



# Workload Assessment Using Speech-Related Neck Surface Electromyography

Aaron Novstrup<sup>1</sup>(✉), Terrance Goan<sup>1</sup>, and James Heaton<sup>2</sup>

<sup>1</sup> Stottler Henke Associates, Inc., Seattle, WA, USA  
{anovstrup,goan}@stottlerhenke.com

<sup>2</sup> Massachusetts General Hospital,  
Center for Laryngeal Surgery and Voice Rehabilitation, Boston, MA, USA  
James.Heaton@mgh.harvard.edu

**Abstract.** This paper presents preliminary findings of an ongoing effort to evaluate the application of face and neck surface electromyography (sEMG) to real-time cognitive workload assessment. A retrospective analysis of anterior neck sEMG signals, recorded from 10 subjects during a time-pressured mental arithmetic task with verbal responses during a previous study by Stepp et al. [52], suggests that a measure known as neck intermuscular beta coherence (NIBcoh) may be sensitive to cognitive workload and/or error commission in tasks involving speech production, with sub-second temporal resolution. Specifically, the recent reanalysis indicates that subjects exhibited significantly lower NIBcoh when they produced incorrect verbal responses as compared to NIBcoh associated with correct responses. We discuss this promising application of NIBcoh within the context of our continuing research program and introduce future experiments that will shed light on the relationships among face and neck sEMG signals, task demands, performance, cognitive effort/strain, subjective workload measures, and other psychophysiological measures.

## 1 Introduction

Mental workload researchers have identified a variety of psychophysiological measures that have proven sensitive to cognitive task demands, including indices based on electrocardiography (ECG) [17,44], transcranial Doppler sonography (TCD) [39,53,61], electroencephalography (EEG) [4,14,21], functional near infrared [12,61], and eye tracking [2,5,27,37]. Electromyographic (EMG) measures, based on electrical potentials produced by motor units during muscle contraction, have also demonstrated sensitivity to task demands (e.g., [18,65]), yet *face and neck surface* EMG (sEMG) has received little attention in workload research despite the critical role of face and neck musculature in reflecting and expressing human mental/emotional state (whether through non-verbal cues or spoken expression). The lack of attention is perhaps not surprising given the

obtrusiveness of many existing sEMG sensor designs. However, recent advances in sEMG sensor design/miniaturation and in EMG signal processing technologies are paving the way for a new generation of unobtrusive sEMG sensors that will conform to the skin surface and minimize any negative impact on the wearer. For example, there are new, commercially available miniature differential sensors specifically designed for high-fidelity, wireless recording of facial sEMG signals [41, 42]. These advances have sparked new interest in the application of face/neck sEMG to cognitive workload assessment.

Face and neck sEMG may offer a unique window into human emotional state, complementing or even replacing previously studied psychophysiological measures as sensing modalities in the real-time assessment of cognitive workload. Many facial muscles are situated immediately below the skin surface and are thereby readily accessible for sEMG recording. They have high endurance and show little change in EMG power spectra across repeated facial contractions [57], and a link between emotional responses and high levels of cognitive strain has also been established (e.g., [20]). A study of error-related activity in the corrugator supercillii, a muscle of the upper face (medial eyebrow region) involved in facial expressions, has linked amplified EMG activity to error commission with less than 100 milliseconds of latency [35].

This paper reports findings of an exploratory reanalysis of an existing dataset, originally collected in a previous investigation by Stepp et al. into the modulation of a specific neck sEMG signal, known as neck intermuscular beta coherence (NIBcoh), by speech and non-speech behaviors [52]. In the present work, the authors reanalyzed this dataset to investigate the potential utility of NIBcoh in real-time cognitive workload assessment. Specifically, the sensitivity of NIBcoh to cognitive task demands and the relationships between NIBcoh and task performance were examined. This reanalysis was an initial step in an ongoing program of research intended to shed light on the potential application of face and neck surface EMG to the real-time assessment of cognitive workload.

The paper makes three primary contributions to the sciences of cognitive workload and of EMG-EMG coherence analysis. First, the reanalysis provides limited evidence that NIBcoh is sensitive to variations in task demand (or attention) across similar speech-related tasks. Second, it indicates that NIBcoh may be correlated with error commission within the context of a specific time-pressured mental arithmetic task requiring verbal responses. Finally, the findings offer validation for concerns raised in [45] regarding the common use of full-wave rectification in EMG-EMG coherence analysis.

The rest of the paper is organized as follows. Section 2 provides relevant background concerning EMG and EMG-EMG coherence measures, establishes existing support for a potential connection between cognitive workload and face/neck sEMG, generally, and NIBcoh more specifically, and offers possible advantages of face and neck sEMG as a real-time cognitive workload sensing modality. Section 3 describes the conditions in which the NIBcoh dataset was collected and the methods employed in the recent reanalysis. Section 4 presents and evaluates the results of the analysis. Finally, Sect. 5 concludes by summarizing the key findings and positioning them within the context of ongoing and future research.

## 2 Related Work

### 2.1 Human Mental Workload

The general concept of cognitive workload has been recognized and studied for at least 50 years, although no formal, standard definition of the construct has yet to emerge within the research community [9]. This paper defines cognitive workload similarly to the operational definition proposed by O'Donnell and Eggemeier [46], as “the fractional utilization of an individual’s limited cognitive resources at a particular moment.” As noted above, a number of psychophysiological measures have proven sensitive to variations in cognitive task demands, supporting their potential utility in the measurement of cognitive workload. However, several studies offer evidence of divergence among known psychophysiological workload indices [26,32,38,64]. The hypothesized causes for disassociation among these measures are varied and include both lack of *specificity* (i.e., some measures can be influenced by non-workload factors) and lack of *diagnosticity* (i.e., measures may reflect differing *aspects* of workload, consistent with multi-resource theories (e.g., [40,63]) in which workload is a multi-faceted construct arising from the capacity of and demand for multiple cognitive resources).

In addition to psychophysiological measures, a number of subjective instruments for cognitive workload assessment have been developed and validated [36]. These include both offline, retrospective instruments, such as the NASA Task Load Index (NASA-TLX) [25], the Workload Profile (WP) [56], and the Subjective Workload Assessment Technique (SWAT) [48], and online (real-time) techniques, such as the Instantaneous Self-Assessment of Workload (ISA) [30]. While subjective instruments may help to validate other measures or models of cognitive workload, a valid *objective* measure offers obvious relative benefits, including correspondence to reality uncontaminated by subjectivity and the avoidance of self-assessment procedures that may distract from primary tasks. The assessment of cognitive workload has found applications in a wide variety of human endeavors, including: manual assembly/manufacturing [43], medical education [7], air traffic control [15], and vehicle operation including that of trains [1,47,50], aircraft [8], and motor vehicles [4,55]. Therefore, the development of a robust, objective, real-time measure of cognitive workload can be expected to have widespread benefits.

