International Journal of Industrial Ergonomics xxx (2016) 1–16

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

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

Workload profiles: A continuous measure of mental workload

Christina F. Rusnock*, Brett J. Borghetti

Air Force Institute of Technology, 2950 Hobson Way, Building 640, WPAFB, OH 45433-7765, United States

ARTICLE INFO

Article history:
Received 13 July 2015
Received in revised form
27 June 2016
Accepted 11 September 2016
Available online xxx

ABSTRACT

The required frequency and disruptive method in which existing subjective measures of mental workload are collected make them infeasible for many types of task allocation decisions. In this paper, we present a method for continually estimating workload without interrupting the operator. When expressed as a time-series, this continual workload assessment becomes a workload profile which can serve purposes before, during, and after task execution. We identify five thrusts areas for using workload profiles which cannot be accomplished using existing workload measures. These thrust areas include characterization of workload over time; identifying the impacts of task management strategy on mission accomplishment; evaluating potential effects of systems design options—including automation—on task performance; informing manpower allocation decisions; and enhancing physiological computing and neuroergonomic research.

Published by Elsevier B.V.

1. Introduction

Effective system performance relies upon both the reliable performance of all subsystems and an effective interaction between all subsystems. When a system includes a human operator, that operator should be considered a subsystem. Because the human operator subsystem is likely to have high variability and low reliability, Human Factors professionals have long recognized that the human operator is one of the most critical subsystems, and must be accounted for when evaluating system performance.

One method commonly used to understand and predict the performance of the human operator is to measure, estimate, or analyze operator workload. Workload is the amount of effort experienced by the operator when performing a task, and is thus affected by both operator context and external factors. Operator context that influences workload include individual capabilities (both physical and mental), training, experience, fatigue, stress, and personality. External factors influencing workload include task quantity, task difficulty, and time available, as well as environmental factors such as temperature and lighting. The workload experienced due to task and environmental factors is largely determined by the system design decisions regarding human-machine interaction. Recognizing that the workload experienced by the operator impacts the operator's performance–which in turn impacts system performance–is key to effective system design.

In order to study or evaluate system performance, we must have a means to measure and predict operator workload. There is a large body of work focused on the measurement of mental workload through subjective-empirical measures (e.g. self-report questionnaires, such as Boles and Adair, 2001; Hart and Staveland, 1988). However, subjective-empirical measures are problematic because they typically result in a single, cumulative value measured at the completion of the task. End-of-task workload measurements fail to capture workload variability and timing throughout the task. In an effort to overcome these issues, more recent efforts have thus turned to objective-empirical measures, specifically neurological and physiological measurements (Parasuraman and Wilson, 2008). While the work using these objective-empirical tools is still nascent, this exploration is promising. However, these tools are not by themselves a direct measure of mental workload, and thus require some other measurement to effectively interpret and corroborate their outputs.

This paper focuses on using analytical measures as an alternative means for measuring and predicting mental workload. Analytical tools can overcome a number of the weaknesses of subjective-empirical tools, because they can reveal the variability or steadiness of workload over the course of a task, enable identification of workload drivers, and enable workload-based interventions. Furthermore, analytical tools can be used in combination with empirical tools to produce a more robust, holistic view of workload (Rusnock et al., 2015). One such tool is the

* Corresponding author.
E-mail address: christina.rusnock@afit.edu (C.F. Rusnock).

http://dx.doi.org/10.1016/j.ergon.2016.09.003
0169-8141/Published by Elsevier B.V.

Improved Performance Research Integration Tool (IMPRINT) (Alion, 2015), a discrete event simulation tool specifically designed for cognitive workload modeling. IMPRINT enables workload prediction using task networks and the Visual Auditory Cognitive Psychomotor (VACP) method (Bierbaum et al., 1989; McCracken and Aldrich, 1984), a quantitative implementation of multiple resource theory. Workload modeling using IMPRINT results in the creation of a continuous analytic workload profile (CAWP). Unlike subjective workload measures, this workload profile provides a continuous estimate of workload with a unique workload value for each point in time.

2. Purpose

The purpose of this study is to demonstrate the utility of an analytically-generated continuous workload profile. To do this, we explore five different potential thrust areas for the workload profile that could not be effectively achieved using empirical measures alone. For concreteness, these five thrust areas are presented in the context of remotely piloted aircraft (RPA) surveillance tasks, which cover phases before, during, and after task execution. The five thrust areas are:

1) characterize workload for a task;
2) identify how overloaded operators' task management strategies impact task/mission accomplishment;
3) analyze potential effects of system design options and automation;
4) evaluate manpower allocations; and
5) enhance physiological computing and neuroergonomic research.

