2009

Vigilant Warrior™: A Selection Tool for Vigilance Performance

David E. Weldon Ph.D.

Clark Shingledecker Ph.D.

VIGILANT WARRIOR™: A SELECTION TOOL FOR VIGILANCE PERFORMANCE

David E. Weldon, Ph. D.
JXT Applications, Incorporated
Beavercreek, Ohio
Clark Shingledecker, Ph. D.
Wright State University
Dayton, Ohio

In this paper, we describe an individual differences model of vigilance performance—the ability to maintain one's focus of attention and remain alert for prolonged periods of time—and summarize our model evaluation research. Our goal was an automated test battery (*Vigilant Warrior*™) that could be employed to select personnel with superior abilities for assignment to critical vigilance duties. Thus, we conducted extensive laboratory research to identify an optimal set of vigilance predictors and validate them against a simulated, real-world, electronic-display, battlefield-monitoring task with high vigilance requirements. The results confirmed that an objective, Short Vigilance Task (SVT), coupled with analytic skill and stress-coping measures, could account for 33% or more of the criterion variance. Moreover, the SVT was the most powerful predictor in the battery. Analytic skill and situational variables contributed to vigilance performance, but to a lesser degree. *Vigilant Warrior*™ is currently receiving extensive field testing in military settings.

Vigilance is the ability to maintain one's focus of attention and remain alert for prolonged periods of time. As such, vigilance is a key cognitive attribute for exceptional performance over a wide range of work domains where the ability to detect and respond to relatively rare and sometimes obscure events must be sustained despite lengthy duty requirements. Tasks requiring a high degree of vigilance are an integral to warfare. In addition to conventional visual monitoring activities, the modern warfighter is likely to engage in computer-mediated monitoring tasks associated with control of aircraft, missiles, unmanned aerial vehicles, or combat robots, and perform detection tasks in efforts to counter enemy threats. Past research has shown that individuals vary widely in their capacity to be vigilant in these situations. Therefore, a need exists to identify and selectively assign individuals with exceptional vigilance performance capabilities to critical jobs with high, sustained attention demands. This paper summarizes the theoretical basis for the development of *Vigilant Warrior*™: a new personnel selection battery designed to identify individuals who display exceptional vigilance performance. It also describes the results of research conducted to refine and validate the predictive abilities of the *Vigilant Warrior*™ battery.

A Model for Development of a Vigilance Selection Test Battery

Previous attempts to identify measures or factors reflecting differences among individuals that reliably predict vigilance performance have been largely unsuccessful. One likely reason for this failure is that approaches that were taken to the problem were typically based solely on single personality characteristics. We developed the *Vigilant Warrior*™ test battery to remedy this shortfall by adopting a multidimensional view of the prediction problem, guided by current theoretical treatments of vigilance and a by a broad examination of past vigilance research findings. This perspective raises the possibility that improved vigilance prediction may be possible by combining information derived from classical personality variables with measures of intelligence, sample vigilance task performance, and measures of the person's characteristic responses to vigilance task demands. We summarize the literature supporting this approach to predicting individual differences in vigilance in the following paragraphs.

Personality factors. Davies & Parasuraman (1982) summarize the findings for personality dimensions related to vigilance performance; including introversion-extraversion (introverted observers outperform their extraverted cohorts), field dependence-independence (field-independent individuals outperform field-dependent observers), internal-external locus of control (individuals with an internal locus of control outperform those with an external locus of control), and the Type A (coronary-prone) behavior pattern (achievement-oriented Type-A individuals outperform their more relaxed, Type-B counterparts). In addition, Thackray, Bailey, & Touchstone (1977) found that boredom prone individuals may be poorer monitors than those less boredom prone while Robertson, et al. (1997) found that absent-minded individuals, defined by high scores on the Cognitive Failures Questionnaire, did more poorly in than non-absent minded observers and reported higher levels of perceived mental workload than the non-absent minded. Finally, Helton, Dember, Warm, & Matthews (1999) found that optimists perform more effectively on vigilance tasks than do pessimists. Such results indicate that personality profiles should be included as candidates for any approach for developing a vigilance test with reliable predictive features.