### 2.2 Face and Neck Surface Electromyography

Previous research has demonstrated the utility of facial sEMG for recognizing and classifying emotional responses [11,54,58]. In contrast to image-based expression assessment, sEMG has the potential to identify rapid or slight facial expressions, including subtle muscle contractions *below the threshold necessary to generate visible changes in the surface contours of the face* [57]. Further, sEMG may provide greater sensitivity regarding the locations and magnitude of facial contraction when compared to image-based assessment. Automatic video quantification of facial movements is relatively difficult for features aside from

high-contrast tissue edges, whereas sEMG can discern a broad combination of facial muscle actions across both high- and low-contrast regions.

Beyond the identification of facial expressions, face/neck sEMG can quantify neuromuscular activity related to the vocalization and articulation of speech, *even for utterances that are not vocalized* (i.e., sub-vocal speech). Several research groups have demonstrated the utility of non-acoustic speech recognition technologies based on surface EMG signals [13,31,41,42], confirming that surface EMG provides ample speech-related information whether speech is spoken aloud or only “mouthed”. Because computer-based speech recognition is possible from EMG alone, it is reasonable to suggest that these signals may also change under varying cognitive load conditions in a manner resembling acoustic markers, such as those employed in voice stress analysis. The characteristics of face and neck sEMG make it particularly well suited to specific operational contexts. In noisy environments (such as an aircraft cockpit) in which an acoustic signal might be compromised, for instance, subtle acoustic features relating to cognitive load might be more readily gleaned from activity in the musculature involved in speech articulation than from degraded acoustic signals. The ability to detect subtle contractions associated with sub-vocal speech or slight (perhaps even involuntary) facial expressions could prove beneficial even in tasks that *do not involve significant amounts of speech*. Most facial expression assessment and eye blink tracking is done through image-based data collection, but this approach is problematic when individuals are freely moving and thereby changing head orientation relative to video capture sources. In addition, flight equipment such as helmets, glasses, and face-masks can preclude visualization of facial movements. In contrast, sEMG sensors can reside under headgear [3] or be incorporated into face masks, chin straps, etc. [10].

Finally, the dimensionality of sEMG, particularly when considering signals from multiple locations on the face and neck, offers a distinct advantage over many of the low-dimensional measures that are unobtrusive enough for operational use in real-time workload assessment. If face/neck sEMG can support speech recognition and the classification of emotional responses, perhaps it is also capable, alone or in concert with other measures, of distinguishing cognitive states that are conflated by other physiological signals. That is, it may help to overcome the lack of specificity and diagnosticity noted in Sect. 2.1.

Although sEMG appears to offer distinct advantages over other sensing modalities in some recording contexts, it also has potential drawbacks. Physiological measures typically require some degree of instrumentation, and even though sEMG is less prone to noise from movements and environmental sources compared to electrically weaker EEG, it is perhaps more cumbersome and prone to noise than other measures such as heart and respiration rate. Moreover, while modern sEMG recording systems do not require the use of conductive gels [41,42], the recorded skin surface should nevertheless be clean and free from hair that can impede adequate electrode contact. This precludes some potential speech-related neck/face recording locations in individuals with beards or when proper skin preparation is impractical (e.g., military field deployment, extremely

dirty or wet environments, etc.). In addition, the degree to which NIBcoh measurement is potentially degraded by motion artifact or environmental sources of electrical noise in the non-laboratory setting is still unknown.

### 2.3 Intermuscular Beta Coherence

The present study focuses on a particular measure, known as neck intermuscular beta coherence (NIBcoh), derived from the surface EMG signal at two anterior neck recording locations superior to (above) neck strap muscles involved in speech. Coherence, generally, is a frequency domain measure of the linear dependency or strength of coupling between two processes [24,62]. The coherence function,  $|R_{xy}(\lambda)|^2$ , can be defined as in Eq. 1 below, where  $f_{xx}$  represents the auto-spectra of a time series  $x(t)$ ,  $f_{yy}$  the auto-spectra of  $y(t)$ , and  $f_{xy}$  the cross-spectra of the two. *Intermuscular* coherence, the coherence between EMG signals, is a measure of the common presynaptic drive to motor neurons [6].

$$|R_{xy}(\lambda)|^2 = \frac{|f_{xy}(\lambda)|^2}{f_{xx}(\lambda)f_{yy}(\lambda)} \quad (1)$$

Muscle is thought to be driven by a number of different physiological oscillations at varying frequencies (see [23] for a review). The frequencies at which physiological oscillations occur appear to be characteristic of the function of distinct neural circuits and have been categorized into distinct bands such as alpha (8–13 Hz), beta (15–35 Hz), gamma (30–70 Hz), and others. It is generally thought that the beta and low gamma bands originate primarily from the primary motor cortex [23]. The beta band is typically associated with production of static motor tasks and is reduced with movement onset (e.g., [33]). Intermuscular coherence measurements reflect all oscillatory presynaptic drives to lower motoneurons. However, the intermuscular coherence in the beta band has been shown to be qualitatively similar to corticomuscular coherence, both in healthy individuals as well as in individuals with cortical myoclonus [6,33], supporting the hypothesis that *beta-band intermuscular coherence* is due to oscillatory drives originating in the motor cortex and is thereby likely influenced by cognitive state. The coherence of neuromuscular oscillations, whether measured through MEG-EMG, EEG-EMG or EMG-EMG, are affected by concurrent cognitive demands differently across the distinct frequency bands, making measures of coherence potentially useful for detecting changes in cognitive workload. For example, although alpha-band coherence is not dominant during motor tasks, it is known to increase when attention is drawn specifically to motor task execution [19,34]. Beta-band coherence is the dominant signal during synchronized oscillatory discharges of corticospinal or corticobulbar pathways onto lower motor neurons, and is likewise reduced when attention is divided or otherwise drawn away from the motor task at hand [29,34,52]. In addition, beta-band coherence is negatively correlated with motor output errors during concurrent cognitive tasks in young adults [28], suggesting that it may be predictive of both cognitive workload and motor performance in younger individuals. In contrast, beta-band

coherence is not necessarily correlated with motor performance in the elderly during divided-attention tasks (61–75 yr; [28]), perhaps due to reduced attentional resources [59] and motor coordination [60] with advancing age.

