# AJP

Volume 124 • Number 4 • Winter 2011

# The American Journal of Psychology

Edited by Robert W. Proctor *Purdue University*

Book Reviews Dominic W. Massaro *University of California, Santa Cruz*

Obituaries and History of Psychology Alfred H. Fuchs *Bowdoin College*

# Workload and Stress in Vigilance: The Impact of Display Format and Task Type

JAMES L. SZALMA
University of Central Florida

Signal salience was manipulated using configural and object displays to examine their effects on the performance, workload, and stress of vigilance. Improving performance and reducing the workload and stress of vigilance are crucial concerns. Signal salience improves performance and reduces stress, but to date there have been no salience manipulations using configural displays in a vigilance task. Two task types (individual variable monitoring and midpoint identification) and 3 display formats (bar graph–different baselines, bar graph–common baseline, and a polygon graph display) were examined. Configural displays improved performance in the midpoint identification task but not in the individual variable monitoring task. Workload depended on the form of display features (bar graph vs. polygon). Stress increased across all conditions, but task and display format did not affect stress. The midpoint identification task was associated with more emotion-focused and avoidant coping. Increasing signal salience in a vigilance task using configural displays with emergent features or physical contours can improve performance and reduce the decrement. These displays may not reduce the stress of vigilance or encourage task-focused coping. Therefore, there may be hidden costs to vigilance performance even when highly salient configural displays are used.

Vigilance, or sustained attention, is the ability to attend to displays over prolonged periods on watch. This capability is relevant to many operational environments in which people monitor displays and take action only when deviations from normal state occur (e.g., Groemer et al., 2010; Hwang et al., 2008; Wiggins, 2011). The most ubiquitous finding in vigilance research is the decline in performance with time on watch (the vigilance decrement) (Davies & Parasuraman, 1982; See, Howe, Warm, & Dember, 1995). In addition, the psychophysical characteristics of vigilance tasks systematically affect performance (for a review, see Warm & Jerison, 1984). Sustained attention tasks also impose high mental workloads and are quite stressful (Warm, Matthews, & Finomore, 2008; Warm, Parasuraman, & Matthews, 2008). Furthermore, these effects can be brought under psychophysical control. Manipulations that improve performance also reduce the perceived workload and stress of vigilance. For instance, one of the most potent variables is signal salience (Warm & Jerison, 1984). Increasing signal salience improves vigilance performance and reduces the perceived workload and stress reported by operators engaged in sustained attention (Helton, Matthews, & Warm, 2009; Temple et al., 2000; Warm, Parasuraman, et al., 2008). However, in most studies examining these effects salience was manipulated by modulating simple psychophysi-

**American Journal of Psychology**
Winter 2011, Vol. 124, No. 4 pp. 441–454 • 

cal parameters such as luminance or target size. To date, no studies have manipulated signal salience using configural displays with features well mapped to system dynamics. One goal for the current investigation was to evaluate the effects of such a manipulation on the performance, perceived workload, and stress associated with vigilance.

### *Configural Displays*

Configural displays consist of elements organized so that relationships between system variables are represented by easily perceived features (for a review of this literature, see Bennett & Flach, 1992, 2011). These features may emerge from the arrangement of separate display elements, or they may be physical features (e.g., contours in an object display). In either case, changes in these features correspond to changes in the relationships between system variables. An important consideration for application of these display formats has been whether they could simultaneously support tasks that entail the extraction of low-level data (i.e., focused attention to specific variables, sometimes called focused attention tasks) and tasks that entail the synthesis of information from multiple display elements (i.e., information based on relationships between multiple variables, sometimes called information integration tasks; Bennett & Flach, 1992). These displays have been shown to improve performance in the latter (Peebles, 2008), but the effect of configural display format on performance of tasks that entail extraction of specific display elements has not been entirely consistent across studies. In their review of the literature, Bennett and Flach (1992) concluded that when differences between display formats are observed, they usually favor separable displays. However, they also noted that the evidence does not "strongly support the existence of an inherent and unavoidable cost for the extraction of low-level data with configural formats" (p. 528).