3. Background

We explore the background of research in workload estimation, modeling, and validation using three lenses. The first lens examines methods which are used to estimate workload directly through self-reporting or indirectly via the relationship between workload and task performance, or the relationship between workload and physiological or neurological response. With the second lens we explore workload modeling methods which produce estimates of an individual's workload, focusing heavily on a multiple resource paradigm modeling tool. Finally, we conclude with methods used to validate workload models.

3.1. Estimating workload directly

Over the past 50 years, researchers have used a variety of subjective measurement tools to measure cognitive workload. These subjective-empirical measures involve asking the subject to rate their mental effort using various scales. Due to the nature of these tools, the measurements are typically conducted just after the conclusion of the task. Because of the timing of the measurement, these tools typically provide a single, cumulative estimate of workload. This cumulative estimate likely represents an average of the operator's workload, although it may be subject to memory biases that give disproportionate weight to the most recent events, or excessively high/low periods of workload. The cumulative nature of these instruments also makes it difficult to identify variability in the workload and the corresponding timing for workload variations.

NASA-TLX is one of the most widely used subjective measures (Hart and Staveland, 1988; Hart, 2006), and features six dimensions: mental, physical, temporal, frustration, performance, and effort. Subjects rate their workload on each of these dimensions separately on a 100 point scale (in increments of 5), as well as evaluate pair-wise comparisons for each of these dimensions. The pair-wise comparisons are used to create a weighted aggregate score. The dimensionality of the tool is helpful in narrowing down the potential source(s) for excess workload.

The Multiple Resources Questionnaire (MRQ) is another well-established subjective measure that features the evaluation of workload across seventeen dimensions (Boles and Adair, 2001). The MRQ is based on an expansion and reinterpretation of Wickens' Multiple Resource Theory (Wickens, 1984), and thus, unlike NASA-TLX, has dimensions that directly relate to various mental resources (referred to as "processes"). Users are instructed to rate the amount of usage for each resource as an average over the whole time the task is performed, from the following Likert scale: no usage, light usage, moderate usage, heavy usage, and extreme usage. These Likert scale options are given values from 0 to 100 in 25 point increments. The resource processes are: auditory emotional, auditory linguistic, facial figural, facial motive, manual, short term memory, spatial attentive, spatial categorical, spatial concentrative, spatial emergent, spatial positional, spatial quantitative, tactile figural, visual lexical, visual phonetic, visual temporal, and vocal. While highly detailed, the instrument can be cumbersome to use and interpret.

Additional commonly used subjective workload measures include the Subjective Workload Assessment Technique (SWAT) (Luximon and Goonetilleke, 2001; Reid and Nygren, 1988), the Cooper-Harper Rating Scale (Cooper and Harper, 1969), the Bedford Scale (Roscoe and Ellis, 1990), Overall Workload (Jung and Jung, 2001), Workload Profile (Rubio et al., 2004; Tsang and Velazquez, 1996), and the Integrated Workload Scale (Pickup et al., 2005).

In addition to subjective-empirical measures, workload researchers are increasingly turning to objective-empirical measures by directly measuring an operator's behavioral performance or physiological state. Objective-empirical measures have a number of advantages over subjective-empirical measures, including the ability to measure them in real-time, elimination of operator response bias, and relatively low task-interference.

Behavioral performance measures typically use either primary or secondary task performance as a proxy for workload. With task performance, the underlying assumption is that performance (error rate, response time, response accuracy) decreases with increasing workload. However, the relationship between workload and performance is likely closer to the inverted U-shape described by the Hebb-Yerkes-Dodson Law, rather than linearly decreasing (Teigen, 1994). The law suggests that both very high workload (overload) and very low workload (underload) are associated with lower performance, and that higher performance is achieved in a medium workload condition. Thus, depending on where the individual is on the curve, increasing workload could increase or decrease performance. Similarly, an increase in performance could have been caused by an increase or decrease in workload, making it hard to use performance as a proxy for estimating workload. Using performance as a proxy for workload can also be problematic as performance effects are not immediate, and may experience a sizeable time-lag after workload has changed.

Using physiological or neurological responses to estimate workload provides an opportunity for timely, sensitive workload information that is specifically tailored to the individual. One of the most common physiological measures of cognitive workload is the measurement of brain activity. This can be accomplished through the use of electroencephalography (EEG), functional magnetic resonance imaging (fRMI), or transcranial Doppler (TCD) sonography. Researchers have successful correlated EEG, fMRI, and TCD results with different resource channels, thus enabling the tracking

Please cite this article in press as: Rusnock, C.F., Borghetti, B.J., Workload profiles: A continuous measure of mental workload, International Journal of Industrial Ergonomics (2016), http://dx.doi.org/10.1016/j.ergon.2016.09.003