A MULTI-SENSORY PROTOCOL FOR EVALUATING WORKLOAD ASSOCIATED WITH NEW ARMY TECHNOLOGIES

J. Christopher Brill*, Mustapha Mouloua, Peter A. Hancock, and Richard D. Gilson
University of Central Florida
Orlando, FL 32816

Robert S. Kennedy and Marty G. Smith
RSK Assessments, Inc.
Orlando, FL 32803

Although a variety of workload measures exist, none are presently capable of assessing the capacities of cognitive resource pools across multiple sensory modalities and across various tasks (Meshkati, Hancock, Rahimi, & Dawes, 1995). Assessment of workload is critical to Army operations given their increasing reliance on advanced technologies. The goal of the present effort is to develop both a device and a protocol to assess simultaneously the workload associated with complex tasks involving more than one sensory system. Such a methodology may be used to determine whether the operator demands of advanced technologies exceed human capabilities. This type of macro-workload assessment is critical for achieving optimal human performance levels, and consequently, mission success in a variety of military systems. More specifically, effective workload assessment can be used to improve the mental agility and lethality of Army forces by removing elements that overtax soldier cognitive-perceptual capacities. Alleviation of such elements can lead to better decision-making and improved information management. Ultimately, this will have an impact on soldier survivability, effectiveness, and adaptability.

Five design criteria were established to guide the development of a new workload assessment system. First, portability is essential to maximize the utility of the protocol across research scenarios, regardless of whether it is used in the field or in the laboratory. Second, the protocol should be independent of language requirements and cultural bias. Third, use of the methodology should not require abilities such as mathematical computation or reading. Fourth, the protocol should be easy to learn, and there should be an absence of practice effects. Fifth, the device should be relatively inexpensive to build and maintain to increase its accessibility. It is believed that the workload assessment device and protocol envisioned below meets all five of these criteria.

The workload assessment device consists of a wearable display as an input to the least used sensory modality: touch. The system is comprised of three vibrotactile actuators (called tactors), a relay-based tactor controller board, a three-button response box, and a PC using proprietary software developed by our research team. The system is used for presenting a multi-sensory counting task for assessing workload while operators perform other complex tasks such as using telecommunications systems, monitoring radar systems, or even operating unmanned aerial vehicles (UAVs).

The protocol for using the workload assessment device is generally the same regardless of which sensory modality (i.e., visual, auditory, or tactile) is targeted, but the following description depicts the tactile component of the protocol. Participants are randomly presented with a series of vibratory taps situated at three loci (left, center, or right) on the abdomen. They count the frequencies of taps (each four count) for the assigned position (e.g., left tactor), and respond by pressing the response button corresponding to the position of the stimuli (e.g., left button). Since participants are asked to monitor a single locus of stimulation, all other signals act as random distractors. The number of counting errors serves as a measure of workload, wherein low workload tasks should be associated with few counting errors and high workload tasks should be associated with greater frequencies of counting errors. The difficulty of the counting task is scalable depending upon the number of stimulus sites participants are instructed to monitor (one site = low difficulty, two sites = moderate difficulty, etc.). The counting task for the other two sensory modalities is similar wherein participants are counting from among either three lights (left side of screen, center of screen, or right side of screen) or three auditory pitches (low, middle, and high). The counting task is an adaptation of the one used earlier by Kennedy (1971).

Aside from meeting the five aforementioned design criteria, there are other significant advantages of the multi-sensory workload assessment protocol

(M-SWAP) over other established workload measures. For example, the NASA Task Load Index (NASA-TLX) is a commonly used subjective self-report workload measure in which operators rate items on several task dimensions such as mental demands, temporal demands, effort, and frustration (Hart & Staveland, 1988). The NASA-TLX, however, is limited in that it cannot be administered while someone is actually performing a task, as can M-SWAP. In addition, the NASA-TLX treats cognitive resources as a unitary pool (e.g., Broadbent, 1958; Kahneman, 1973), which restricts its utility for situations in which one wishes to determine the differential demands of a task on various cognitive resources, especially if the research is rooted in theories postulating the existence of multiple resource pools of attention (e.g., Wickens, 1984). In fact, assessing resources across and within modalities is a distinct advantage of M-SWAP. With M-SWAP, workload can be assessed on-line (meaning while the operator performs the task) and in real-time, without the limitations of task-specificity associated with primary task measures (e.g., Linton's (1975) statistical workload assessment model (SWAM) or North and Riley's (1988) Workload Index (W/INDEX)). This real-time assessment is essential for the design of adaptive technologies in which elements of a complex task might be automated through feedback if excessive workload levels are detected (Hancock & Chignell, 1988). Finally, the secondary task paradigm used by M-SWAP can be used to determine the availability of resources across or within sensory modalities isolated from or vying for resources within a modality or resource pool. These data can also be used for system design to maximize the likelihood of signal detection while performing complex tasks. For example, one might use M-SWAP as a design decision aid to confirm of the availability of resources within a sensory modality that is theoretically untaxed by a task (e.g., given task "x," which sensory modality should an alarm system target?).

Testing to establish the validity of M-SWAP is continuing, but initial evaluation of this methodology in both a driving simulator and in an on-road vehicle suggests M-SWAP is sensitive to different levels of workload. For example, while operating an automobile in light traffic participants committed an average of one counting error (visual) every four minutes. When given the additional task of either operating a navigation system or talking on a cell phone, mean counting errors increased 2.5 and 4.5 per four-minute block, respectively. These findings and the results of further testing of M-SWAP will be presented in forthcoming research publications.

With the increasing sophistication and pervasive use of Army technologies, a universal protocol for assessing cognitive workload is needed to help prevent human capabilities from being exceeded. It is believed that M-SWAP can assist with this function so that troops can perform their best, aided by new technology, rather than be hindered by it.

ACKNOWLEDGEMENTS

This research was facilitated by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the ARO under grant DAAD19-01-1-0621. Professor P. A. Hancock was the Principle Investigator.

REFERENCES

Broadbent, D. (1958). *Perception and communication*. New York: Pergamon.

Hancock, P. A., & Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *IEEE Transactions of systems, man, and cybernetics, 18,* 647-658.

Hart, S., & Staveland, L. (1988). Development of the NASA-TLX: Results of empirical and theoretical research. In P. Hancock and N. Meshkati (Eds.), *Human mental workload*. Amsterdam: North-Holland.

Khaneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kennedy, R. S. (1971). A comparison of performance on visual and auditory monitoring tasks. *Human Factors*, *13(2),* 93-97.

Linton, P. (1975). *VFA-STOL crew loading analysis* (NADC-57209-40). Warminster, PA: U.S. Naval Air Development Center.

Meshkati, N., Hancock, P. A., Rahimi, M., & Dawes, S. M. (1995). Techniques in mental workload assessment. In J. R. Wilson and E. N. Corlett, (Eds.) *Evaluation of human work: A practical ergonomics methodology* (pp. 749-782). Taylor & Francis, Philadelphia, PA.

North, R., & Riley, V. (1988). W/INDEX: A predictive model of operator workload. In G. MacMillian (Ed.), *Human performance models*. Orlando, FL: NATO AGARD Symposium.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63-101). Hillsdale, NJ: Erlbaum.