Measures of neck and face intermuscular beta coherence might be particularly well suited to real-time workload assessment given the bilateral symmetry of contraction typical for neck midline and facial muscles. For example, superficial facial muscles involved in speech articulation and neck midline strap muscles typically contract symmetrically across the right and left sides during speech and swallowing [41], providing an opportunity for coherence measurement during these synchronous contractions. Stepp and colleagues found that NIBcoh measured from ventral neck strap muscles (sternohyoid, sternothyroid, and thyrohyoid) can distinguish not only individuals with disordered (strained, hyperfunctional) versus healthy voice production [51], but also healthy individuals when they mimic a strained voice versus natural speech [52]. Vocal hyperfunction is associated with heightened speaking effort and anxiety [22], which may represent increased cognitive demand during speech and thereby reduce NIBcoh, regardless of whether the hyperfunction is pathological or mimicked. Stepp and colleagues [52] also found that NIBcoh decreases when speech is produced under divided attention (cognitive load imposed by rapid, backwards skip-counting), consistent with prior reports of divided attention effects on beta coherence in different motor systems [29, 34]. The goal of the present study was to re-examine the Stepp et al. [52] dataset of NIBcoh during their normal versus divided-attention speaking conditions, with the hypothesis that (1) their finding of reduced NIBcoh under the divided attention condition would be replicated, and (2) the commission of cognitive errors (miss counting) could be detected in the NIBcoh measure as errors occurred during their recordings of running speech (e.g., at a sub-second time resolution). If NIBcoh indeed correlates with cognitive errors, this measure would have important implications for real-time monitoring of cognitive load and performance.

## 3 Design and Methodology

### 3.1 Data Collection Procedures

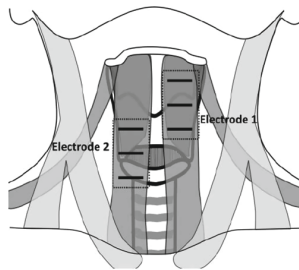
The dataset analyzed in this study consists of simultaneous neck surface EMG (sEMG) and acoustic signals recorded during an earlier investigation by Stepp et al. into the modulation of neck intermuscular beta coherence (NIBcoh) by speech and non-speech behaviors [52]. The signals were recorded under a variety of speech and non-speech task conditions, including a normal speech condition involving both spontaneous and scripted speech and a “divided-attention” condition in which participants were instructed to rapidly skip-count backwards from 100 by 7s. In the present study, these data were reanalyzed (as detailed in Sects. 3.2 and 3.3) to explore the relationships among the neck sEMG signals, the acoustic signal, task demands, and task performance in order to shed more light on the possible relationship between neck sEMG and mental workload. Because this research is ultimately focused on *real-time* workload assessment,

the previous analysis was also extended by considering *time-varying* measures and not only summary statistics over the entire time series.

**Participants.** The participants were ten (10) vocally healthy female volunteers (mean age: 25 years, standard deviation: 2.6 years). They reported no complaints related to their voice, and no abnormal pathology of the larynx was observed during standard digital video endoscopy with stroboscopy performed by a certified speech-language pathologist (SLP). Informed consent was obtained from all participants in compliance with the Institutional Review Board of the Massachusetts General Hospital.

**Recording Procedures.** As reported by Stepp et al. [52], simultaneous neck sEMG and acoustic signals from a lavalier microphone (Sennheiser MKE2-P-K, Wedemark, Germany) were filtered and digitally recorded at 20 kHz with Delsys hardware (Bagnoli Desktop System, Boston, MA) and software (EMGworks 3.3). The neck of each participant was prepared for electrode placement by cleaning the neck surface with an alcohol pad and “peeling” (exfoliating) with tape to reduce electrode-skin impedance, DC voltages, and motion artifacts. Neck sEMG was recorded with two Delsys 3.1 double differential surface electrodes placed on the neck surface, parallel to underlying muscle fibers. Each electrode consisted of three 10-mm silver bars with interbar distances of 10 mm. Double differential electrodes were chosen instead of single differential electrodes in order to increase spatial selectivity and to minimize electrical cross-talk between the two electrodes.

The two electrodes were placed on the right and left anterior neck surface, as depicted by the schematic in Fig. 1. Electrode 1 was centered approximately 1 cm lateral to the neck midline, as far superior as was possible without impeding the jaw opening, superficial to fibers of the thyrohyoid and sternohyoid muscles, and to some degree the omohyoid. Electrode 2 was centered vertically on the gap between the cricoid and thyroid cartilages of the larynx, and centered 1 cm lateral to the midline contralateral to Electrode 1, superficial to the cricothyroid,



**Fig. 1.** sEMG electrode placement [52]. Copyright 2011 by the American Speech-Language-Hearing Association. Reprinted with permission.

sternothyroid, and sternohyoid muscles. However, based on previous examinations of sEMG recordings during pitch glides [51], it is doubtful that cricothyroid contraction contributed much energy to the sEMG due to its relatively deep position. The platysma muscle likely contributed to some degree to the activity recorded at both electrode locations. A ground electrode was placed on the superior aspect of the participant's left shoulder. The sEMG recordings were pre-amplified and filtered using the Delsys Bagnoli system set to a gain of 1,000, with a bandpass filter with roll-off frequencies of 20 Hz and 450 Hz. All recordings were monitored by the experimenters in real time to ensure signal integrity, and no recordings included movement artifacts.

**Tasks.** Participants completed eleven separate speech and non-speech tasks, broadly organized into six task conditions. Only two conditions are relevant to the present study, however: a “normal speech” condition and a “divided attention” condition. The normal speech condition consisted of two tasks—a scripted task in which participants read “The Rainbow Passage” [16], and a spontaneous speech task in which participants produced speech spontaneously in response to a variety of available prompts, selected by participants (e.g., “What did you do last weekend?”). The Rainbow Passage was typically produced for 30–45 s. Spontaneous speech samples were approximately 1 min in length. No participant had any problems completing these speech tasks correctly. In order to collect speech under divided attention, participants were given 60 s to count backwards from 100, aloud, as quickly as possible in decrements of 7. These recordings were typically approximately 45 s in length. Participants uniformly reported this task as difficult, but all were able to produce continuous speech during the recording. The primary cognitive demand in this task is a (non-verbal) one imposed by time pressured arithmetic computation. The production of verbal responses, which imposes modest demands for linguistic processing and motor control resources, can be considered a secondary task. Since EMG-EMG coherence measures have been observed to decrease when attention is diverted from the motor task involving the instrumented muscle, the authors hypothesized that NIBcoh would decrease in response to increased demands of the primary, mental arithmetic task. From this perspective, NIBcoh was expected to function as a measure of secondary task attention.

### 3.2 Data Analysis

The original data consisted of discrete multivariate time series for ten subjects under five speech-related conditions and one non-speech condition, sampled at a rate of 20 kHz. Each time series included EMG variables from the two anterior neck surface recording locations depicted in Fig. 1 and an acoustic variable. From these “raw” time series, the authors derived several dependent EMG and acoustic measures and down-sampled to a rate of approximately 1.83 Hz (i.e., three samples for every 32,768 samples of the “raw” time series). The derived EMG-based variables were: NIBcoh, average magnitude, and gradient. The acoustic variables



were: average amplitude, peak amplitude, spectral roll-off, cepstral peak prominence, and sound intensity.

