



# The Effects of Workload Transitions in a Multitasking Environment

Margaret A. Bowers<sup>1,2</sup>, James C. Christensen<sup>2</sup>, and F. Thomas Eggemeier<sup>1</sup>

<sup>1</sup>University of Dayton

<sup>2</sup>US Air Force Research Laboratory

## ABSTRACT

Interest in workload transitions is centered on the hypothesis that transitions from one level of task difficulty to another may negatively impact performance on those tasks and result in potentially critical impacts on professionals such as air traffic controllers and emergency medical staff. The current study sought to determine the effect of workload transitions on participants' neurophysiological signals and performance. Participants completed trials in the Air Force Multi-Attribute Task Battery (AF-MATB; Miller, 2010), while Electroencephalography (EEG), Electrocardiography (ECG) and Electrooculogram (EOG) signals were recorded. All participants completed AF-MATB trials that delivered consistently low or high task difficulty, and trials that transitioned from an easy level to a difficult level and vice versa. Additionally, participants completed the NASA Task Load Index to assess subjective workload, and the shortened Dundee Stress State Questionnaire to measure subjective task-related stress during their testing sessions. Analyses of the performance data provide limited support for a negative impact of transitions from hard-to-easy, and the analysis of both the NASA Task Load Index and of the shortened Dundee State Questionnaire did not reveal any significant differences related to workload transitions. Analysis of the EEG data revealed that temporal gamma oscillations rapidly changed following a transition and settled after easy-to-hard changes in task difficulty, but settled more slowly after the task transitioned from hard-to-easy. Frontal theta oscillations, in contrast, exhibited consistently rapid settling which may indicate rapid changes in working memory utilization and conflict resolution (Gevins, et al., 1997). These EEG results suggest potential for further research.

## INTRODUCTION

A workload transition is a shift in the difficulty of a cognitive task. Historically, these transitions have been of interest because of the possibility that a performer's effectiveness and accuracy may be affected by these changes in workload. Morgan and Hancock (2011) use driving as an example and point out that, "...the history of previously experienced events may be as influential on driver response and levels of workload as are current levels of demand. The term most applicable for the ongoing influence of such prior historical influences is *hysteresis*" (p76). Any effects of prior task workload on post transition workload or performance represent so-called hysteresis effects.

For many professions, a transient change in effectiveness is not cause for much concern; however, in some professions, there is little room for error. For instance, Farrell (1999) noted that, "an investigation of air traffic control (ATC) operational errors showed that a high proportion of near misses occurred after a period of sustained high workload, suggesting that the hysteresis effect may have been a strong contributor." Huey and Wickens, in their 1993 report for the National Research Council, theorize that workload transitions could similarly affect "commercial airline crews, nuclear power plant control room crews, railroad freight train crews, merchant and military ship crews, natural disaster relief teams, emergency medical services crews, and trauma center and emergency room crews" (pg 3).

These examples demonstrate that the effects of workload transitions may have a far-reaching impact on our safety. As a

result, there have been several lines of research examining the impact of workload transitions on performance. However, the overall results have been mixed. There has been some agreement that a shift from high-to-low task demand results in a performance penalty as compared to constant low demand (Cox-Fuenzalida, 2007; Cumming and Croft, 1973; Gluckman, Warm, Dember, and Rosa, 1993; Goldberg and Stewart, 1980; Matthews, 1986; Ungar, 2008), while a shift from low-to-high task demand has variously resulted in improvement (Matthews, 1986), no change (Cox-Fuenzalida, 2007) or a decline in performance (Krulowitz, Warm, and Wohl, 1975) as compared to an unshifted baseline, where available.

One possible explanation for the variability in results obtained by different researchers and with different tasks is that the exact magnitude of workload transition delivered has been variable; if the task demand is changed for each participant by a fixed amount, the impact on workload will vary across individuals and tasks. Consequently, the present study adjusted the task demand to produce a consistent difference in performance between conditions across participants.

