St. John, M. & Risser, M. R. (2009). Sustaining vigilance by activating a secondary task when inattention is detected. In *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

SUSTAINING VIGILANCE BY ACTIVATING A SECONDARY TASK WHEN INATTENTION IS DETECTED

Mark St. John & Matthew R. Risser
Pacific Science & Engineering Group, Inc.
San Diego, CA, USA

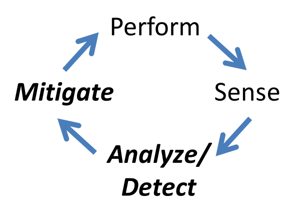
Vigilance tasks, from driving a vehicle to surveillance to security monitoring, are both commonplace and high-stakes. Yet users have well known difficulties sustaining vigilance. We evaluate the ability of an augmented cognition closed-loop attention management (CLAM) system to sustain vigilance and task performance by monitoring operator's psychophysiology, detecting inattention, and activating a countermeasure when inattention crosses a threshold. Eighteen participants performed a vigilance task and were monitored for inattention via a combination of eye, head, and electroecephalographic (EEG) measures. A cognitively demanding secondary task was activated either when inattention was detected or randomly throughout a 40 minute session. While participants in both conditions demonstrated a vigilance decrement, as measured by an increase in misses over the course of the session, the CLAM condition produced 17% fewer misses overall than the random condition. This improvement was not due to the countermeasure, per se, but to the timing of the countermeasure to participant's detected inattention. The advantage for a tailored presentation of the secondary task is noteworthy because prior evaluations of *continuous* secondary tasks demonstrated degraded vigilance performance. The results inform our understanding of how human vigilance operates and the technology for its detection and manipulation.

Vigilance tasks, such as video surveillance from remote vehicles, security operations, automation supervision, and long-distance driving, are both commonplace and high-stakes. Unfortunately, vigilance tasks are highly repetitive and understimulating, and operators struggle to sustain vigilance for even short stretches of time. A relatively new method for sustaining attention is to monitor operators' psychophysiology for signs of inattention and trigger a countermeasure in order to sustain it at an appropriate level for good task performance. We call this concept Closed-Loop Attention Management (CLAM).

A CLAM system is a form of adaptive automation (e.g., Byrne & Parasuraman, 1996; Scerbo, 2005), though in this case, the "automation" activates a countermeasure to re-engage the operator's attention rather than automating task control when the system detects user overload.

Psychophysiologically-based closed-loop systems have been developed for both versions of adaptive automation. Regarding inattention, Berka et al. (2005) measured EEG correlates of vigilance and sounded alarms when participants in a driving simulator became drowsy, and Mikulka, Scerbo, and Freeman (2002) measured EEG correlates of vigilance and changed the presentation rate of stimuli when participants lost vigilance. Regarding overload, Prinzel et al. (2003) and Wilson, Lambert, & Russel (1999), for example, measured EEG correlates of mental workload and automated some tasks when excessive workload was detected.

An effective closed-loop system requires overcoming a variety of technical and scientific hurdles including the development of comfortable and wireless equipment, psychophysiological measures that are sensitive to vigilance levels, methods for combining multiple measures in real-time for accurate and timely detection of inattention, and effective countermeasures to rouse the operator and mitigate any vigilance decrement. Much recent progress has been made toward addressing each of these hurdles. Our research has focused primarily on developing methods for combining multiple contemporary psychophysiological measures and developing effective countermeasures to mitigate the vigilance decrement.



Here, we report an evaluation of a complete closed-loop system composed of 1) a combination of eyelid opening, head posture, and EEG measures of inattention and 2) a novel countermeasure composed of a cognitively demanding secondary task. The results enrich our understanding of how human vigilance operates and the technology for its detection and manipulation.

A number of psychophysiological measures have shown promise for detecting inattention either alone or in combination including EEG, eye movements, and head and body posture (e.g., Balaban et al., 2004; Duta, Alford, Wilson, & Tarassenko, 2004; Jung, Makeig, Stensmo, & Sejnowski, 1997; Van Orden, Jung, & Makeig, 2000). Recently, St. John, Risser, and Kobus (2006) found that a combination of a derived EEG measure of task engagement, percent eye opening, and variability of head pitch (nodding) predicted 42% of the variance in the miss rate during a vigilance task. No individual measure accounted for more than 13% of the variance.