Performance sampling as a predictor. A second promising source of predictors of sustained attention ability is the objective measurement of an individual's performance on vigilance tasks themselves. However, traditional laboratory vigilance tasks require a lengthy watch period that would make them impractical as selection tests for large groups of examinees. Recent research, however, shows that brief, highly-demanding, vigilance tasks can be constructed that produce performance that mirrors the vigilance decrements typically observed in long-term vigils (e.g., Matthews, Davies & Lees, 1990; Temple et al., 2000). These tasks show rapid perceptual sensitivity decrements over a period of 10 minutes or less. They also they demonstrate the key diagnostic indicators of being resource-limited: sensitivity decrement, high subjective workload, and sensitivity to stress and arousal factors. Thus, a high level of performance on a short task may be a good indicator of aptitude for longer vigilance tasks.

Differences in subjective responses to vigilance task demands. Finally, recent studies indicate that the perceived workload of vigilance tasks is quite substantial and that workload grows linearly over time (Warm, Dember, & Hancock 1996). Johnson & Proctor (2003) conclude that, rather than being under-stimulating, vigilance tasks place high information-processing demands upon observers. Thus, Resource Theories appear to take precedence over Arousal Theory as models of the factors controlling vigilance performance. However, following Kahneman (1973), Matthews and Davies (2001) argued that Arousal and Resource Theories are not necessarily mutually exclusive, and that they can be integrated by viewing arousal as the agent responsible for resource production. The finding that there seems to be agreement between psychophysiological measures, subjective self-reports, and performance, as predicted by the integrated models, is of considerable significance for selection test development. In addition to workload response differences, Hancock & Warm (1989) found that operators differed in the way they deployed compensatory effort and coping strategies to adapt to demanding performance environments. Short tasks are sometimes insensitive to stressor effects, but as time progresses it becomes increasingly more difficult for the operator to maintain successful coping. Therefore, it may be possible to identify useful predictor measures from an operator's reactions to performing a short vigilance task, which may offer early warning signs of difficulties in coping.

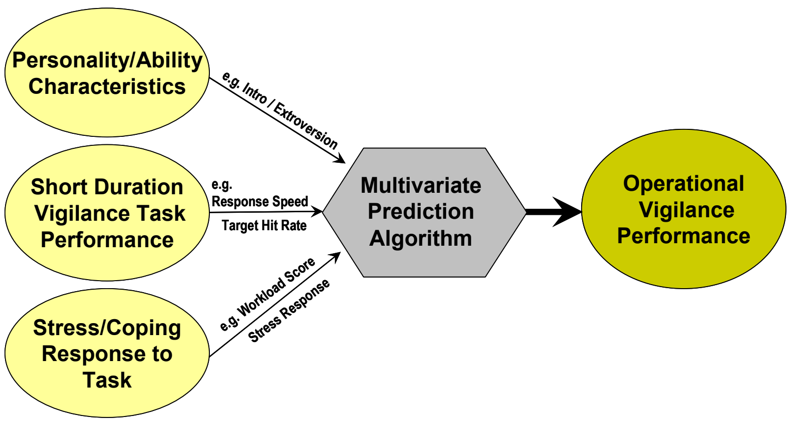

Figure 1. *Vigilant Warrior*™: A model approach to developing a personnel selection tool for sustained attention ability

The proposed model. The challenge presented for developing *Vigilant Warrior*™ was to apply the concepts of vigilance and its measurement discussed above to develop a reliable and valid vigilance prediction toolset. The multidimensional solution to vigilance prediction that was conceived to meet this challenge was to sample key constructs related to (1) personality and analytic skill, (2) objective task performance, and (3) stress, workload and coping responses to vigilance tasks. A primary goal was to extract the optimal measurement instruments from these complimentary approaches and blend them to produce an efficient personnel selection system capable of predicting vigilance performance. A graphic representation of the *Vigilant Warrior*™ personnel selection battery concept is shown in Figure 1.