Additionally, unlike most previous studies on the topic, this study utilized a multi-task environment. Using the Air Force Multi-Attribute Task Battery (AF-MATB; Miller, 2010), we sought to determine how workload transitions in a dynamic environment would affect participants' EEG, ECG and performance.

## METHODS

## Participants

Sixteen paid participants from local universities completed the study. Eleven were male and five were female. Their ages ranged from 18 to 28. Participants completed a comprehensive written informed consent, and all study procedures were reviewed and approved by the Air Force Research Laboratory Institutional Review Board.

## Apparatus

This study employed the Multi-Attribute Task Battery (AF-MATB; Miller, 2010; Comstock & Arnegard, 1992), which was run on a Micron Personal Computer (MPC) Client Pro 565, and the task was mirrored to a standard LCD, 19 inch monitor situated (18 inches) in front of the participant. Participants interacted with the AF-MATB tasks through a keyboard, mouse and a Logitech Extreme 360 Pro joystick.

Physiological data recording and analysis replicated Wilson and Russell (2007). EEG was recorded from five channels, at F7, Fz, Pz, T5, and O2, positioned according to the International 10-20 electrode system (Jasper, 1958), using an Electro-Cap (Electro-Cap International, Inc., Eaton, OH). Reference and ground electrodes were positioned on the mastoid processes with impedances verified below 5 kilohms. Horizontal and vertical Electrooculogram (HEOG and VEOG, respectively) and ECG were also recorded with the use of standard Ag/AgCl electrodes. A Cleveland Medical Devices, Inc. (now Great Lakes NeuroTechnologies, Cleveland, OH), BioRadio 110 telemetry unit was used to acquire these data channels with a sampling rate of 200 Hz (12-bit resolution, band-pass filtered between 0.5 and 52.4 Hz). Corrections for eye movement and blinks were made with the use of an online implementation of an adaptive filter with HEOG and VEOG used as reference noise channels (He, Wilson, & Russell, 2004; He, Wilson, Russell, & Gerschutz, 2007). Interbeat interval (IBI), calculated across a 10-second window, was derived from the ECG channel with the use of an open source algorithm (Hamilton & Tompkins, 1986; Pan & Tompkins, 1985). Similarly, blink rate, calculated across a 30-second window, was derived with the use of an open source algorithm developed by Kong and Wilson (1998).

The EEG data were filtered into separate band-limited channels with elliptical infinite impulse response filter banks. The passbands for each channel were consistent with the five traditional bands of EEG: delta (0.5–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz), and gamma (31–42 Hz).

## Procedure

Participants were trained on the Air Force Multi-Attribute Task Battery (AF-MATB; Miller, 2010), which was adapted from the original version of the MATB (Comstock & Arnegard, 1992) to be compatible with current computer hardware. The AF-MATB is an experimental task that requires participants to perform a tracking task while concurrently monitoring warning lights and dials, responding to computer-generated auditory requests to adjust radio

frequencies, and managing simulated fuel flow rates using various key presses. Each subtask is individually scored. Participants completed six training sessions at 2 hours each. During those sessions, participants completed trials in a wide range of difficulties. Difficulty was manipulated by increasing the event rate in each of the subtasks, or in the case of the tracking component, difficulty was manipulated by adjusting speed and number of directional changes. After completing this training, individualized high difficulty levels were established via a simple staircase procedure that increased difficulty until proportions correct in the system management task fell below 65%. Low difficulty was consistent across participants and set at a level likely to produce ceiling performance. This process was intended to ensure that a large difference in workload was consistently imposed across participants despite varying aptitude for the task.

