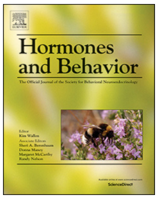
Hormones and Behavior 106 (2018) 52–61

Contents lists available at ScienceDirect

Hormones and Behavior

journal homepage: www.elsevier.com/locate/yhbeh



# Multimodal stress detection: Testing for covariation in vocal, hormonal and physiological responses to Trier Social Stress Test

Katarzyna Pisanski[a,b,*], Aleksander Kobylarek[c], Luba Jakubowska[d], Judyta Nowak[e], Amelia Walter[e], Kamil Błaszczyński[a], Magda Kasprzyk[a], Krystyna Łysenko[a], Irmina Sukiennik[e], Katarzyna Piątek[a], Tomasz Frackowiak[a], Piotr Sorokowski[a]

[a] *Institute of Psychology, University of Wrocław, Poland*
[b] *Mammal Vocal Communication and Cognition Research Group, School of Psychology, University of Sussex, Brighton, UK*
[c] *Institute of Pedagogy, University of Wrocław, Poland*
[d] *Department of Health Promotion, Faculty of Health Sciences, Wroclaw Medical University*
[e] *Department of Human Biology, University of Wrocław, Poland*

ARTICLE INFO

*Keywords:*
Stress
TSST
Voice pitch
Fundamental frequency
Nonverbal communication
Cortisol
Testosterone
Polygraph
Skin temperature
Hand movement

ABSTRACT

Examining the effects of acute stress across multiple modalities (behavioral, physiological, and endocrinological) can increase our understanding of the interplay among stress systems, and may improve the efficacy of stress detection. A multimodal approach also allows for verification of the biological stress response, which can vary between individuals due to myriad internal and external factors, thus allowing for reliable interpretation of behavioral markers of stress. Here, controlling for variables known to affect the magnitude of the stress response, we utilized the Trier Social Stress Test (TSST) to elicit an acute stress response in 80 healthy adult men and women. The TSST involves an interview-style oral presentation and critical social evaluation, and is highly effective in inducing psychosocial stress. Participants completed the study in individual 2 h sessions, during which we collected voice, polygraph and salivary hormone measures in baseline, stress, and relaxation phases. Our results show sizeable systematic increases in voice pitch (mean, minimum and variation in fundamental frequency, *F*0), hormone levels (cortisol) and decreases in skin temperature and hand movement during psychosocial stress, with striking similarities between men and women. However, cortisol and skin temperature only weakly predicted changes in voice pitch during stress, in either women or men, respectively. Thus, while our results provide compelling evidence that psychosocial stress manifests itself behaviorally by increasing voice pitch and its variability alongside simultaneous activation of physiological and endocrinological stress systems, our results also highlight a relatively weak degree of intra-individual 'response coherence' across these stress systems, with dissociations among different stress measures related most strongly to sex.

## 1. Introduction

Stress detection has been a central topic of research for many decades, and for good reason. Acute increases in stress levels can drastically affect behavior, memory, and cognitive reasoning (Lupien et al., 2009), and while potentially adaptive in the short-term (McEwen, 2007), repeated or chronic stress can suppress immunological, metabolic, and cardiovascular functioning (McEwen, 1998). The capacity to effectively measure intra-individual changes in stress levels from multiple sources (behavioral, physiological, and endocrinological), although challenging (Andrews et al., 2013), could improve the efficacy of stress detection, lead to a better understanding of mechanisms, promote stress reduction therapies, and would find a range of practical applications in medicine, forensics, and in the development of remote monitoring and voice recognition technologies (Campbell and Ehlert, 2012; Hellhammer and Schubert, 2012 for discussion).

Stress is a complex and multifaceted phenomenon, the definition of which depends on discipline and profession (Murray et al., 1996). In biology and psychology, stress can be defined as a disruption in physiological homeostasis caused by an external or internal stressor (Lazarus, 2000). Psychosocial stress in particular arises in response to real or perceived psychological or social threat, especially the threat of social evaluation (Dickerson and Kemeny, 2004). The brain's neural response to such a threat sets off a characteristic cascade of bodily

---

[*] Corresponding author at: University of Sussex, School of Psychology, Pevensey Building Room 2C2, Falmer BN1 9QH, UK.
*E-mail address:* kasiapisanski@gmail.com (K. Pisanski).

https://doi.org/10.1016/j.yhbeh.2018.08.014
Received 23 March 2018; Received in revised form 30 July 2018; Accepted 31 August 2018

reactions, including activation of the hypothalamus-pituitary-adrenal (HPA) axis and the sympathetic nervous system (SNS) (Andrews et al., 2013; Lupien et al., 2009). This typically causes the release of stress hormones such as cortisol (Takahashi et al., 2005), and causes changes in physiology (e.g., heart rate, skin temperature and perspiration). In some but not all people, the stress response also manifests itself behaviorally through changes in body movement (e.g., decreased hand movement, Vrij et al., 1997) and vocal production (e.g., decreased speech fluency, Buchanan et al., 2014; increased voice pitch, Pisanski et al., 2016).

Although many studies have examined speech under stress (Giddens et al., 2013; Kirchhübel et al., 2011 for reviews), these studies have produced inconsistent results. The most consistent finding is that voice pitch (the acoustic correlate of fundamental frequency, *F*0) increases under stress; however, many studies have failed to replicate even this most consistent effect (e.g., Dietrich and Abbott, 2012; Paulmann et al., 2016; Streeter et al., 1983; Tolkmitt and Scherer, 1986; Van Lierde et al., 2009). In large part, the mixed results of past work could be due to widely varying methodologies (e.g., real-world versus lab settings) as well as to a general lack of controls for inter-individual differences in the biological stress response (e.g., whereas some people show a strong cortisol response to stress, others show no response at all: Alberdi et al., 2016; Andrews et al., 2013; Berger et al., 1987; Kudielka et al., 2007). Indeed, recent work has shown that increases in cortisol levels prior to an oral examination predict corresponding increases in voice pitch among female students (Pisanski et al., 2016), indicating that voice changes may be observed only in those who demonstrate a physiological stress response, and that these individual differences should be taken into account in stress research.