### *Using Configural Displays to Improve Vigilance*

The utility of configural displays in attenuating the vigilance decrement and mitigating workload and stress has not been examined empirically (but see Hancock & Szalma, 2003). In one noteworthy exception, Molloy and Parasuraman (1994) observed that an integrated display improved performance in an automation failure detection task, but the display had no significant effect on perceived workload. They attributed the absence of workload effects to high intersubject variability and to the insensitivity of the perceived workload measure they used (NASA Task Load Index). Note, however, that they did not compare different display formats or task types, nor did they examine the stress associated with using such displays for different kinds of task. Szalma, Hancock, Dember, and Warm (2006) evaluated stress response to a vigilance task with an object display. They reported that no performance decrement was associated with the object display, but the task was nevertheless perceived as stressful. However, Szalma et al. did not manipulate the display format or task type.

### *The Current Study*

The current experiment was conducted to investigate the effect of task type and display format on the performance, workload, and stress of vigilance. A second goal for this study was to measure the effect on vigilance of increasing signal salience using configural displays and determine whether this effect depends on the discrimination requirement of the task. If configural displays increase the salience of critical signals in a task requiring the identification of changes in the relationships between variables (e.g., an averaging or midpoint identification task), then it was expected that their use would reduce the vigilance decrement and perceived workload and stress reported by participants, relative to a display with features that did not provide an easily perceived configuration of elements. No such display differences were expected in the context of a task requiring attention to individual variables because the configuration of display elements was not relevant to the discrimination required. Alternatively, following Bennett and Flach (1992), if significant differences do occur, the displays in which the elements are separable (e.g., a configural display with emergent features) would be associated with better performance and lower workload and stress relative to displays in which the elements are not easily separable (e.g., an object display). In either case, a display format by task type interaction was expected. The workload and stress response associated with the different task–display combinations was expected to mirror those of the performance effects, such that conditions in which performance is improved should lead to lower workload and stress (i.e., performance–

workload associations; Hancock, 1996). With respect to coping response, it was predicted that if the use of highly salient emergent features in displays reduced stress and improved performance, they might also facilitate the adoption of task-focused coping strategies to deal with task demands rather than emotion-focused or avoidant coping methods. Although the latter two strategies might reduce some symptoms of stress, they would probably be associated with performance decrements. Alternatively, if the coping strategy observers use to deal with the stress of vigilance were similar across display formats, then configurality might be unrelated to coping strategy.

## EXPERIMENT

### METHOD

#### *Participants*

Ninety-six undergraduates (48 men and 48 women) at State University of New York, Farmingdale, participated in the study in exchange for course credit. They ranged in age from 18 to 46 years old ($M$ = 20.8). All participants had normal or corrected-to-normal vision and were free of known hearing impairments.

#### *Experimental Design*

Two levels of task type (individual variable monitoring and midpoint identification or data averaging) were factorially combined with three display types (bar graph with different baselines, bar graph with a common baseline and easily perceived changes in emergent features, and a polygon graph display with an easily perceived contour changes) to yield six experimental groups. Sixteen participants were assigned at random to each of the six conditions, with the restriction that the groups were equated for participant sex.

#### *Displays and Tasks*

The displays used are shown in Figure 1. The stimuli consisted of three black dots and either three rectangles (bar graph displays) or a polygon (object display). Each black dot was positioned above a rectangle or the polygon. Both tasks required participants to monitor the state of an imaginary system consisting of two inputs and an output. In each display the output value was located between the two inputs. The two bar graph displays differed only in the position of the output along the $y$ coordinate. The three rectangles of the bar graph with easily perceived emergent features were arrayed along a common baseline, whereas in the bar graph–different baselines display the output rectangle was placed 3.1 cm above the common baseline. The latter display was adapted from previous research in which the effectiveness of the configurality was disrupted by moving the baseline of the center rectangle (e.g., Carswell & Wickens, 1987). The purpose of the display manipulation was not to test differences in configurality per se but to use the features of these displays as a manipulation of signal salience in a vigilance task. Hence, the focal interest was in how the configurations and task demand combinations affect sustained attention and operator workload and stress. Pilot work established that the two tasks were psychophysically equated for difficulty. In that study, a sample of 16 participants discriminated critical from neutral events for each task and display format using a two-alternative forced-choice procedure.

