International Journal of Psychophysiology 83 (2012) 96–101

Contents lists available at SciVerse ScienceDirect

International Journal of Psychophysiology

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

Respiratory sinus arrhythmia as a measure of cognitive workload

Eric R. Muth *, Jason D. Moss, Patrick J. Rosopa, James N. Salley, Alexander D. Walker

Department of Psychology, Clemson University, Clemson, SC, United States

ARTICLE INFO

Article history:
Received 10 August 2011
Received in revised form 26 October 2011
Accepted 30 October 2011
Available online 10 November 2011

Keywords:
Respiratory sinus arrhythmia
Workload
Heart rate variability
NASA-TLX

ABSTRACT

The current standard for measuring cognitive workload is the NASA Task-load Index (TLX) questionnaire. Although this measure has a high degree of reliability, diagnosticity, and sensitivity, a reliable physiological measure of cognitive workload could provide a non-invasive, objective measure of workload that could be tracked in real or near real-time without interrupting the task. This study investigated changes in respiratory sinus arrhythmia (RSA) during seven different sub-sections of a proposed selection test for Navy aviation and compared them to changes reported on the NASA-TLX. 201 healthy participants performed the seven tasks of the Navy's Performance Based Measure. RSA was measured during each task and the NASA-TLX was administered after each task. Multi-level modeling revealed that RSA significantly predicted NASA-TLX scores. A moderate within-subject correlation was also found between RSA and NASA TLX scores. The findings support the potential development of RSA as a real-time measure of cognitive workload.

1. Introduction

Measuring cognitive workload has broad appeal because of its apparent link to performance. Cognitive workload (hereafter referred to as "workload") refers to the difference between the cognitive demands of a task and the cognitive resources available to the operator (Gopher and Donchin, 1986). Both tasks that have too high workload (Gopher and Donchin, 1986) and too low workload (Pattyn et al., 2008; Warm et al., 2008; Young and Stanton, 2002) have been shown to cause performance decrements. Measuring workload in real or near real-time could afford adjustment of task demands in a way to potentially optimize performance (Byrne and Parasuraman, 1996).

The NASA-TLX, developed by Hart and Staveland (1988), has become the standard for measuring subjective workload. It measures workload on six subscales: mental demand, physical demand, temporal demand, effort, performance, and frustration. Each subscale is individually rated on an analog scale, and through a series of weighted comparisons, each subscale is weighed relative to the other subscales, and an overall workload score is calculated. The NASA-TLX has been used to assess workload in a variety of tasks including those in very cognitively demanding jobs such as aircraft pilots (Karavidas et al., 2010; Lehrer et al., 2010; Sohn and Jo, 2003) and air traffic controllers (Brookings et al., 1996). Within the realm of psychophysiology, it has been used as a benchmark for assessing psychophysiological measures of workload (Fournier et al., 1999; Miyake et al., 2009; Prinzel et al., 2000). It has also been used in conjunction with other measures of workload to improve their overall diagnosticity (Lehrer et al., 2010).

Subjective measures like the NASA-TLX have some limitations. Subjective measures require a conscious response from the user that is subject to personal perceptions (O'Brien and Charlton, 1996). In addition, subjective measures require an interruption of the task, and therefore, cannot provide feedback in real-time (Kramer, 1991; Wilson and Eggemeier, 1991; Yeh and Wickens, 1988). This can severely limit the real-world application of these measures, particularly for real-time assessment. An objective, psychophysiological measure could be preferable for this purpose.