The testing session was composed of four blocks of six trials, and each trial was six minutes long. Two blocks were control trials with all easy or all hard trials (constant difficulty). The other two blocks were trials that transitioned from hard-to-easy or from easy-to-hard, with the transition occurring after minute three. The order of blocks was randomized for each participant. The NASA TLX (Hart & Staveland, 1988), which is a subjective workload measure, was administered after every trial. Participants were asked to focus on the second half of the just-completed trial while completing the TLX. Additionally, the shortened version of the Dundee Stress-State Questionnaire (DSSQ-3 State Questionnaire; Matthews, Emo & Funke, 2005) was administered at the beginning of the testing session and after the completion of each block to measure subjective task-related stress.

## RESULTS

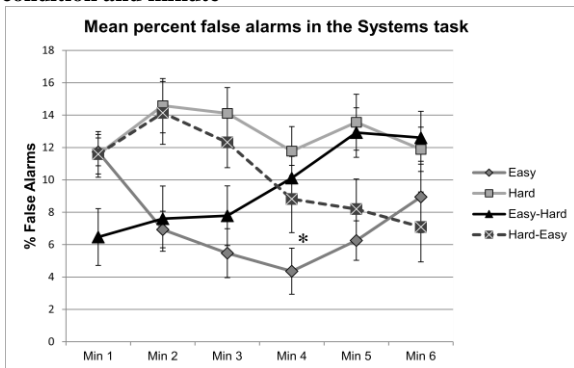
To analyze the performance data, trials were partitioned into 1-minute increments. Within each condition, minutes across trials were collapsed to produce minute-by-minute performance means for each participant. Significant 4 (difficulty condition) x 6 (minute) repeated measures ANOVAs for each performance measure were followed up with one-way ANOVAs to compare performance within each minute as a function of condition. For the measures with significant main effects of difficulty condition in this comparison, Tukey's HSD tests were used to determine which conditions had significantly different means. Because transitions in task difficulty occurred at the beginning of Minute 4 in the transition conditions, performance in Minutes 4, 5, and 6 were of particular interest in regard to the possible effects of workload transitions on performance. Specifically, it was meaningful if performance during Minutes 4, 5, or 6, in the transition scenarios was significantly different from its comparable control.

The analyses revealed that, in general, the significant differences between conditions, within each minute, were primarily demonstrated between easy and hard portions of the trials. However, significant differences were found between the easy control and hard-easy transition conditions in the Systems task in percent false alarms (Figure 1) and in the Tracking root

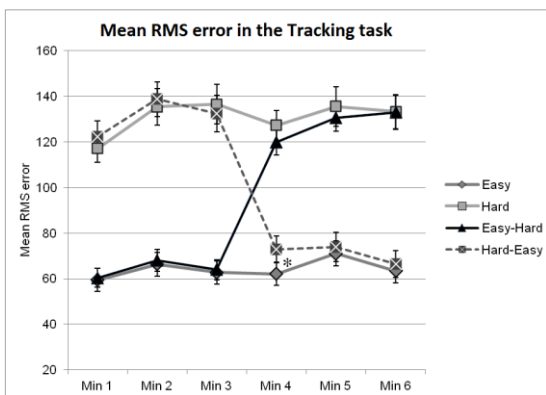
mean square (RMS) error (Figure 2). Tukey's HSD tests revealed that the RMS error in the Tracking task and the percent false alarms in the Systems task were both higher in the hard-easy transition condition than in the easy control condition during Minute 4 ( $p < .05$ ). An additional significant difference was demonstrated in the mean correct reaction times between the easy control and hard-easy transition conditions in Minute 6 of the Systems task, with higher mean correct reaction times occurring in the hard-easy condition as compared to the easy control condition (Figure 3). These results can be interpreted as lending some support to the position that workload transitions impacted performance in these conditions.

However, as illustrated in Figure 1 and Figure 3, unanticipated trends in performance in the easy condition may have had an impact on the noted differences between the easy and easy-hard conditions in Minutes 4 and 6 of the Systems task, thereby making clear interpretation of these results somewhat problematic as they pertain to transition effects.