Additionally, many countermeasures for sustaining vigilance have been evaluated, though most have involved changes to the vigilance task, itself (for a review, see See, Howe, Warm, & Dember, 1995). St. John and Risser (2007) evaluated the potential for an *intermittent* secondary task to serve as a vigilance countermeasure. Prior evaluations of *continuous* secondary tasks, both cognitive and vigilance tasks, have demonstrated worse performance and larger vigilance decrements (e.g., Alluisi & Hall, 1963; Craig, 1981). However, our concept was to tailor the presentation of the secondary task specifically to times when participants were inattentive, so as not to tax participants' resources, yet still increase mental demands at critical moments. In the 2007 study, the secondary task was activated when participants missed a target – a sign of inattention. In the current study, the secondary task was activated when psychophysiological measures indicated inattention.

The secondary task was a digit task that required auditory processing and verbal working memory to store and resequence digits, while the vigilance task required visual processing and spatial working memory to compare the visual stimulus against the remembered target. The digit task was as effective as an alarm for reducing the vigilance decrement, but it was rated as less annoying by participants – an important practical consideration.

To truly demonstrate the effectiveness of a closed-loop augmented cognition system, the psychophysiologically triggered countermeasure must be more effective at sustaining vigilance than the same countermeasure triggered randomly via an open-loop. Therefore, the present study compared the closed-loop-triggered countermeasure to a randomly triggered countermeasure. We hypothesized that the closed-loop-triggered countermeasure would sustain vigilance and reduce misses better than the randomly triggered counter-measure because the increased demands from the secondary task would be synchronized with periods of inattention.

METHOD

Participants. Participants were recruited from www.craigs list.org/software-QA-DBA and the University of California, San Diego in order to find participants who were knowledgeable of computers and high-tech equipment. Participants were further screened for a good knowledge of English and having or in the process of obtaining a bachelor degree. Twenty-eight people consented to participate in the study. Ten participants were excluded for a variety of reasons: four failed to pass the EEG baseline tasks due to equipment problems or unusual EEG patterns, four had incomplete eye tracking data due to excessive body motion or unusual interference from their glasses, and two failed to pass the training criterion for the primary vigilance task. Lowering this exclusion rate is an important practical goal.

Of the 18 participants who completed the experiment and are included in the analysis, 10 were male and eight were female. They had a mean age of 32, ranging from 19 to 59. Participants reported getting slightly over seven hours of sleep per night on average, and all participants scored normally ($\leq$ 10) on the Epworth Sleepiness Scale (Johns, 1992), indicating that no participants reported excessive daytime sleepiness that might affect the EEG or vigilance task results.

The experiment lasted approximately four hours, from 10am to 2pm. EEG baseline measures were obtained in the morning and the vigilance study was conducted in the early afternoon. Participants were paid $60 and provided a caffeine-free lunch.

Apparatus. A SmartEye Pro desktop eye tracking system, version 4.5.4, was used to record head movement and percent eye opening. The system includes two 60 Hz cameras located on either side of the computer monitor that allowed the participant to move freely with minimal data loss. Head pitch was measured as a rotation around the X axis (through the ears). Head pitch variability (nodding) was computed as the standard deviation of head pitch over a 60-second window. Percent eye opening was measured as the vertical distance between the upper and lower eyelids. It was made relative to each participant by computing the percentage deviation from a baseline open eye average. The open eye average was obtained during an early training session when participants were alert and engaged. Artifacts were removed in real-time from the head and eye data by discarding data when the system could not find the head and at least one eye.

A B-Alert wireless EEG system computed an index of task engagement (Berka, et al., 2007). Their index consists of a four-class quadratic discriminant function analysis tailored for each participant from a set of baseline tasks conducted prior to the experiment. The model was constructed using absolute and relative power spectra variables from Fz-POz and Cz-POz obtained using stepwise regression on a database of fully rested and sleep-deprived participants. The index computes the probability that each second of EEG data falls into each of four levels of task engagement: high engagement (HE), low engagement (LE), relaxed wakefulness (predominance of alpha), or sleepy (predominance of theta).