Preliminary Research and Test Battery Selection

To identify preliminary components for each of the three vigilance prediction dimensions discussed above, we examined the literature addressing the relationship between various personality and analytic skill variables and vigilance performance and documented the limitations and strengths of identified vigilance predictors. Finally, a panel of experts rated the degree of research support and projected utility for each personality dimension. In addition, available brief vigilance tasks were assessed for inclusion in the battery, as well as subjective rating dimensions and scales that could be used to determine an examinee's perceived workload, coping responses, and attitudes associated with performing the vigilance task. Based on the results of these analyses, we developed a candidate vigilance prediction battery composed of personality/analytic skill metrics, brief vigilance-task performance metrics, and resource depletion and allocation metrics. The personality dimensions selected for preliminary research were: Introversion/Extraversion, Intelligence Quotient, Boredom Proneness, Cognitive Failures, Conscientiousness, Trait Sleepiness, Attention Deficit Hyperactivity Disorder (ADHD), Schizotypy, and Propensity to Daydream. Two measures of Analytic skill, Fluid and Crystallized Intelligence, rounded out this group of measures. Two versions of a Short Vigilance Task (SVT) were created for the battery in order to account for the well-known differences in performance and sensitivities to stimulus and environmental variables observed in tasks with (*simultaneous*) and without (*successive*) a comparison stimulus available to classify an event as a signal or a non-signal. The task is a brief (12-minute), paired-symbol vigilance task. Events are presentations of letter pairs in any combination drawn from the letters D, O, and backward D. In the *simultaneous* trials, the signal is any matching pair (e.g., "DD"). In the *successive* version, the signal is defined as the occurrence of the pair "OO." Finally, the Dundee Stress State Questionnaire (DSSQ), the Coping Inventory for Task Situations (CITS), the National Aeronautics and Space Administration Task Load Index (NASA TLX) workload scale, and the Boles Multiple Resource Questionnaire were selected to assess subject attitudes toward, and responses to, performing the SVT. Dimensions assessed by these instruments are Task Engagement, Distress, Worry, Coping (task focused), Coping (avoidance), Coping (emotion focused), Workload, and Multiple Resource Usage.

Refinement Of The Initial Battery

The goals of the main preliminary investigation of the candidate vigilance test battery was to confirm the qualities of the SVT, assess the psychometric properties of the personality, intelligence, and stress/attitude/coping measures to be included in the battery, and to assess their differential abilities to predict vigilance performance on the SVT. The study was conducted with a sample of 210 participants recruited from psychology classes at the University of Cincinnati.

Method. Participants completed a series of questionnaire and performance-based assessments in the following sequence: personality tests, intelligence tests; pre-task stress state, 12-minute SVT; and post-task stress state and coping. During the SVT the character pairs were presented against a masking background at a high event rate. One hundred five (105) participants performed the *simultaneous* version of the task, requiring a comparative judgment to detect the target, while 105 additional participants performed the *successive* version of the task, requiring an absolute judgment to detect the target.

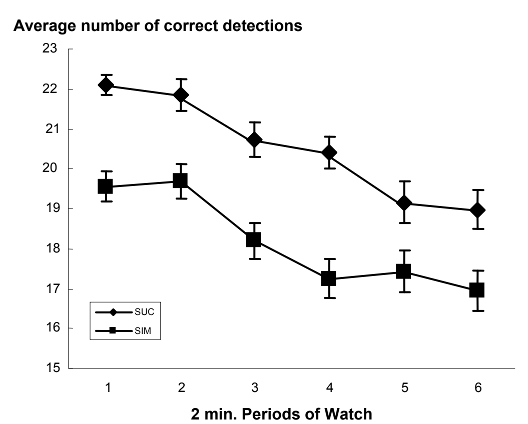

Figure 2. Mean number of correct detections as a function of periods of watch for both *simultaneous* (SIM) and *successive* (SUC) conditions.