Two versions of the intermuscular beta coherence measure were computed, with and without full-wave rectification (NIBcoh-rect and NIBcoh, respectively) of the EMG signals. It is common practice in EMG-EMG coherence analysis to apply full-wave rectification as an EMG pre-processing step, and this step was performed in the research in which the NIBcoh dataset originated [52]. Other work, however, has called this practice into question, demonstrating that full-wave rectification may impair the identification of common oscillatory inputs to muscle pairs [45]. Therefore, the present study experimented with both rectified and unrectified EMG signals.

The NIBcoh and NIBcoh-rect time series were computed as follows. First, any DC offset was removed from the raw sEMG signals. In the case of the NIBcoh-rect measure, the resulting signals were then full-wave rectified. The signals were segmented by sliding a 16,384-point ( $\approx 820$  ms) rectangular window over each of the resulting EMG-EMG bivariate time series, with 50% overlap. Coherence between the two EMG signals, as defined in Eq. 1, was then estimated within each rectangular window using Welch’s overlapped averaged periodogram method [62], with sliding 8,192-point ( $\approx 410$  ms) Hamming windows, a 8,192-point fast Fourier transform, and 50% overlap (i.e., three Hamming windows per rectangular segment). Finally, the beta-band coherence values were computed by averaging the coherence values over the 15–35 Hz frequency range. Based on the findings of Neto and Christou [45] suggesting that oscillatory activity in the 100–150 Hz frequency band of the unrectified signal may drive variations in the beta band of rectified EMG signals, coherence in the 100–150 Hz band was also computed.

In order to associate the acoustic and sEMG signals with time-varying performance metrics, the acoustic signals for the divided-attention condition were manually annotated with labels indicating the participants’ verbal responses—each interval  $i$ , terminated by the completion of a response, was labeled with the number uttered  $l_i$ . Performance in the backwards-skip-counting task is characterized by both speed and accuracy. Accuracy is quantified as a function of error commission, with “errors” defined relative to the most recent element that a subject produced (e.g., 80 was regarded as the correct successor to 87, despite 87 not being an element of the correct sequence) to avoid an error-compounding effect. Speed is quantified by the duration of time required for a subject to produce each element of the sequence (i.e., response time). From the manual response annotations, discrete time series were generated capturing the error commission  $\epsilon$  and response time  $r$  performance metrics. Specifically, the value of the error commission indicator variable  $\epsilon_i$  for a given labeled interval  $i$  reflects whether the response label  $l_i$  for the interval is 7 less than the response label  $l_{i-1}$  for the previous labeled interval. The value of the response time variable for a given labeled interval is simply the duration of the interval. That is:

$$\epsilon_i = \begin{cases} 1 & \text{if } l_i \neq l_{i-1} - 7 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$r_i = \text{len}(i)$$

The label boundaries did not, in general, align with the  $\approx 820$  ms segment boundaries used to derive down-sampled time series from the raw EMG and acoustic signals. Therefore, a time-weighted average of the down-sampled time series over each labeled interval was computed to assign a single value of each dependent variable (e.g., NIBcoh) to the interval.

### 3.3 Statistical Analysis

As an initial step in the present analysis, the authors sought to determine how well a key finding in [52] held up in light of Neto and Christou’s [45] criticism of full-wave rectification in EMG-EMG coherence analysis—specifically, Stepp et al. [52] had found a significant effect of task condition on neck intermuscular beta coherence. This result was re-examined by using ANOVA to quantify the effect of condition on four EMG-EMG coherence measures: NIBcoh, NIBcoh-rect, and the corresponding intermuscular coherence measures for the 100–150 Hz frequency band. For consistency with the methods employed in [52], coherence was estimated over each signal as a whole with Welch’s overlapped averaged periodogram method [62], with a sliding 16,384-point Hamming window, 16,384-point FFT, and 50% overlap. The recent analysis differs from that of [52] in that their two-factor ANOVA was replaced with a more conservative one-factor repeated measures ANOVA, which makes weaker independence assumptions. The results of this analysis are reported in Table 1, below.

Since the authors’ primary interest was in evaluating the coherence measure’s utility in workload assessment, a post hoc two-tailed  $t$ -test was performed to contrast the cognitively demanding divided-attention condition and the normal speech condition under the assumptions of the ANOVA model, as a means to evaluate the sensitivity of NIBcoh to varying task demands. The relationships between the performance measures for the backwards skip-counting task (defined by Eq. 2) and the measures derived from the EMG and acoustic signals were then investigated. Specifically, associations between response time and each of the EMG/acoustic variables were evaluated by using Student’s  $t$ -tests to test the null hypotheses that each Pearson’s product-moment correlation coefficient was 0 (i.e.,  $H_0 : \rho_{r,v} = 0$ , where  $\rho_{r,v}$  denotes the correlation between response time ( $r$ ) and an EMG/acoustic variable  $v$ ). Similarly, Student’s two sample  $t$ -tests were used to evaluate the hypotheses that the distribution of each EMG/acoustic variable had unequal means for correct versus incorrect responses (i.e.,  $H_0 : \bar{v}_{\epsilon=1} = \bar{v}_{\epsilon=0}$ ). The results of these statistical tests are shown in Tables 2 and 3, in Sect. 4 below, along with the estimated correlation coefficients and differences in means. The  $t$ -tests had 123 degrees of freedom, corresponding to 125 observations (i.e., “responses”) across the 10 participants. The reported  $p$ -values are *not* adjusted for multiple comparisons, since any such adjustment could

itself be misleading due to correlations among several of the variables. In order to statistically control for substantial variability across subjects, both in terms of performance on the skip-counting task and in terms of the EMG/acoustic measures, the EMG/acoustic variables were normalized before computing  $t$ -values and correlation coefficients. Specifically, given a value  $x$  of a variable for subject  $s$  and the within-subject sample mean  $\mu_s$  and standard deviation  $\sigma_s$  of that variable over the normal speech condition, the standardized value  $z(x)$  was computed using Eq. 3. The statistical tests thus reflect the “effects” of performance on the other measures *relative to each subject’s “baseline”* from the normal speech condition.

$$z(x) = \frac{x - \mu_s}{\sigma_s} \quad (3)$$

The analysis of the relationships between performance and the EMG/acoustic variables is motivated by the hypothesis that within-subject variations in performance on the skip-counting task result, at least in part, from variations in the difficulty of the task. Given the nature of the task, within-subject variations in cognitive demand might reasonably be expected to be quite small. However, some such variation may arise from differences in the arithmetic problems that each participant encountered while completing the task, especially in subjects with less developed mental arithmetic skills. Some subjects may find it easier to compute  $100 - 7 = 93$  than to compute  $93 - 7 = 86$ , for instance. Whether this was actually the case was examined by comparing mean response times for decrementing numbers with a ones digit of 7, 8, or 9 (which can be computed without regard to the tens digit) versus other numbers (which require “borrowing” from the tens digit). Further, participants were encouraged to count as quickly as possible and to maintain continuous speech during the divided-attention task, which might be expected to lead them to sacrifice accuracy for speed and thereby to experience amplified within-subject variations in cognitive demands. This hypothesis was examined by analyzing the distributions of errors and response times and by comparing mean response times for correct versus incorrect responses within and across subjects. If within-subject variations in performance during the task can be plausibly connected to time-varying cognitive demands, then it is plausible that any observed relationship between performance and a physiological indicator can be interpreted as evidence of a possible relationship between that indicator and cognitive workload (i.e., that task performance acts as a proxy for cognitive workload within the context of this task).