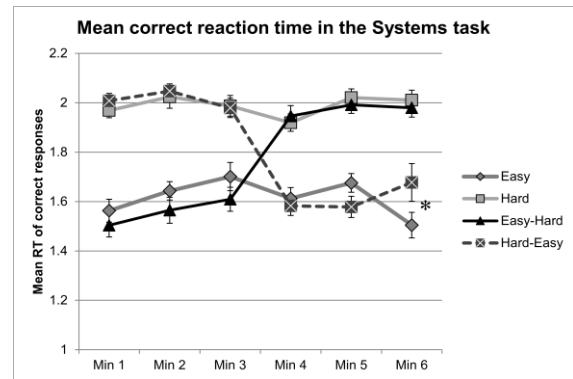
**Figure 1. Mean percent false alarms in the Systems task by condition and minute**



**Figure 2. Mean root mean square (RMS) error for the Tracking task by condition and minute**



**Figure 3. Mean correct reaction time in the Systems task by condition and minute**



Analysis of the TLX unweighted composite scores and of the six TLX subscale scores found only the expected main effect of workload, with no significant differences related to workload transitions. Likewise, the results of the shortened DSSQ were also not significant in terms of the workload transitions.

Preliminary analyses of the EEG data focused on frontal (Fz) theta and lateral (T5) gamma oscillations, based on prior research linking these measures with mental workload (e.g., Gevins, et al., 1997; Wilson & Russell, 2003). Parietal (Pz) alpha was also analyzed, but produced no noticeable effect of workload (consistent with Wilson & Russell, 2003). To reduce noise and stabilize time course estimates, power values in these bands were derived from non-overlapping, five second windows. These values were normalized within each participant via Z-transform, and then means and standard errors were calculated across participants for each condition. Figures 4 and 5 plot the resulting time courses for frontal (Fz) theta. In Figure 4, the constant difficulty trials are overlaid, showing that theta activity was overall higher in the hard trials. Figure 5 overlays both transition types, and illustrates rapid and complete crossover in this measure. Figure 6 plots the constant hard trials for T5 gamma activity and illustrates a strong decline at the beginning of the run that stabilizes to higher power in the hard condition. Figure 7 illustrates that the transition conditions produced crossover; however, the change following a hard-to-easy transition was relatively slow. This was quantified via calculation of 10-90 settling time, defined as the time between when the signal went from 10% below the mean value of the first 20 samples to within 90% of the last 20 samples; for T5 gamma following a hard-to-easy transition this was 45 seconds, whereas the easy-to-hard transition required only 10 seconds.

**Figure 4. EEG frontal theta power for easy and hard trials, plotted**

in this figure and subsequent as the mean and standard error across participants

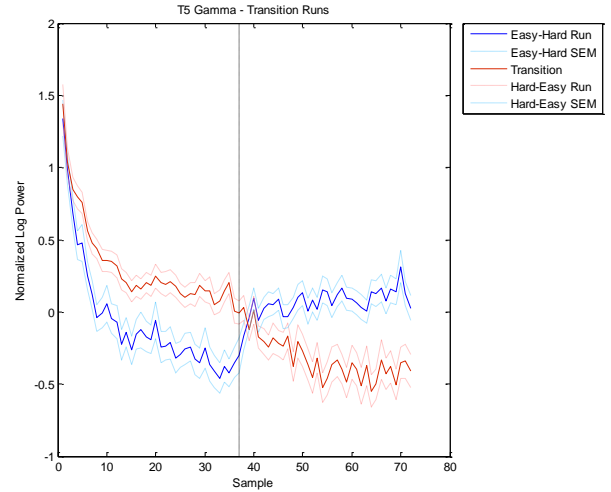
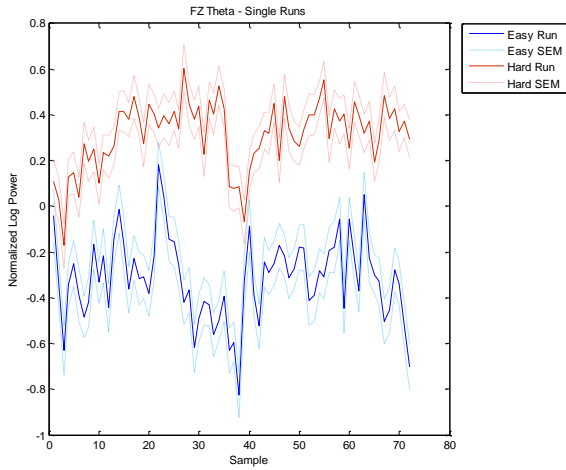