In the CLAM condition, timestamped data packets from the eye tracker and the EEG system were queued and synched to the master experiment clock in real-time using custom software. The synched data were sampled every two seconds and a mean value was computed over the prior 60 seconds. This procedure created a 1-minute-wide sliding window that was updated every two seconds. The measure of inattention, called the predicted miss rate (PMR), was computed by multiplying each measure by a weight and adding a constant (see equation 1). The values for the weights and the constant were computed by a step-wise regression using data obtained from nine participants performing the identical vigilance task without any countermeasures (St. John, Risser, & Kobus, 2006). The original data were recomputed using the relative eye opening measure. Head pitch variability (nodding) increased the predicted miss rate, percent eye opening (fully open equals 1.0 and fully closed equals 0.0) decreased the predicted miss rate, and low engagement increased the predicted miss rate.

(1) $PMR = 2.1 * HeadPitchVar – 1.11 * PercentEyeOpening + .62 * LowEngagement + .99$

The countermeasure threshold was determined through pilot testing and was triggered when the predicted miss rated exceeded 0.5. The countermeasure was also triggered if the predicted miss rate could not be computed for 60 consecutive seconds, typically because the participant fell out of view of the eye tracker. Once triggered, the countermeasure could not be triggered again for 60 seconds. This period provided time for the countermeasure to be presented and affect participants' vigilance levels. Onset of the countermeasure could co-occur with a vigilance target, as would be the case in the real world. This co-occurrence of stimuli could cause interference (St. John & Risser, 2007).

In the random condition, the countermeasure was triggered on a pseudorandom schedule with an interval of 120 seconds plus a random deviation of +/- 30 seconds. Consequently, the countermeasure was triggered 19 or 20 times per 40-minute session. The interval was set through pilot testing to roughly match the average number of countermeasures occurring in the CLAM condition.

Task and Stimuli. Participants performed a 40-minute vigilance task that was couched as an unmanned aerial vehicle (UAV) surveillance task, monitoring a simulated video feed taken from a UAV flying along a highway (St. John, Risser, & Kobus, 2006). Participants monitored snapshots of the highway for long-bed trucks because local insurgents were known to use long-bed trucks for subversive activities (see Figure 2). In these respects, the task was more visually complex than typical vigilance tasks, but it maintained the essential characteristics of a vigilance task (See, Howe, Warm, & Dember, 1995).

The target trucks had 10% longer beds (110 pixels) than the distracter trucks (100 pixels). One truck appeared every two seconds for 400ms at one of six locations along a vertically oriented road displayed on a 17” LCD monitor set at a resolution of 1024 x 768 pixels. Three targets appeared per minute and were presented randomly and nonconsecutively among 27 distracters, a target rate of 1 in 10. Participants responded to targets by pressing the space bar. No feedback was provided. A miss was recorded when a participant failed to respond to a target within two seconds of target onset. Hits, misses, false alarms, and response times were recorded for each trial.

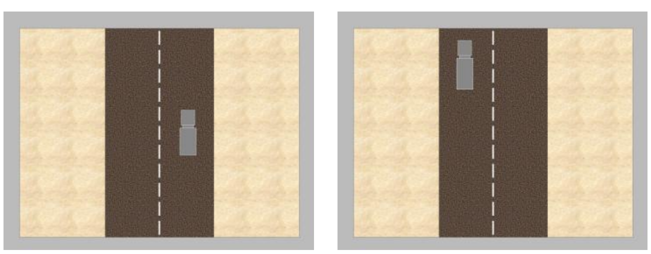


Figure 2. Screenshots of the vigilance task. Long and short-bed trucks appeared briefly at six locations along the road; (left) a distracter short-bed truck, (right) a target long-bed truck.

The countermeasure was an auditory secondary task. Participants heard a string of three digits, and they were asked whether they could be reordered to be consecutive with no gaps, for example, “1-3-2” versus “1-4-2.” Participants responded by pressing the left and right arrow keys. A miss was recorded when a participant failed to respond within 3.5 seconds of the auditory offset. Auditory feedback was provided for correct and incorrect responses. Each countermeasure consisted of three trials of the digit task. The total time for the three trials, including time for participants’ responses, was about 18 seconds.

Procedure. After signing consent forms, participants were prepped and fitted with the B-Alert EEG cap. They then performed a series of three 5-minute B-Alert baseline tasks to derive their unique EEG profile. Next, participants were trained on the vigilance task.