## 4 Results and Evaluation

Replication analysis (Table 1) confirmed the finding in [52] that the experimental conditions had a significant effect on neck intermuscular beta coherence, despite the use of a more conservative statistical test than that employed by Stepp et al. The results also lend credence to the concerns of [45] regarding the common use

of full-wave rectification in EMG-EMG coherence analysis. Rectifying the EMG signals prior to estimating coherence appears to dilute the estimated difference in coherence across conditions, suggesting that full-wave rectification may result in a harmful loss of information—note the difference between the effect sizes and p-values between the unrectified and full-wave rectified beta-band EMG-EMG coherence signals. No significant effect of the task conditions on neck intermuscular coherence in the 100–150 Hz frequency band was found.

**Table 1.** Summary of replication analysis. Effect of task condition on EMG-EMG coherence measures.

Coherence measure	Size of effect (generalized $\eta^2$ )	p-value
Full-wave rectified		
Beta band	0.120	0.0715
100–150 Hz	0.0892	0.0905
Unrectified		
Beta band	0.176	<b>0.0194<sup>†</sup></b>
100–150 Hz	0.0651	0.146

<sup>†</sup>Denotes a Huynh-Feldt sphericity-corrected p-value, applied because Mauchly’s test for sphericity failed to reach a 5% significance level.

A post hoc comparison of the normal speech and divided-attention conditions revealed a significant difference between the two ( $p < 0.001$ ). The linear model corresponding to the repeated measures ANOVA implied that NIBcoh was lower by an average of 0.0956 in the divided-attention condition, with a standard error of 0.0290. This difference *may* indicate sensitivity of NIBcoh to the change in cognitive demands between the two conditions. Statistical analysis of the performance metrics for the backwards skip-counting task did seem to support the use of these metrics as proxies for cognitive workload within the context of this task. First, it was found that participants tended to compute the successors for numbers with a ones digit of 7, 8, or 9 faster on average, with a mean time of 1.9 s, than for other numbers (3.5 s). This result is consistent with the hypothesis that cognitive demands vary between these two conditions and that the varying cognitive demands are reflected in task performance. Additionally, the data are consistent with the expectation that subjects sacrificed accuracy for speed—time pressure that would tend to amplify the effects of within-subject variations in cognitive demands on accuracy. A full 36% (45 out of 125) of the subjects’ responses were erroneous, despite the presumed simplicity of the task. The mean response time was 3.1 s with a standard deviation of 2.8, and the response time distribution was heavily left-skewed with a median response time of only 2.1 s. Furthermore, longer response times were strongly associated with *incorrect* responses both between and within most subjects, indicating that participants took longer to respond when they were struggling but that delayed

responses did not generally result in greater accuracy. Analysis of cross-subject variation suggests that standardization of the EMG and acoustic variables, using Eq. 3 as described in the previous section, was justified. Subjects were found to vary substantially, both in terms of their performance on the skip-counting task (the total number of errors committed by each subject ranged from 0 to 10, with a mean of 5 and a median of 4) and on the EMG/acoustic measures (e.g., the mean within-subject standard deviation of NIBcoh was 0.104, while the overall mean was 0.446, overall standard deviation was 0.116, and standard deviation of within-subject means was 0.0497).

Table 2 shows the results of the  $t$ -tests comparing the distributions of the standardized EMG and acoustic variables across correct and incorrect responses. NIBcoh, EMG magnitude, sound intensity, average sound amplitude, peak amplitude, and cepstral peak prominence all exhibited significant ( $\alpha = 0.05$ ) differences in estimated means across correct and incorrect responses. The NIBcoh-rect, EMG gradient, and acoustic spectral roll-off measures showed no significant difference across correct and incorrect responses. NIBcoh associated with incorrect responses was lower than NIBcoh associated with correct responses, as may be expected if workload or simply reduced attention to speech reduces intermuscular coherence of oscillatory drives in speech-related muscles during speech. The results for several acoustic measures frequently used in voice stress analysis also accord with expectations.

**Table 2.** Difference in mean standardized features values for correct vs. incorrect responses.

Measure	Estimated difference	$t$ -statistic	$p$ -value
NIBcoh	0.505	2.81	<b>&lt;0.006</b>
NIBcoh-rect	0.078	0.548	0.585
EMG magnitude (site 1)	0.432	-2.22	<b>0.028</b>
EMG magnitude (site 2)	0.532	-2.50	<b>0.014</b>
EMG gradient (site 1)	0.120	-0.772	0.44
EMG gradient (site 2)	0.064	-0.414	0.68
Mean acoustic amplitude	0.852	-2.35	<b>0.020</b>
Peak acoustic amplitude	2.52	-2.96	<b>&lt;0.004</b>
Sound intensity	0.345	-1.97	0.051
Spectral roll-off	0.276	1.21	0.230
Cepstral peak prominence	0.258	-1.67	0.097

Correlations between response time and the EMG/acoustic variables are shown in Table 3. NIBcoh, peak acoustic amplitude, sound intensity, and cepstral peak prominence exhibited significant ( $\alpha = 0.05$ ) associations with response time.

**Table 3.** Correlations between response time and the standardized EMG and acoustic features.

Measure	Pearson's $r$	$t$ -statistic	$p$ -value
NIBcoh	-0.196	-2.22	<b>0.028</b>
NIBcoh-rect	-0.078	-0.868	0.387
EMG magnitude (site 1)	-0.176	-1.98	0.050
EMG magnitude (site 2)	0.008	0.084	0.933
EMG gradient (site 1)	0.076	0.849	0.397
EMG gradient (site 2)	0.111	1.24	0.218
Mean acoustic amplitude	-0.139	-1.56	0.122
Peak acoustic amplitude	0.185	2.08	<b>0.040</b>
Sound intensity	-0.223	-2.54	<b>0.012</b>
Spectral roll-off	-0.005	-0.053	0.958
Cepstral peak prominence	-0.219	-2.49	<b>0.014</b>

#### 4.1 Caveats and Limitations

While the results are promising, several caveats must be acknowledged and the exploratory nature of the analysis must be stressed. Because the original experiment in which the data were collected was designed to shed light on the modulation of NIBcoh by speech and non-speech behaviors and *not* to explore the sensitivity of NIBcoh to cognitive demands, the recent reanalysis was necessarily ad hoc. The findings should therefore be considered only suggestive rather than conclusive—robust conclusions will require well controlled experiments in which cognitive demands are manipulated directly (Sect. 5 briefly describes such an experiment, which the authors plan to conduct in the near future).