A variety of physiological measures have been used to quantify workload, including the electroencephalogram (Brookings et al., 1996; Fournier et al., 1999; Prinzel et al., 2000; Ullsperger et al., 1988; Wilson, 2002), eye movement activity and pupilometry (Brookings et al., 1996; Dahlstrom and Nahlinder, 2009; Fournier et al., 1999; Hankins and Wilson, 1998; Wilson, 2002), skin conductance (Wilson, 2002), metabolic measures (Fairclough and Houston, 2004), and heart rate variability (HRV) (Brookings et al., 1996; Dahlstrom and Nahlinder, 2009; Fournier et al., 1999; Wilson, 2002). Because HRV is regulated by autonomic nervous system activity (ANS) and workload has been correlated with ANS activity (Comens et al., 1987; Hart and Hauser, 1987; Lindholm and Cheatham, 1983; Nicholson, et al., 1970; Wilson, 1993, 2002), HRV is one of the most often studied physiological indices of workload (Backs, 1995). Further, the acquisition of HRV data is now possible in most, if not all, environments.

HRV is influenced by both sympathetic nervous system (SNS) activity and parasympathetic nervous system (PNS) activity. Backs (1995) points out that examining heart rate alone is not enough to index workload, as heart rate is influenced by factors other than

* Corresponding author at: 418 Brackett Hall, Clemson, SC 29634-1355, United States. Tel.: +1 864 656 6741.

E-mail addresses: muth@clemson.edu (E.R. Muth), jdmoss25@gmail.com (J.D. Moss), prosopa@clemson.edu (P.J. Rosopa), jnsalle@clemson.edu (J.N. Salley), alexanderwalker@gmail.com (A.D. Walker).

0167-8760/$ – see front matter
doi:10.1016/j.ijpsycho.2011.10.011

workload and does not differentiate between SNS and PNS effects. PNS and SNS influences on HRV can be parsed by examining more specific parameters. The HRV frequency corresponding to changes in respiration, known as respiratory sinus arrhythmia (RSA) has been validated as an index of PNS activity when respiratory influences are accounted for (Grossman et al., 1991; Grossman et al., 1990). While other frequencies of HRV have been examined and shown to be sensitive to changes in cognitive workload, they require more complicated instrumentation and knowledge to implement as well as more sophisticated data reduction and analysis techniques. Measurements of RSA alone require minimal instrumentation, and can be integrated with current technology.

Respiration controlled RSA has been shown to be sensitive to changes in workload (Althaus et al., 1998; Backs et al., 1999). Fishel et al. (2006) monitored changes in RSA during a dual task paradigm with high task load and low task load conditions, and found that for tasks lasting longer than 2 min, RSA was significantly lower in the high task load condition than the low task load condition. In a study investigating patterns in RSA while performing tasks of varying task load during sleep deprivation, Walker et al. (2009) found that RSA was significantly lower in tasks with a high cognitive demand than in tasks with a low cognitive demand. In another study comparing the sensitivity of the different bands of HRV, Althaus et al. (1998) found that, when controlling for respiration, only the RSA frequency of HRV was sensitive to changes in task load. Other research has shown that as the complexity of a task increased, HRV in the RSA frequency decreased, showing a withdrawal of PNS activity (Althaus et al., 1998; Backs et al., 1999; Veltman and Gaillard, 1996).

The purpose of this study was to compare measures of workload gathered from the NASA-TLX to measures of workload gathered from RSA. Given the previous work, it was hypothesized that RSA would be negatively correlated with the NASA-TLX measure of workload. If RSA is shown to be comparable to NASA-TLX as a measure of workload, it has the potential to provide a continuous, non-invasive measure of workload that requires no conscious input from the examinee that can be used reliably in real-world situations (Grossman et al., 1990; Porges and Byrne, 1992).

2. Method

2.1. Participants

Although 213 participants completed the study, HRV data from 12 participants were discarded because of uncorrectable errors in the recordings. Therefore, data from 201 (108 female) participants ranging in age from 18 to 34 years ($M = 20$, $SD = 2.47$) were analyzed. Participants were recruited using an advertisement posted throughout Clemson University and through the psychology department's Human Participation in Research website. Individuals who did not self-report normal or corrected vision or normal hearing, did not have English as a first language, or had previous flight experience were ineligible to participate in the study. Participants were asked to refrain from caffeine, alcohol, and tobacco use, and from any vigorous physical activity on the day of their scheduled participation. All participants gave signed informed consent to participate in the study, which was approved by Clemson University's Institutional Review Board. Participants were compensated with $20 for their time. Participants who were recruited through the website received course credit in addition to the $20.