Figure 5. EEG frontal theta power, for both types of transition trials

## DISCUSSION

The current study sought to discover how transitions in workload might affect physiological and subjective responses as well as performance in a complex multitasking environment. The results of the NASA TLX and shortened DSSQ did not provide support for the position that workload transitions affect either subjective assessments of workload or task-related stress. It is possible that the subjective workload and stress measures were insufficiently sensitive to detect such effects; however, the lack of significant differences in most performance measures suggests that there was, indeed, no difference. Specific to the DSSQ, the shortened measure may be less sensitive than the full measure, and thus, may have obscured possible differences.

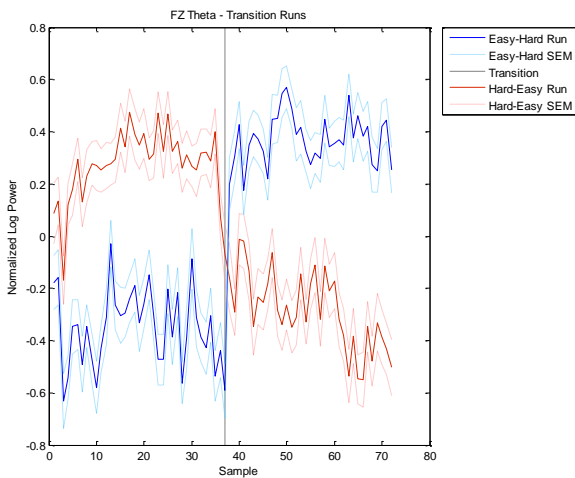


Figure 6. EEG temporal gamma power, for easy and hard trials.

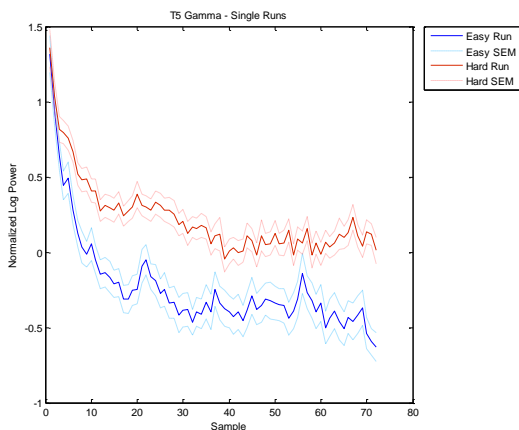


Figure 7. EEG temporal gamma power, for both types of transition trials.

Across the performance measures available in the AF-MATB task, most of the significant differences between conditions were between the easy and hard portions of trials. However, there was some evidence to suggest the presence a hysteresis effect in the hard-to-easy transitions, with significantly worse performance in RMS tracking error and in false alarms in the Systems task as compared to the consistently easy trials. These results could be explained as the participants' inability to rapidly adapt their approach to monitoring and responding to the Systems task when difficulty decreased, which has been previously described as strategic persistence (Matthews, 1986). However, the lack of a similar effect following an increase in task difficulty suggests that there may be some carryover or depletion of task performance ability, previously termed cognitive resource depletion (Ungar, 2008). Neither explanation is fully supported by these results. It is also possible that performance results were masked by individual differences in subtask prioritization and by variable prioritization with changes in task difficulty. Assigning primary-task status to one subtask and rank ordering the importance of the other subtasks could provide an interesting view into workload transitions in real-world situations where some subtasks are more critical than others. However, research pertaining to resource theory that has used primary and secondary tasks (Parasuraman & Hancock, 2001) has shown

that there are pitfalls associated with this method as well.