The small sample size of only 10 participants and the primary focus on the (single) backwards skip-counting task further limit the generalizability of the results. In particular, although the task did require mental arithmetic, the verbal response format makes it impossible to determine from this study whether the NIBcoh measure may offer any insight into cognitive workload in *non-speech-involving tasks*. Additionally, it was found that the statistical hypothesis tests were quite sensitive to outliers—for instance, the omission of one subject in particular, who rapidly produced the entire sequence without errors, changes the  $p$ -value for the  $t$ -test comparing mean NIBcoh in correct versus incorrect responses from less than 0.006 to 0.074. This lack of robustness underscores the limits of this reanalysis and justifies cautious optimism in interpreting the results.

Finally, it must be noted that several of the EMG and acoustic measures were themselves somewhat correlated (e.g., NIBcoh and peak acoustic amplitude), and the design of the original experiment makes it impossible to infer the causal factors that underlie these correlations. Were NIBcoh and peak acoustic amplitude correlated because of a mutual causal relationship to cognitive

demands / workload? It seems intuitively likely that the effect on peak acoustic amplitude is an artifact of the experimental conditions, rather than an indication of any general utility in predicting workload or error commission. Bursts of nervous laughter accompanying high workload might, for instance, contribute to high peak amplitudes in the specific conditions of this experiment, but one would not expect to find the same effect in other settings. Could such an artifact also explain lower intermuscular beta coherence associated with erroneous responses? Questions such as these cannot be addressed adequately with the existing data.

## 5 Conclusions and Future Work

The data reanalysis reported herein offers some evidence bearing on the utility of neck surface EMG for detecting cognitive strain in real time. Specifically, the analysis demonstrates that time-varying EMG-derived measures, with a sub-second temporal resolution, are correlated with error commission and response time in a backwards skip-counting task. Although the task exhibits only mild variations in task difficulty over the sub-problems that comprise the task, the analysis indicates that a time-varying NIBcoh measure was lower by about half a standard deviation on average during intervals in which subjects produced incorrect responses. Further research is necessary to confirm the effect in a well controlled context, determine whether it is due to a causal relationship with error commission, cognitive demands, and/or workload, investigate whether the effect is limited to tasks involving speech, and generalize the work to other EMG sensing locations on the face/neck surface where other physiological responses to variations in cognitive workload may be detected. A planned study, commencing in 2018, will establish more conclusively whether intermuscular beta coherence or other measures derived from face and neck sEMG signals are sensitive to cognitive task demands by recording these, and other, psychophysiological signals while participants complete tasks with varying levels of difficulty in the NASA Multi-Attribute Task Battery (MATB) [49]. The research will investigate multiple EMG sensing locations on the face and neck surface, relating to muscles involved in facial expression, mastication/jaw clenching, speech articulation, and voice production, and a psychometric analysis will establish relationships to more conventional workload indicators (including subjective workload as measured by the NASA Task Load Index administered within the MATB).

The experiments will employ a novel protocol designed to establish whether a perceived risk of aversive consequences affects the measured psychophysiological responses to cognitive task demands. Specifically, after half of the task blocks, identified to participants *before* and *during* each such block, a series of mildly noxious electrical stimuli will be delivered to participants, with the number of stimuli ostensibly associated with task performance—but *actually* determined by the (manipulated) level of task demand in the block. The protocol will thus test how the presence of perceived risks mediates the relationship between task demands and psychophysiological responses. It is our hope that the technique of employing aversive consequences in order to elevate physiological responses

to workload will resolve two key challenges for workload researchers—namely, the risk that muted responses may lead to Type I errors in laboratory studies and the problem that laboratory models may not transfer well into operational environments.

**Acknowledgements.** This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. D17PC00119 (Distribution Statement: approved for public release; distribution unlimited). The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

## References

1. Balfe, N., Crowley, K., Smith, B., Longo, L.: Estimation of train driver workload: extracting taskload measures from on-train-data-recorders. In: Longo, L., Leva, M.C. (eds.) H-WORKLOAD 2017. CCIS, vol. 726, pp. 106–119. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_7](https://doi.org/10.1007/978-3-319-61061-0_7)
2. Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* **91**(2), 276–292 (1982). <https://doi.org/10.1037/0033-2909.91.2.276>
3. Betts, B., Binsted, K., Jorgensen, C.: Small-vocabulary speech recognition using surface electromyography. *Interact. Comput.* **18**(6), 1242–1259 (2006). <https://doi.org/10.1016/j.intcom.2006.08.012>
4. Borghini, G., et al.: Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, pp. 6442–6445 (2012). <https://doi.org/10.1109/EMBC.2012.6347469>
5. Brookings, J.B., Wilson, G.F., Swain, C.R.: Psychophysiological responses to changes in workload during simulated air traffic control. *Biol. Psychol.* **42**, 361–377 (1996). [https://doi.org/10.1016/0301-0511\(95\)05167-8](https://doi.org/10.1016/0301-0511(95)05167-8)
6. Brown, P., Farmer, S.F., Halliday, D.M., Marsden, J., Rosenberg, J.R.: Coherent cortical and muscle discharge in cortical myoclonus. *Brain* **122**(3), 461–472 (1999). <https://doi.org/10.1093/brain/122.3.461>
7. Byrne, A.: Mental workload as an outcome in medical education. In: Longo, L., Leva, M.C. (eds.) H-WORKLOAD 2017. CCIS, vol. 726, pp. 187–197. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_12](https://doi.org/10.1007/978-3-319-61061-0_12)
8. Cahill, J., et al.: Adaptive automation and the third pilot. In: Aircraft Technology. IntechOpen (2018)
9. Cain, B.: A review of the mental workload literature. In: Defence Research and Development Toronto (Canada), pp. 4-1–4-34 (2007). <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA474193>
10. Chan, A., Englehart, K., Hudgins, B., Lovely, D.: Myo-electric signals to augment speech recognition. *Med. Biol. Eng. Comput.* **39**(4), 500–504 (2001). <https://doi.org/10.1007/BF02345373>
11. Cheng, B., Liu, G.: Emotion recognition from surface EMG signal using wavelet transform and neural network. In: Proceedings of The 2nd International Conference on Bioinformatics and Biomedical Engineering (ICBBE), pp. 1363–1366 (2008). <https://doi.org/10.1109/ICBBE.2008.670>