First, participants completed a two-minute demonstration trial with auditory feedback and a target rate higher than that used during the experimental task (i.e., 1 in 3). Second, participants completed a three-minute practice trial with auditory feedback, but with the target ratio the same as it would be during the experimental task (1 in 10). Participants repeated this trial until they reached criterion performance of three or fewer errors (misses or false alarms). Participants who failed to reach the criterion after four tries were dismissed. Third, participants were provided a 30-minute caffeine-free lunch. Fourth, participants were reacquainted with the vigilance task with another three-minute trial. They were not required to meet the criterion again. Fifth, the secondary task was described, and participants practiced for 10 trials with feedback. Sixth, the vigilance task and the secondary task were then combined for a final three-minute practice trial. During this trial, the secondary task occurred on a fixed schedule so that all participants received equal experience. Feedback was provided for both tasks.

Finally, participants received two sessions of the experimental task, once in the CLAM condition and once in the random condition. The conditions occurred in a counterbalanced order. This within-participant design controlled for the large individual differences in both task performance and ability to sustain vigilance.

Participants were told there were two control procedures for triggering the countermeasure, and we were testing which one worked better. Participants were blind to the test conditions, and EEG, head, and eye data were recorded for both sessions. Feedback was provided only for the secondary task. Each session lasted 40 minutes. Participants received a short break between sessions and were instructed to walk around the building for five minutes. Before starting the next session, the participant was retrained with the three-minute vigilance task and a three-minute combined vigilance and secondary task, with feedback, to establish the same level of training going into each experimental session.

RESULTS AND DISCUSSION

The primary hypothesis of the experiment was that CLAM control over the countermeasure trigger would sustain vigilance and task performance better than random control over the countermeasure trigger because the CLAM system would tailor the countermeasure to occur specifically when participants were inattentive and the countermeasure was most needed. To test this hypothesis, as well as examine the time course of performance, each participant’s miss rate was computed in five-minute consecutive blocks across the 40-minute session (see figure 3).

The five-minute data were submitted to a repeated measures ANOVA with condition (CLAM or random) and block as repeated factors and order of sessions as a between factor. The effect of condition was significant, $F(1,16) = 5.3$, $p = .035$. CLAM control over the countermeasure led to fewer misses than random control. Across all eight blocks (40 minutes) the average rate of misses under CLAM control was 0.30, and the average rate of misses under random control was 0.36. Therefore, CLAM improved the miss rate by 17%.

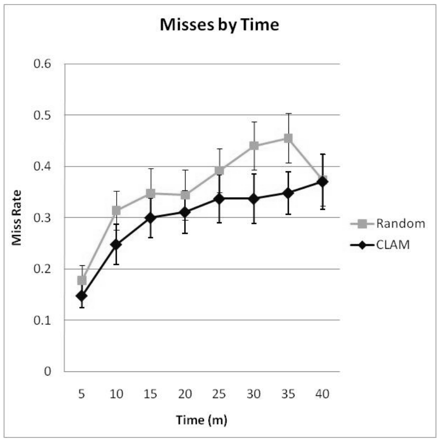


Figure 3. The time course of misses.

The effect of block was also significant, $F(7, 112) = 9.3$, $p < .0001$. The miss rate in both conditions followed the classic vigilance decrement pattern, with a sharp increase in the miss rate during the first 10 to 15 minutes of a session, followed by a more subtle increase or steady state. The order of conditions was not significant, $F(1, 16) = .91$.

For response times, only the effect of block was reliable, $F(7, 112) = 9.7$, $p < .0001$, with response times increasing across blocks.

The predicted miss rate also showed a significant difference between conditions $F(1,15) = 9.0$, $p = .009$. The number of participants in this analysis was reduced by 1 because one participant was out of view of the eye tracker for one block, and no predicted miss rate could be computed for that block. When participants were in the random condition, they demonstrated an immediate and consistently higher predicted miss rate compared with the CLAM condition.

