The Psychometrics of Mental Workload: Multiple Measures Are Sensitive but Divergent

Gerald Matthews, Lauren E. Reinerman-Jones, Daniel J. Barber, and **Julian Abich IV,** University of Central Florida, Orlando

Objective: A study was run to test the sensitivity of multiple workload indices to the differing cognitive demands of four military monitoring task scenarios and to investigate relationships between indices.

Background: Various psychophysiological indices of mental workload exhibit sensitivity to task factors. However, the psychometric properties of multiple indices, including the extent to which they intercorrelate, have not been adequately investigated.

Method: One hundred fifty participants performed in four task scenarios based on a simulation of unmanned ground vehicle operation. Scenarios required threat detection and/or change detection. Both single- and dual-task scenarios were used. Workload metrics for each scenario were derived from the electroencephalogram (EEG), electrocardiogram, transcranial Doppler sonography, functional near infrared, and eye tracking. Subjective workload was also assessed.

Results: Several metrics showed sensitivity to the differing demands of the four scenarios. Eye fixation duration and the Task Load Index metric derived from EEG were diagnostic of single-versus dual-task performance. Several other metrics differentiated the two single tasks but were less effective in differentiating single- from dual-task performance. Psychometric analyses confirmed the reliability of individual metrics but failed to identify any general workload factor. An analysis of difference scores between low- and high-workload conditions suggested an effort factor defined by heart rate variability and frontal cortex oxygenation.

Conclusions: General workload is not well defined psychometrically, although various individual metrics may satisfy conventional criteria for workload assessment.

Application: Practitioners should exercise caution in using multiple metrics that may not correspond well, especially at the level of the individual operator.

Keywords: mental workload, dual-task performance, psychometrics, psychophysiology, neuroergonomics

Address correspondence to Gerald Matthews, University of Central Florida, 3100 Technology Pkwy, Orlando, FL 32826, USA; e-mail: gmatthews@ist.ucf.edu.

HUMAN FACTORS
Vol. 57, No. 1, February 2015, pp. 125–143
DOI: 10.1177/0018720814539505

INTRODUCTION

It is standard practice to assess mental workload during system design and evaluation in order to avoid operator overload in a variety of industrial, transportation, military, and medical contexts (Vidulich & Tsang, 2012). Recent interest in augmented cognition has highlighted the importance of continuous workload monitoring for detecting transient overload and for regulating function allocation in automated systems (Parasuraman, Sheridan, & Wickens, 2008). Workload assessment in all practical contexts requires reliable and valid metrics. From a theoretical standpoint, workload is characterized in relation to the demands imposed by tasks on the operator's limited information-processing resources (Wickens, 2008).

Workload is not directly observable. Instead, it is an abstract property of human-machine interaction that is inferred through multiple methods, including subjective report, psychophysiological measurement, and performance itself (Hancock & Chignell, 1988; O'Donnell & Eggemeier, 1986; Vidulich & Tsang, 2012). To date, the most widely used and extensively validated workload instruments are subjective scales, such as the NASA Task Load Index (NASA-TLX; Hart, 2006; Hart & Staveland, 1988; Stanton, Salmon, Walker, Baber, & Jenkins, 2005). The pragmatic utility of self-report instruments is well established (Vidulich & Tsang, 2012), but they have several limitations. First, especially in high-stakes settings, responses may be subject to biases, such as social desirability. Second, self-reports are not well suited to continuous monitoring of workload. Third, there may be important facets of workload that are inaccessible to consciousness. In principle, these limits may be overcome through the use of objective, psychophysiological indices obtained through recordings of the electroencephalogram (EEG), electrocardiogram (ECG), eye tracking, and other response systems (see Cain, 2007; Vidulich & Tsang, 2012; Wilson & Eggemeier, 1991). Such recordings can be secured in real-world