12. De Joux, N., Russell, P.N., Helton, W.S.: A functional near-infrared spectroscopy study of sustained attention to local and global target features. *Brain Cogn.* **81**(3), 370–375 (2013). <https://doi.org/10.1016/j.bandc.2012.12.003>
13. Denby, B., Schultz, T., Honda, K., Hueber, T., Gilbert, J., Brumberg, J.: Silent speech interfaces. *Speech Commun.* **52**(4), 270–287 (2010). <https://doi.org/10.1016/j.specom.2009.08.002>
14. Dussault, C., Jouanin, J.C., Philippe, M., Guezennec, C.Y.: EEG and ECG changes during simulator operation reflect mental workload and vigilance. *Aviat. Space Environ. Med.* **76**(4), 344–351 (2005)
15. Edwards, T., Martin, L., Bienert, N., Mercer, J.: The relationship between workload and performance in air traffic control: exploring the influence of levels of automation and variation in task demand. In: Longo, L., Leva, M.C. (eds.) *H-WORKLOAD 2017*. CCIS, vol. 726, pp. 120–139. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_8](https://doi.org/10.1007/978-3-319-61061-0_8)
16. Fairbanks, G.: The rainbow passage. In: *Voice and Articulation Drillbook*, vol. 2 (1960)
17. Fairclough, S.H., Venables, L., Tattersall, A.: The influence of task demand and learning on the psychophysiological response. *Int. J. Psychophysiol.* **56**(2), 171–184 (2005). <https://doi.org/10.1016/j.ijpsycho.2004.11.003>
18. Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A.R., Miyake, S.: Effects of mental workload on physiological and subjective responses during traffic density monitoring: a field study. *Appl. Ergon.* **52**, 95–103 (2016). <https://doi.org/10.1016/j.apergo.2015.07.009>
19. Feige, B., Aertsen, A.D., Kristeva-Feige, R.: Dynamic synchronization between multiple cortical motor areas and muscle activity in phasic voluntary movements. *J. Neurophysiol.* **84**(5), 2622–2629 (2000). <https://doi.org/10.1152/jn.2000.84.5.2622>
20. Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., McLaughlin, K.: Emotion, cognitive load and learning outcomes during simulation training. *Med. Educ.* **46**(11), 1055–1062 (2012). <https://doi.org/10.1111/j.1365-2923.2012.04355.x>
21. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* **4**(1–2), 113–131 (2003). <https://doi.org/10.1080/14639220210159717>
22. Goldman, S.L., Hargrave, J., Hillman, R.E., Holmberg, E., Gress, C.: Stress, anxiety, somatic complaints, and voice use in women with vocal nodules: preliminary findings. *Am. J. Speech-Lang. Pathol.* **5**(1), 44–54 (1996). <https://doi.org/10.1044/1058-0360.0501.44>
23. Grosse, P., Cassidy, M.J., Brown, P.: EEG-EMG. MEG-EMG and EMG-EMG frequency analysis: physiological principles and clinical applications (2002). [https://doi.org/10.1016/S1388-2457\(02\)00223-7](https://doi.org/10.1016/S1388-2457(02)00223-7)
24. Halliday, D., Rosenberg, J., Amjad, A., Breeze, P., Conway, B., Farmer, S.: A framework for the analysis of mixed time series/point process data: theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. *Prog. Biophys. Mol. Biol.* **64**(2–3), 237–278 (1995)
25. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv. Psychol.* **52**(C), 139–183 (1988). [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
26. Horrey, W.J., Lesch, M.F., Garabet, A.: Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions. *J. Saf. Res.* **40**(1), 7–12 (2009). <https://doi.org/10.1016/j.jsr.2008.10.011>

27. Jacob, R.J., Karn, K.S.: Eye tracking in human-computer interaction and usability research. Ready to deliver the promises. In: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, pp. 531–553. Elsevier, New York (2003). <https://doi.org/10.1016/B978-044451020-4/50031-1>
28. Johnson, A.N., Shinohara, M.: Corticomuscular coherence with and without additional task in the elderly. *J. Appl. Physiol.* **112**(6), 970–981 (2012). <https://doi.org/10.1152/jappphysiol.01079.2011>
29. Johnson, A.N., Wheaton, L.A., Shinohara, M.: Attenuation of corticomuscular coherence with additional motor or non-motor task. *Clin. Neurophysiol.* **122**(2), 356–363 (2011). <https://doi.org/10.1016/j.clinph.2010.06.021>
30. Jordan, C., Brennen, S.: *Instantaneous self-assessment of workload technique (ISA)*. Defence Research Agency, Portsmouth (1992)
31. Jou, S., Schultz, T., Walliczek, M., Kraft, F., Waibel, A.: Towards continuous speech recognition using surface electromyography. In: *Ninth International Conference on Spoken Language Processing* (2006)
32. Kamzanova, A.T., Kustubayeva, A.M., Matthews, G.: Use of EEG workload indices for diagnostic monitoring of vigilance decrement. *Hum. Factors* **56**(6), 1136–1149 (2014). <https://doi.org/10.1177/0018720814526617>
33. Kilner, J.M., Baker, S.N., Salenius, S., Jousmäki, V., Hari, R., Lemon, R.N.: Task-dependent modulation of 15–30 Hz coherence between rectified EMGs from human hand and forearm muscles. *J. Physiol.* **516**(2), 559–570 (1999). <https://doi.org/10.1111/j.1469-7793.1999.0559v.x>
34. Kristeva-Feige, R., Fritsch, C., Timmer, J., Lücking, C.H.: Effects of attention and precision of exerted force on beta range EEG-EMG synchronization during a maintained motor contraction task. *Clin. Neurophysiol.* **113**(1), 124–131 (2002). [https://doi.org/10.1016/S1388-2457\(01\)00722-2](https://doi.org/10.1016/S1388-2457(01)00722-2)
35. Lindström, B.R., Mattsson-Mårn, I.B., Golkar, A., Olsson, A.: In your face: risk of punishment enhances cognitive control and error-related activity in the corrugator supercilii muscle. *PLOS One* **8**(6), e65692 (2013). <https://doi.org/10.1371/journal.pone.0065692>
36. Longo, L.: Subjective usability, mental workload assessments and their impact on objective human performance. In: Bernhaupt, R., Dalvi, G., Joshi, A., Balkrishan, D.K., O'Neill, J., Winckler, M. (eds.) *INTERACT 2017*. LNCS, vol. 10514, pp. 202–223. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-67684-5\\_13](https://doi.org/10.1007/978-3-319-67684-5_13)
37. Marshall, S.P.: The index of cognitive activity: measuring cognitive workload. In: *Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants*, pp. 7–9 (2002). <https://doi.org/10.1109/HFPP.2002.1042860>
38. Matthews, G., Reinerman-Jones, L.E., Barber, D.J., Abich, J.: The psychometrics of mental workload. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **57**(1), 125–143 (2015). <https://doi.org/10.1177/0018720814539505>
39. Matthews, G., Warm, J.S., Reinerman-Jones, L.E., Langheim, L.K., Washburn, D.A., Tripp, L.: Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *J. Exp. Psychol.: Appl.* **16**(2), 187–203 (2010). <https://doi.org/10.1037/a0019572>
40. McCracken, J., Aldrich, T.B.T.: *Analyses of Selected LHX Mission Functions: Implications for Operator Workload and System Automation Goals*. Technical report, Anacapa Sciences Inc., Fort Rucker, AL (1984)
41. Meltzner, G.S., Heaton, J.T., Deng, Y., De Luca, G., Roy, S.H., Kline, J.C.: Development of sEMG sensors and algorithms for silent speech recognition. *J. Neural Eng.* **15**(4), 046031 (2018). <https://iopscience.iop.org/article/10.1088/1741-2552/aac965/meta>