2.2. Workload task manipulation

The workload task manipulation consisted of seven subsections of a performance based selection test being developed by the US Navy. All seven subsections were performed on a computer. The software was installed on four desktop computer stations; each equipped with a Thrustmaster HOTAS™ (Hands on Throttle and Stick) Cougar flight control system. The tasks were displayed on 19 in. LCDs. Partitions were used to separate each station and provide privacy. The tasks were composed of seven subsections consisting of four single tasks and three multi-task subsections. In order of performance, the single task subsections were: 1) a cardinal direction task; 2) a dichotic listening task; 3) a one-dimensional tracking task; and 4) a two-dimensional tracking task. The multi-task subsections were: 5) one-dimensional tracking + two dimensional tracking; 6) one-dimensional tracking + two dimensional tracking + dichotic listening; and 7) one-dimensional tracking + two-dimensional tracking + an emergency scenario task. With the exception of subsection 7, each subsection was composed of three sessions: instructional, practice, and the full test session. Subsection 7 did not include a practice session.

In the cardinal direction task subsection, participants had to view an overhead map depicting an aircraft's heading and then locate a target (i.e., north, east, south, or west parking lot) from the view of the aircraft's perspective. This task consisted of 48 target location problems with no time constraints. During the dichotic listening task subsection, participants were instructed to monitor the left or right ear for 50 s and had to respond as quickly and accurately in one of two ways: 1) pull the trigger of the joystick when an even number was presented; or 2) press a button on the throttle when an odd number was presented.

The one-dimensional tracking task (vertical tracking task) subsection consisted of using the throttle to control an airplane icon in order to track a moving target of a cross hair along the vertical axis. This task lasted 60 s. The two-dimensional tracking task (airplane tracking task) subsection measured the ability to track a moving target in two dimensions. Participants used the flight stick to control an airplane icon in order to track a crosshair moving along the horizontal and vertical axes. This task lasted for 60 s.

The one-dimensional + two-dimensional multi-task subsection combined the vertical tracking and airplane tracking tasks. Participants had to use the throttle to perform the vertical tracking component and the flight stick to perform the airplane tracking component. The duration of this subsection was 120 s. This subsection was further combined with the dichotic listening task to compose the sixth subsection, which lasted 180 s.

The last subsection combined the one-dimensional + two-dimensional tracking task subsection with an emergency scenario task. While performing the tracking tasks, participants had to monitor several display indicators to detect three potential emergency situations; i.e., engine fire, engine malfunction, and propeller malfunction. Each emergency situation called for a unique response in order to properly resolve the emergency situation. Participants responded by manipulating the appropriate controls on the throttle in the correct sequence for each emergency situation. The appropriate manipulations and sequences for each emergency situation were presented to the participants during the instructional phase of the subsection. Further, if participants failed to respond, responded too slowly, or responded incorrectly, the background color changed to red indicating a system operating under duress. This subsection lasted 180 s.

2.3. NASA TLX workload measure

An automated version of the NASA-TLX (Hart and Staveland, 1988) was installed on each computer and used to collect self-report workload data. The NASA-TLX provides a global measure of task workload derived from: 1) participants' ratings on mental demand, physical demand, temporal demand, performance, effort, and frustration; and 2) participants' personal weights for each of the above dimensions of workload. Participants are instructed to rate each sub dimension on a visual analog scale. Participants are then presented with 15 paired comparisons of each subsection and asked to choose which factor had a greater impact on their performance. The results of these comparisons apply a specific weight to each factor from 0 to 5. Each subsection is rated on a 100 point