Following both sessions, participants were asked several questions about their subjective experiences. Thirteen out of 18 participants (72%) reported that they performed better in the CLAM condition than in the random condition, $p = .0002$ by a one-sample sign test. These 13 participants showed an 18% benefit from CLAM over the random condition while the other 5 participants showed only an 8% benefit from CLAM, but this difference was not significant, $F(1,16) = .37$. A somewhat different set of 13 out of 18 participants (72%) reported the CLAM condition provided better-timed countermeasures than the random condition. These participants showed a 19% benefit from CLAM over the random condition while the other 5 participants showed only a 10% benefit from CLAM, again this difference was not significant, $F(1,16) = .10$.

The countermeasure was triggered an average of 18 times in the CLAM condition and 19 times in the random condition. This similarity means that the benefits of CLAM were not due to differences in countermeasure presentation rates between the two conditions.

While the number of countermeasures did not vary across conditions, their time-courses did vary. The interaction between block and condition was significant, $F(7, 112) = 13.7$, $p < .0001$. As shown in Figure 4, the number of countermeasures per five minute block in the random condition held fairly constant across the session, in accord with the uniform random distribution that triggered the countermeasures. The number of countermeasures per five minute block in the CLAM condition, however, started low and increased sharply across the session, in accord with the predicted miss rate and task performance. This distribution demonstrates the tailoring of the closed-loop countermeasure schedule to participants' levels of inattention.

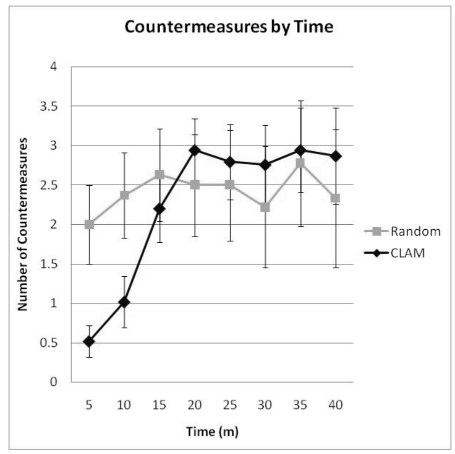


Figure 4. The time course of countermeasures between the CLAM and random conditions.

CONCLUSIONS

The UAV vigilance task was relatively difficult. Even in the CLAM condition, participants missed 30% of the targets. Nonetheless, the 17% improvement over the random condition represents a substantial improvement. This improvement is especially impressive since the random condition contained the same type and number of countermeasures as the CLAM condition. Therefore, the improvement is not due to the countermeasures, *per se*, but rather to tailoring the timing of the countermeasures to participants' detected levels of inattention.

However, because the study did not include a baseline condition with no countermeasures, this study cannot determine whether there is an overall benefit of the countermeasure or just a relative improvement of tailored presentations versus random presentations. A prior study (St. John & Risser, 2007) did demonstrate an improvement in vigilance performance for the tailored secondary task compared with a no countermeasure baseline, although in that study, the countermeasure was activated by missed targets rather than psychophysiology. It can also be argued that the benefit of tailoring the presentation of the countermeasure would only occur if the countermeasure was, in fact, beneficial. If it didn't help rouse participants, then it would not prove especially effective when synchronized with periods of inattention. In any case, a future study to explicitly test CLAM-driven tailoring against a baseline is planned.

Finally, it is intriguing and theoretically significant that tailoring the presentation of a secondary task increases vigilance task performance. Recent research has documented that vigilance tasks are stressful and workload increases as participants struggle to remain vigilant (Grier, Warm, Dember, Matthews, Galinsky, & Parasuraman, 2003). Similarly, resource theories suggest that the vigilance decrement occurs as resources are drained and not replenished (Warm, Dember, & Hancock, 1996; Wickens, 1991). Nonetheless, we found that adding an intermittent secondary task sustains vigilance and task performance. These results suggest that the additional demands of the secondary task actually replenish resources, perhaps by a generalized activation of resources, to both support the secondary task as well as re-engage participants with the vigilance task (St. John & Risser, 2007). The finding is especially interesting since prior vigilance studies demonstrated that continuous secondary tasks degraded vigilance performance, as noted above. The tailoring and intermittency of our secondary task apparently reversed the effect, perhaps because its synchronization with periods of inattention proved more stimulating than demanding. More study is needed to explicitly test this hypothesis and understand the theoretical implications.

These results constitute strong support for the closed-loop attention management concept. Achieving this result required 1) contemporary, sophisticated psychophysiological measures of inattention, 2) their real-time combination and computation to trigger a countermeasure, and 3) a secondary cognitive task tailored to inattention and designed to rouse participants by increasing effort and improve vigilance task performance.