Validity of the SVT. One objective of this study was to ensure that the SVT developed for the battery would show the classic performance changes over time that are characteristic of typical longer tasks. Figure 2 shows the average number of correct detections made by subjects performing the *successive* (SUC) and *simultaneous* (SIM) versions of the test over the six continuous 2-min. watch periods. As the graph suggests, the short tasks yielded a common decrement in performance over the 12-min. watch ($F_{(5, 1248)} = 44.74$, $p < .001$.) and a clear difference between the task conditions (F (1, 208) = 19.80, $p < .001$).

Factor analysis of the personality scales. A factor analysis was conducted to test whether the initial set of personality dimensions could be reduced to a smaller number of underlying factors. Analysis of the personality scales showed that these individual difference indicators were intercorrelated. A principal factor analysis was run, followed by an oblique (direct oblimin) rotation. On the basis of the scree test and factor interpretability, a four-factor solution was extracted, explaining 63.7% of the variance. Factor 1 (labeled Cognitive

Disorganization) is defined by various scales linked to disruption of attentional focus, including cognitive failures, mind wandering, and daydreaming, as well as the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE) Disorganization scale and the Young ADHD Questionnaire-Self-Report (YAQ-S). Factor 2 (Heightened Experience - i.e., enjoyment of events) is defined by O-LIFE unusual experiences and sensation-seeking subscales, and low internal boredom score of the Boredom Proneness Scale. This factor appears to indicate a vivid, excitable mental life. Factor 3 (Sleep quality) brings together the 3 subscales of the Pittsburgh Sleep Quality Index used in the study. Surprisingly, the Epworth Sleepiness Scale (Johns, 1994) fails to load on this factor. Factor 4 (Impulsivity) contrasts the sensation-seeking subscale of the Urgency, Premeditation, Perseverance, and Sensation Seeking (UPPS) scale with the low-premeditation subscale on the I_7 Impulsiveness Questionnaire. The factors were intercorrelated, with the highest correlations found between factors 1 and 4 (r = .51) and between 1 and 3 (r = .44). Factor 2 was largely uncorrelated with the remaining factors.

Correlates of SVT performance. Satisfied that the SVT possesses the fundamental characteristics of a more classical extended-duration task, we examined the Pearson correlations between the SVT and the personality and situational measures. Personality was represented by regression-model factor scores computed on the basis of the factor analysis. Detection frequencies within each 2-min. period were highly intercorrelated (alpha = .93), so average target detection frequency was used as the performance measure for this analysis. Table 1 provides a summary of the correlations of the various scales with performance, for *simultaneous* and *successive* conditions.

Table 1. *Correlations Of Intelligence And Stress Variables With Performance.*

Test Type	Test/Questionnaire	Simultaneous	Successive
Intelligence	Advanced Vocabulary	.084	.294**
	Letter Sets	.274**	.259**
Personality	Cognitive disorganization	-.099	-.089
	Impulsivity	-.170	-.132
	Heightened awareness	.090	.048
	Sleep quality	-.033	.077
Stress (pre)	Engagement	.359**	.122
	Distress	-.135	-.089
	Worry	-.156	-.152
Stress (post)	Engagement	.456**	.402**
	Distress	-.199*	-.180
	Worry	-.120	-.172
Coping	Task-focused	.284**	.402**
	Emotion Focused	-.230*	-.181
	Avoidance	-.429**	-.303**

Note: **Correlation is significant at the .01 level.
*Correlation is significant at the .05 level.

Table 1 shows that the two measures of Analytic skill positively correlate with performance on the SVT. The Educational Testing Service (ETS) Advanced Vocabulary test (Crystallized Intelligence) is a better predictor of performance on the *successive* task, while the ETS Letter Sets test (Fluid Intelligence) correlates with both the *simultaneous* and *successive* tasks. The other correlates with the SVT were the subjective stress states and coping-style measures. Table 1 also suggests that, while *simultaneous* and *successive* tasks have some common correlates, the set of correlates for each type of task may differ somewhat.

