

[15] Charles D Spielberger, "State-trait anxiety inventory," *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.

[16] Ligia Batrinca, Giota Stratou, Ari Shapiro, Louis-Philippe Morency, and Stefan Scherer, "Cicero-towards a multimodal virtual audience platform for public speaking training," in *International workshop on intelligent virtual agents*. Springer, 2013, pp. 116–128.

[17] Lei Chen, Gary Feng, Jilliam Joe, Chee Wee Leong, Christopher Kitchen, and Chong Min Lee, "Towards automated assessment of public speaking skills using multimodal cues," in *Proc. ICMI*. ACM, 2014, pp. 200–203.

[18] Andreas Schwerdtfeger, "Predicting autonomic reactivity to public speaking: don't get fixed on self-report data!," *International Journal of Psychophysiology*, vol. 52, no. 3, pp. 217–224, 2004.

[19] "Empatica e4," `https://www.empatica.com/`.

[20] Azadeh Tavoli, Mahdiyeh Melyani, Maryam Bakhtiari, Gholam Hossein Ghaedi, and Ali Montazeri, "The brief fear of negative evaluation scale (BFNE): translation and validation study of the iranian version," *BMC psychiatry*, vol. 9, no. 1, pp. 42, 2009.

[21] James C McCroskey, "Measures of communication-bound anxiety," 1970.