Still, much remains to be accomplished to refine the choice of measures, account for individual differences in predicting inattention, and understanding the mechanisms underlying the effect of the tailored secondary task.

ACKNOWLEDGEMENTS

The authors would like to thank Dan Manes for programming the experiment and the real-time data analysis and Linda Wu for conducting participants through this complex study. This research was supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number W911NF-05-C-0094.

REFERENCES

Alluisi, E. A. & Hall, T. J. (1963). Declines in auditory vigilance during periods of high multiple-task activity. *Perceptual and Motor Skills, 16,* 739-740.

Balaban, C. D., Cohn, J. V., Redfern, M. S., Prinkey, J., Stripling, R., & Hoffer, M. (2004). Postural control as a probe for cognitive state: Exploiting human information processing to enhance performance. *International Journal of Human-Computer Interaction, 17,* 275-286.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine, 78(5), Section II, Supplement,* B231-B244.

Berka, C., Levendowski, D.J., Westbrook, P., Davis, G., Lumicao, M.N., Ramsey, C., Petrovic, M.M., Zivkovic, V.T., & Olmstead, R.E. (2005). Implementation of a closed-loop real-time EEG-based drowsiness detection system: Effects of feedback alarms on performance in a driving simulator. In *Proceedings of the 11th Annual Conference on Human-Computer Interaction*, Las Vegas, NV: LEA.

Byrne, E. A. & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42,* 249-268.

Craig, A. (1981). Monitoring for one kind of signal in the presence of another: The effects of signal mix on detectability. *Human Factors, 23,* 191-197.

Craven, P. L., Belov, N., Tremoulet, P., Thomas, M., Berka, C., Levendowski, D., & Davis, G. (2007). Cognitive workload gauge development: Comparison of real-time classification methods. In D.D. Schmorrow, D.M. Nicholson, J.M. Drexler, & L.M. Reeves (Eds.), *Foundations of Augmented Cognition, 4th Edition* (pp. 66-74). Arlington, VA: Strategic Analysis, Inc.

Duta, M., Alford, C., Wilson, S., & Tarassenko, L. (2004). Neural network analysis of the mastoid EEG for the assessment of vigilance. *International Journal of Human-Computer Interaction, 17,* 171-199.

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors, 45*(3), 349-59.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep, 15,* 376-381.

Jung, T-P., Makeig, S., Stensmo, M., & Sejnowski, T. (1997). Estimating alertness from the EEG power spectrum, *IEEE Transactions on Biomedical Engineering, 44,* 60-69.

Mikulka, P. J., Scerbo, M. W., & Freeman, F. G. (2002). Effects of a biocybernetic system on vigilance performance. *Human Factors, 44,* 654-664.

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component. *Human Factors, 45,* 601-613.

Scerbo, M. W. (2005). Adaptive automation. In R. Parasuraman & M. Rizzo (Eds.), *Neuroergonomics: The brain at work.* Oxford, Great Britain: Oxford University Press.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). A meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin, 117,* 230-249.

St. John, M., & Risser, M. R. (2007). Mitigating the vigilance decrement: Perceptual versus cognitive interventions. In D.D. Schmorrow, D.M. Nicholson, J.M. Drexler, & L.M. Reeves (Eds.), *Foundations of Augmented Cognition, 4th Edition* (pp. 81-89). Arlington, VA: Strategic Analysis, Inc.

St. John, M., Risser, M. R., & Kobus, D. A. (2006). Toward a usable closed-loop attention management system: Predicting vigilance from minimal contact head, eye, and EEG measures. In D. Schmorrow, K. Stanney & L. Reeves (Eds.), *Foundations of Augmented Cognitio*n, 2nd Ed. (pp. 12-18). Arlington, VA: Strategic Analysis.

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 183-200). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-task Performance* (pp. 3-34). London: Taylor and Francis.

Wilson, G. F. & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors, 45,* 635-643.

Van Orden, K. F., Jung, T. P., Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology, 52,* 221–240.

Wilson, G. F., Lambert, J. D., & Russell, C.A. (1999). Performance enhancement with real-time physiologically controlled adaptive aiding. *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting,* 3-61 to 3-64.