Key factors known to influence the severity of HPA activation and the stress response across individuals include: sex (men generally show a stronger cortisol response than do women); lactation and breastfeeding (suppressants of the cortisol stress response); the use of drugs including nicotine, coffee, alcohol, steroids and hormonal contraception (stimulants or suppressants); the use of medications for mental disorders (particularly depression and anxiety); and various personality traits (Allen et al., 2017; Hellhammer and Schubert, 2012; Kudielka et al., 2007; Kudielka and Kirschbaum, 2005 for reviews). External factors such as time of day can also affect hormonal responses to psychosocial stress (Kudielka et al., 2004). Studies in search of behavioral indices of stress should control for these factors.

The goal of most studies examining behavioral correlates of stress is to offer an easy, cheap and non-invasive tool for measuring and monitoring stress, ideally in real time. However, to demonstrate that a behavioral measure (e.g., voice) is a reliable proxy of stress, it should be shown to vary systematically in response to a proven stressor, and should, in theory, also covary with underlying activity of the HPA axis (e.g., cortisol) and/or the SNS (e.g., heart rate, blood pressure, skin response) (Andrews et al., 2013). However, the nature of such covariation among stress systems is likely to be complex. Although relationships among stress systems have traditionally been thought of as sequentially positive (Lupien et al., 2009), recent evidence indicates the presence of some compensatory mechanisms in which, for example, the physiological stress response may be elevated following the suppression of the endocrinological stress response (Andrews et al., 2013).

Emerging evidence indicates an interplay among physiological, endocrinological and behavioral stress systems, which has led the authors of recent reviews to urge researchers to include multiple measures of stress in their empirical studies, emphasising that a multimodal system for stress detection is the way forward (Alberdi et al., 2016; Andrews et al., 2013). Yet, such studies remain scarce. For example, only approximately 6% of studies on subjective emotional reactions to stress in human subjects have examined covariation with HPA and SNS measures (Campbell and Ehlert, 2012). Most of these studies found no covariation or weak covariations among biological and subjective reactions to stress. Buchanan et al. (2014) examined covariation between biological markers of stress and speech fluency (e.g., word productivity); however, the present study is the first to examine covariation among endocrinological, physiological and nonverbal characteristics of the voice during stress.

Here, we utilized the Trier Social Stress Test (TSST) to elicit an acute stress response in a sample of 80 healthy adult men and women. The TSST is now the gold standard in research on the neurobiology of acute stress (Allen et al., 2017), offering a standardized yet ecologically valid procedure for evoking stress that involves an interview-style oral presentation and critical social evaluation. Compared to other lab procedures, the TSST's original design, used here, evokes the strongest HPA axis response (Dickerson and Kemeny, 2004 for meta-analysis), doubling or tripling cortisol levels in 70% to 80% of adults (Dickerson and Kemeny, 2004; Frisch et al., 2015). While its design offers a natural opportunity to examine variation (and covariation) in voice production under stress, only a handful of studies have analyzed TSST speech, of which the focus has been entirely on linguistic (e.g., word fluency: Buchanan et al., 2014) rather than nonverbal (e.g., pitch) properties of voice.

## 2. Methods

In individual sessions, we collected voice, polygraph and salivary hormone measures from 80 adults before, during, and after the TSST, and examined independent changes in each system as well as covariation among the systems in response to psychosocial stress. In addition to measuring voice pitch parameters (*F*0 mean, range and variation), we tested whether lesser studied voice perturbation and noise parameters (jitter, shimmer, and harmonics-to-noise ratio) change under stress. Hormone measures included free cortisol (biologically active fraction) as well as free testosterone, the latter of which influences the development of *F*0 and may interact with cortisol to influence pitch production in adulthood (Puts, 2016), and thus should be controlled. Finally, using a biofeedback modular polygraph with finger sensors, we tracked changes in participants' pulse, skin temperature, skin conductance and hand movement.

### *2.1. Participants*

Using meta-data computed by Dickerson and Kemeny (2004), we conducted an *a priori* power analysis in G*Power 3 (Faul et al., 2007) to determine the minimum sample size required to detect the established large effect (Cohen's $d = 0.85$, $f = 0.43$) in the cortisol stress response following a task with social-evaluative threat and uncontrollability (i.e., TSST) given alpha = 0.05, power = 0.9, two groups (men, women) and repeated measures. Power analysis indicated a required total minimum sample size of $N = 74$. Data from 80 healthy adults, including 47 women (aged 19–41, $M = 24.8$) and 33 men (aged 21–46, $M = 26.6$) were included in the study (see supplementary Table S1 for sample descriptives).

Participants were recruited from the general population in a large European city using posters and online ads advertising a study on nonverbal communication involving a social evaluative test. Interested individuals were pre-screened for criteria known to influence salivary hormone levels or the physiological stress response including: use of steroids or allergy medications containing corticosteroids; use of hormonal contraception within past three months; pregnant, lactating, or breastfeeding within past three months; smoking more than fifteen cigarettes a day; suffering from any disease of the mouth, ongoing infection or chronic illness. The sample size of 80 therefore excluded data from participants who, in the secondary screening questionnaire, reported using medication (hormonal contraception, $n = 1$; medication for a mental disorder, $n = 3$; steroids for thyroid disorder, $n = 1$) as well as pilot participants who took part for procedural training purposes ($n = 3$).