Candidate Test Battery for Validation

This preliminary study confirmed that the SVT showed the vigilance decrement characteristic of performance of longer monitoring tasks, qualifying it as the performance sampling component of the battery. The data also replicated findings that personality traits are no more than modest predictors of vigilance**.** However, additional analyses showed that some of the personality factors predict stress and coping during vigilance, which may contribute to their utility in prediction for a longer, sustained monitoring task. In addition, the present data support inclusion of short intelligence tests in the predictive battery. Thirdly, both stress states and coping scales correlated with performance, supporting inclusion of these measures in the battery. Finally, the analyses permitted reductions in both the number of tests and the number of test items in the battery. These reductions allowed construction of a 45-minute automated test battery to be used in the *Vigilant Warrior*™ battery validation study.

Criterion Validation Study

The vigilance criterion task designed to test the predictive capabilities of the *Vigilant Warrior*™ battery employed a simulated, tactical, situation display presented on a computer monitor to provide a two-dimensional plan-view map of a geographical area within which the positions of military combat vehicles were represented. Static components of the display included terrain features and reference grid lines. The dynamic components of the display were moving combat vehicles, the positions of which changed with each display update. The symbolic combat vehicles appeared in three columns that moved from left to right across the screen and returned in the opposite direction with unpredictable directional deviations. The center column of combat vehicles was led by a combat tank with two gun barrels. The display was updated every second, with the gun barrels displayed for 50 msec. Participants were required to report a detection whenever the gun barrels were of different lengths (*simultaneous* condition), or are both were longer than the standard length (*successive* condition). Two additional versions of the *successive* criterion task were created to examine the battery's capacity to predict performance under special task conditions and the concurrent cognitive demands that accompany many real-world vigilance tasks. The target cueing version was intended to simulate vigilance tasks augmented by probabilistic information about potential upcoming signals during screen display updates. The second version of the criterion task represented the common vigilance condition in which the worker is engaged in an additional task; in this case, a secondary auditory task to answer queries about the location of specific vehicles on the map. This additional task was designed to increase the mental resource demands imposed upon the subject to permit testing the ability of the battery to predict vigilance performance under multitasking conditions.

Criterion tasks. Task duration was 60 minutes in all cases, analyzed as 6 successive 10-min. periods of work. Correct detections and false positive responses were recorded for all task versions. The signal detection theory index of perceptual sensitivity, *d*' (Macmillan & Creelman. 2005), was calculated from these response data and was employed as the principal performance index in the validation study.

Participants and procedure. A total of 462 participants were recruited. They were allocated at random to the four criterion task conditions as follows: *Simultaneous* detection task (110), *Successive* detection task (122), *Successive* detection task *with cueing* (122), *Successive* detection task *with auditory competing task* (108). Participants first completed the automated *Vigilant Warrior*™ described above. Then, participants participated in two 2-min. practice sessions for the specific criterion task to be performed followed by the task itself for 60 minutes.

Results. Three sets of predictors were available from the tested battery of measures: (1) The dispositional measures (personality and analytic skill), (2) mean *d*' on the SVT, averaged across the six task periods (Cronbach α = 0.95), and (3) the subjective measures taken following the SVT including three stress state factors (Engagement, Distress, and Worry), three coping scales (Task-focused, Emotion-focused, and Avoidance), and overall workload from the modified NASA-TLX, calculated as an unweighted sum of the 6 rating scales. The performance criterion was mean *d*' on the criterion task, averaged across the six task periods (Cronbach α = 0.97) and was calculated separately for each of the four criterion task versions: *simultaneous*, *successive*, *successive* with cueing, *successive* with secondary task. Bivariate correlations showed that SVT *d*', Analytic skil, post-SVT subjective state, and coping all had some capacity to predict performance on the criterion task while the personality variables were unrelated. We then proceeded to a multiple regression analyis using Analytic skill, SVT *d'*, and the strongest stress/coping/workload measure, the task Engagement stress index.

Table 2. *Summary Statistics For The Regression Of Simultaneous Mean d' Onto The Predictor Sets*.