The preliminary analysis of EEG measures reveals a pattern of results somewhat similar to the performance results. Frontal theta oscillations exhibited a mean shift with changes in workload and rapid adaptation, with little or no hysteresis, following a transition. Frontal theta amplitude has previously been linked to activity in anterior cingulate cortex, with functional correlations involving sustained attention, conflict resolution, and rapid updating of working memory (Gevins, et al., 1997). With the increased demands for dividing attention and managing multiple subtasks in the hard AF-MATB condition, this result suggests that participants rapidly increased mental processing associated with balancing these demands.

Temporal gamma oscillations show a strikingly different pattern of results. There appears to have been a strong effect of task onset with declining gamma power for up to 60 seconds thereafter; subsequently, gamma was well-separated by task difficulty, which is consistent with previous results that suggest it is a salient feature for mental workload (Wilson & Russell, 2003). Following a transition, there was an asymmetry in gamma as a function of transition direction; transitions from easy-to-hard resulted in rapid change and settling in gamma, while transitions from hard-to-easy resulted in a slower settling similar to that observed with task onset.

The interpretation of power changes in the gamma band is somewhat unclear. Gamma oscillations in EEG have been linked to cognitive processes such as matching sensory stimuli to representations in memory (Herrmann, Munk, & Engel, 2004), which is most like the actions performed during the Systems task. The hysteresis in gamma following a hard-to-easy transition is consistent with the performance result that found hysteresis in the Systems task when transitioning from a hard to an easy task but not from an easy to a hard task. It is also possible that the observed changes in gamma reflect extracortical muscle artifact, a consequence of sympathetic changes in tonic muscle tension (Whitham, et al., 2007). Under this hypothesis, a slow decline in sympathetic activity following task onset or a decline in task difficulty could produce this pattern of results. The observed signals likely reflect both cortical and extracortical sources, and further analyses of cardiac activity collected in this study may help to support or refute sympathetic activity as a significant factor.

The pattern of EEG results observed following a transition suggests a pragmatic tradeoff for the practitioner: gamma activity was more clearly separated as a function of task difficulty than was frontal theta; however, theta did not exhibit noticeable hysteresis. In mental workload monitoring systems, it may thus be possible to trade speed of detection against overall accuracy. By favoring theta as a workload signal, we would expect rapid detection of a change in workload, but occasional misdetections. Favoring gamma should result in delays in detecting a decline in workload, but would ultimately result in a very stable signal with few misdetections.

In conclusion, this study found some evidence of performance effects of workload transitions; the data from several measures were consistent with the position that performance was negatively affected, post-transition, when

trials began at a high level of workload and then became less demanding. The EEG results provide some additional support for slow adaptation to a sudden decline in workload; however, previous research has been quite mixed as to which types of transitions impact performance. Further research on this topic, particularly in multi-tasking environments, is needed.