42. Meltzner, G.S., Heaton, J.T., Deng, Y., De Luca, G., Roy, S.H., Kline, J.C.: Silent speech recognition as an alternative communication device for persons with laryngectomy. *IEEE/ACM Trans. Audio Speech Lang. Process.* **25**(12), 2386–2398 (2017). <https://doi.org/10.1109/TASLP.2017.2740000>
43. Mijović, P., Milovanović, M., Ković, V., Gligorijević, I., Mijović, B., Mačuzić, I.: Neuroergonomics method for measuring the influence of mental workload modulation on cognitive state of manual assembly worker. In: Longo, L., Leva, M.C. (eds.) *H-WORKLOAD 2017*. CCIS, vol. 726, pp. 213–224. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_14](https://doi.org/10.1007/978-3-319-61061-0_14)
44. Mulder, L., De Waard, D., Brookhuis, K.A.: Estimating mental effort using heart rate and heart rate variability. In: *Handbook of Human Factors and Ergonomics Methods*, pp. 1–20 (2004)
45. Neto, O.P., Christou, E.A.: Rectification of the EMG signal impairs the identification of oscillatory input to the muscle. *J. Neurophysiol.* **103**(2), 1093–1103 (2010). <https://doi.org/10.1152/jn.00792.2009>
46. O'Donnell, R., Eggemeier, F.: Workload assessment methodology. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) *Handbook of Perception and Human Performance, Cognitive Processes and Performance*, vol. 2. Wiley, Hoboken (1986)
47. Pickup, L., Wilson, J.R., Sharpies, S., Norris, B., Clarke, T., Young, M.S.: Fundamental examination of mental workload in the rail industry. *Theor. Issues Ergon. Sci.* **6**(6), 463–482 (2005). <https://doi.org/10.1080/14639220500078021>
48. Reid, G.B., Nygren, T.E.: *The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload*, North-Holland, vol. 52 (1988). [https://doi.org/10.1016/S0166-4115\(08\)62387-0](https://doi.org/10.1016/S0166-4115(08)62387-0)
49. Santiago-Espada, Y., Myer, R., Latorella, K., Comstock, J.: *The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide*. Technical report, NASA, Chicago (2011)
50. Smith, A.P., Smith, H.N.: Workload, fatigue and performance in the rail industry. In: Longo, L., Leva, M.C. (eds.) *H-WORKLOAD 2017*. CCIS, vol. 726, pp. 251–263. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_17](https://doi.org/10.1007/978-3-319-61061-0_17)
51. Stepp, C.E., Hillman, R.E., Heaton, J.T.: Use of neck strap muscle intermuscular coherence as an indicator of vocal hyperfunction. *IEEE Trans. Neural Syst. Rehabil. Eng.* **18**(3), 329–335 (2010). <https://doi.org/10.1109/TNSRE.2009.2039605>
52. Stepp, C.E., Hillman, R.E., Heaton, J.T.: Modulation of neck intermuscular beta coherence during voice and speech production. *J. Speech Lang. Hear. Res.* **54**(3), 836–844 (2011). [https://doi.org/10.1044/1092-4388\(2010\)10-0139](https://doi.org/10.1044/1092-4388(2010)10-0139)
53. Stroobant, N., Vingerhoets, G.: Transcranial Doppler ultrasonography monitoring of cerebral hemodynamics during performance of cognitive tasks: a review (2000). <https://doi.org/10.1023/A:1026412811036>
54. Tassinary, L.G., Cacioppo, J.T.: Unobservable facial actions and emotion. *Psychol. Sci.* **3**(1), 28–33 (1992). <https://doi.org/10.1111/j.1467-9280.1992.tb00252.x>
55. Tong, S., Helman, S., Balfe, N., Fowler, C., Delmonte, E., Hutchins, R.: Workload differences between on-road and off-road manoeuvres for motorcyclists. In: Longo, L., Leva, M.C. (eds.) *H-WORKLOAD 2017*. CCIS, vol. 726, pp. 239–250. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61061-0\\_16](https://doi.org/10.1007/978-3-319-61061-0_16)
56. Tsang, P.S., Velazquez, V.L.: Diagnosticity and multidimensional subjective workload ratings. *Ergonomics* **39**(3), 358–381 (1996). <https://doi.org/10.1080/00140139608964470>
57. Van Boxtel, A., Goudswaard, P., Van der Molen, G., Van Den Bosch, W.: Changes in electromyogram power spectra of facial and jaw-elevator muscles during fatigue. *J. Appl. Physiol.* **54**(1), 51–58 (1983). <https://doi.org/10.1152/jappl.1983.54.1.51>

58. Van Boxtel, A.: Facial EMG as a tool for inferring affective states. In: *Proceedings of Measuring Behavior*, pp. 104–108. Noldus Information Technology Wageningen (2010)
59. Vanneste, S.: Timing in aging: the role of attention. *Exp. Aging Res.* **25**(1), 49–67 (1999). <https://doi.org/10.1080/036107399244138>
60. Voelcker-Rehage, C., Stronge, A.J., Alberts, J.L.: Age-related differences in working memory and force control under dual-task conditions. *Aging, Neuropsychol. Cogn.* **13**(3–4), 366–384 (2006). <https://doi.org/10.1080/138255890969339>
61. Warm, J.S., Tripp, L.D., Matthews, G., Helton, W.S.: Cerebral hemodynamic indices of operator fatigue in vigilance. In: *Handbook of Operator Fatigue*, pp. 197–207 (2012)
62. Welch, P.: The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **15**(2), 70–73 (1967). <https://doi.org/10.1109/TAU.1967.1161901>
63. Wickens, C.D.: Multiple resources and mental workload. *Hum. Factors* **50**(3), 449–455 (2008). <https://doi.org/10.1518/001872008X288394>
64. Yeh, Y.Y., Wickens, C.D.: Dissociation of performance and subjective measures of workload. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **30**(1), 111–120 (1988). <https://doi.org/10.1177/001872088803000110>
65. Zhang, J.Y., Liu, S.L., Feng, Q.M., Gao, J.Q., Zhang, Q.: Correlative evaluation of mental and physical workload of laparoscopic surgeons based on surface electromyography and eye-tracking signals. *Sci. Rep.* **7**(1), 11095 (2017). <https://doi.org/10.1038/s41598-017-11584-4>