Step	Predictors	R^2	ΔR^2	df	F
1	Analytic Skill	.086	.086	2, 107	5.01**
2	SVT *d*'	.283	.198	1, 106	29.22**
3	Engagement	.326	.043	1, 105	6.64*

*p<.05, **p<.01

Table 2 shows the summary statistics for predicting criterion mean *d*' on the *Simultaneous* task. The two Analytic skill variables, SVT *d*', and post-SVT engagement all added to the variance explained, explaining about 33% of the variance in the criterion in total. The final equation attained significance (R = .571; $F_{(4,105)}$ = 12.87, p < .01).

Table 3. *Summary statistics for the regression of Successive Mean d' onto the predictor sets.*

Step	Predictors	R^2	ΔR^2	*df*	*F*
1	Task type	.104	.104	2,349	20.23**
2	Analytic Skill	.254	.150	2,347	18.05**
3	SVT *d*'	.374	.120	1,346	46.01**
4	Engagement	.385	.011	1,345	5.25*

* p < .05, **p < .01

Table 3 shows the summary statistics for the *Successive* criterion tasks. Again, all the predictor sets made a significant contribution, adding 27.1% to the variance explained by task type (the three different *Successive* task versions). The final equation attained significance (R = .620; $F_{(6,345)}$ = 35.95, p < .01).

Conclusions on Assessment of Individual Differences in Vigilance Ability using Vigilant Warrior™

This study validated the *Vigilant Warrior*™ battery against a specific criterion-task simulation in a laboratory setting. While further work will determine the generality of the results, the following conclusions are justified from the large body of data assembled thus far. The results clearly vindicate the multivariate approach to vigilance assessment upon which *Vigilant Warrior*™ was based. Use of multiple objective and questionnaire predictors in *Vigilant Warrior*™ enhances predictive validity. The results also show that the predictor sets are fairly consistent across different versions of the criterion task, implying that the battery has the capacity to predict performance across a range of sustained monitoring tasks and to be practically useful for selecting workers both for superior objective performance on sustained monitoring tasks and for greatest resistance to stress and fatigue.

Acknowledgements

This research was supported by a Phase II Small Business Innovation Research grant from the Army Research Institute under contract number W74V8H-06-C-0049. The views, opinions, and/or findings contained in this chapter are those of the authors and should not be construed as an official Department of the Army position, policy, or decision. The authors also thank Joel Warm, Gerald Matthews, Kyle Behymer, Benjamin Simpkins, Elizabeth Lerner, Victor Finomore, Tyler Shaw, and Mark Crabtree for their contributions to this research.

References

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Hancock, P.A, & Warm, J.S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, *31*, 519-537.

Helton, W.S., Dember, W.N., Warm, J.S., & Matthews, G. (1999). Optimism, pessimism, and false failure feedback: Effects on vigilance. *Current Psychology, 18*, 311-325.

Johnson, A. & Proctor, R.W. (2003). *Attention: Theory and practice.* Thousand Oaks, CA: Sage.

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

Macmillan, N.A. , & Creelman, C.D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Matthews, G. and Davies, D.R. (2001). Individual differences in energetic arousal and sustained attention: A dual-task study. *Personality and Individual Differences, 31,* 575-589.

Matthews, G., Davies, D.R., & Lees, J.L. (1990). Arousal, extraversion, and individual differences in resource availability. *Journal of Personality and Social Psychology, 59*, 150-168.

Robertson, I.H., Manly, T., Andrade, J. Baddeley, B.T., & Yiend, J. (1997). Oops! Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35,* 747-758.

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of caffeine and signal salience on performance, workload, and stress in an abbreviated vigilance task. *Human Factors*, 4*2,* 183-194.

Thackray, R.I., Bailey, J.P., & Touchstone, R.M. (1977). Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar control task. *Vigilance: Theory, operational performance, and physiological correlates*. R.R. Mackie (Ed.). New York, Plenum Publishing Company, pp. 203-215.

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). Mahwah, NJ: Erlbaum.