## REFERENCES

- Comstock, J. L., & Arnegard, R. J. (1992). *The Multi-attribute Task Battery for human operator workload and strategic behavior research*. Technical Report 104174. Hampton, VA: NASA Langley Research Center.
- Cox-Fuenzalida, L. E. (2007). Effect of workload history on task performance. *Human Factors*, 49, 277-291.
- Cumming, R. W., & Croft, P. G. (1973). Human information processing under varying task demand. *Ergonomics*, 16, 581-586.
- Farrell, P.S.E. (1999). The hysteresis effect. *Human Factors*, 41(2), 226-240.
- Gevins, A., Smith, M.E., McEvoy, L., Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7(4), 374-385.
- Gluckman, J.P., Warm, J.S., Dember, W.N., & Rosa, R.R. (1993). Demand transitions and sustained attention. *The Journal of General Psychology*, 120(3), 323-337.
- Goldberg, D. R., & Stewart, M. R. (1980). Memory overload or expectancy effect? "Hysteresis" revisited. *Ergonomics*, 23, 1173-1178.
- Hamilton, P.S., & Tompkins, W.J. (1986). Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *Biomedical Engineering, IEEE Transactions on*, 12, 1157-1165.
- Hart, S.G., & Staveland, L.E. (1988). Development of a multi-dimensional workload scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139- 183). Amsterdam: North-Holland.
- He, P., Wilson, G., & Russell, C. (2004). Removal of ocular artifacts from electro-encephalogram by adaptive filtering. *Medical and Biological Engineering and Computing*, 42(3), 407-412. <http://dx.doi.org/10.1007/BF02344717>
- He, P., Wilson, G., Russell, C., & Gerschutz, M. (2007). Removal of ocular artifacts from the EEG: A comparison between time-domain regression method and adaptive filtering method using simulated data. *Medical and Biological Engineering and Computing*, 45(5), 495-503. <http://dx.doi.org/10.1007/s11517-007-0179-9>
- Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences*, 8(8), 347-355.
- Huey, B.M., & Wickens, C.I. (Eds.). (1993). *Workload transition: Implications for individual and team performance*. Washington DC: National Academy Press. (National Research Council).
- Jasper, H.H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371-375.
- Kong, X., & Wilson, G. F. (1998). A new EOG-based eyeblink detection algorithm. *Behavior Research Methods, Instruments & Computers*, 30, 713-719.
- Krulowitz, J.E., Warm, J.S., & Wohl, T.H. (1975). Effects of shifts in the rate of repetitive stimulation on sustained attention. *Perception & Psychophysics*, 18, 245-249.
- Matthews, M. L. (1986). The influence of visual workload history on visual performance. *Human Factors*, 28, 623-632.
- Matthews, G., Emo, A.K., & Funke, G.J. (2005). A short version of the Dundee Stress State Questionnaire. *Presented at the Twelfth Meeting of the International Society for the Study of Individual Differences*, Adelaide, Australia, July 2005.
- Miller, W. D. (2010). The U.S. Air Force-developed adaptation of the Multi-Attribute Task Battery for the assessment of human operator workload and strategic behavior (Technical Report AFRL-RH-WP- TR-2010-0133). Retrieved from the Defense Technical Information Center website: <http://www.dtic.mil/dtic/tr/fulltext/u2/a537547.pdf>
- Morgan, J.F., & Hancock, P.A. (2011). The effect of prior task loading on mental workload: An example of hysteresis in driving. *Human Factors*, 53(1), 75-86.
- Pan, J., & Tompkins, W.J. (1985). A real-time QRS detection algorithm.

- Biomedical Engineering, IEEE Transactions on*, 3, 230-236.
- Parasuraman, R., & Hancock, P.A., (2001). Adaptive control of mental workload. In R. Hancock & P.A. Desmonds (Eds.) *Stress, workload and fatigue* (pp. 305-320), Mahwah, NJ: Erlbaum.
- Ungar, N.R. (2008). Effects of transitions in task-demand on vigilance performance and stress. Unpublished doctoral dissertation, University of Cincinnati, Cincinnati, OH.
- Whitham, E.M., Pope, K.J., Fitzgibbon, S.P., Lewis, T., Clark, C.R., Loveless, S., et al. (2007). Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology*, 118, 1877-1888.
- Wilson, G. F., & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, 45(4), 635-644.
- Wilson, G.F. & Russell, C.A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Human Factors*, 49(6), 1005-1018. DOI 10.1518/001872007X249875