In the midpoint identification task, input and output values in the bar graph displays were represented by the heights of the rectangles. Neutral events, in

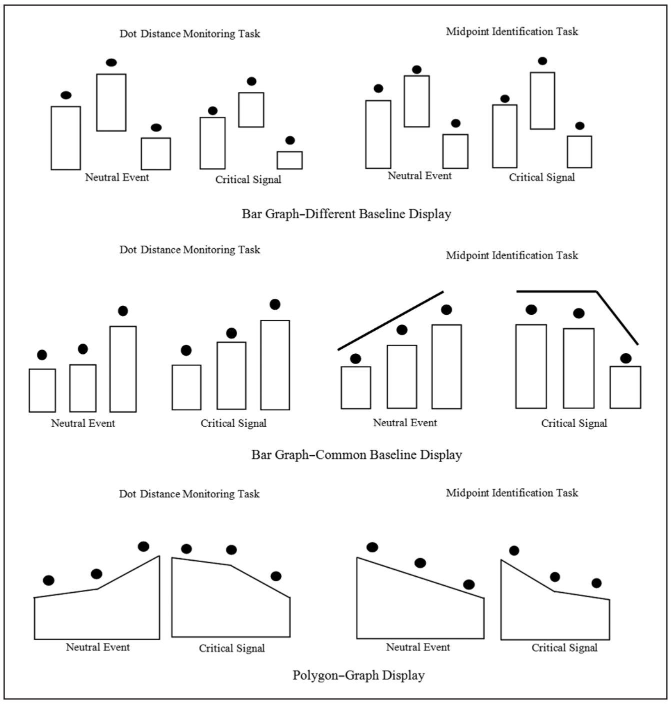


**FIGURE 1.** Displays used in the present experiment. In each case the output is in the middle, flanked by the two inputs. The lines above the rectangles in the bar graph–common baseline display indicate the emergent feature. The labels in each display and the lines above the rectangles are for illustrative purposes and were not present during the task

which the system parameters were defined as "in balance" and required no overt response from the participant, were cases in which the height of the output bar was the average of the heights of the two input bars. Neutral events for the rectangle displays had two different sets of values: (1) Input 1 and input 2 lengths were 5.8 cm and 2.8 cm, respectively, and the output length was 4.3 cm; and (2) input 1 and input 2 lengths were 4.3 cm and 1.4 cm, respectively, and the output value was 2.8 cm. Each rectangle was 1.9 cm in width. Critical signals for detection, in which the system was defined as "off balance," were cases in which the value of the output was 1 cm higher or lower than the average of the two input values. The lengths of the inputs were identical to those of the neutral events. The polygon was designed so that the input and output values, represented as the distance from the base to the top of the polygon at the point below each dot, were equivalent to those of the rectangle displays. The width of the polygon was 7.7 cm.

In the bar graph–common baseline display, the emergent feature was the inferred linearity of the input and output bars, such that the tops of the bars were arrayed along a straight line when the system was in balance (see Figure 1). Violations of the inferred linearity indicated that the system was off balance. Linearity was also a feature in the polygon graph display, a simplified adaptation of the "house" display used by Buttigieg and Sanderson (1991), but in this case the feature was a physical contour. Neutral events were cases in which the top of the polygon was a straight line. a bend or break in the top of the polygon indicated that the system was off balance.

In the task requiring monitoring specific variables (i.e., dot figure distance monitoring), the input and output values were represented as the vertical distance of each 0.6-cm diameter black dot from its respective rectangle or from the polygon, and the task required participants to monitor these relative distances. Neutral events (the system was "in balance") were defined as cases in which each dot was a standard distance (0.7 cm) from its rectangle or polygon. Critical signals for detection were cases in which any one of the dots was 0.5 cm closer than the standard distance to the rectangle or polygon. There were no cases in which more than one dot was closer than the standard distance. In addition, critical signals in this task were unrelated to the relative heights of the rectangles or the shape of the polygon, so that the linearity feature could not serve as a cue in the distance monitoring task.

In all conditions, stimuli were presented at a rate of 26 events/min by setting the stimulus onset asynchrony at 2.3 s. Twelve critical signals appeared at random intervals during each of the four 6-min periods of watch (signal probability = .08). Participants responded to critical signals by pressing the spacebar on a computer keyboard. Responses occurring within 1.5 s after the onset of a critical signal were recorded as correct detections. All other responses were recorded as false alarms. Participants were seated in front of a 43.2-cm (17-in) video display terminal at eye level approximately 45 cm from the participant.

### *Measurement of Perceived Workload, Stress, and Coping Strategies*

Perceived workload was assessed using the NASA Task Load Index (TLX; Hart & Staveland, 1988), which provides an index of global workload and identifies the relative contributions of six sources of workload (mental demand, physical demand, temporal demand, performance, effort, and frustration). Self-reports of stress were assessed using the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 2002), which provides 11 factor analytically determined scales to reflect multiple dimensions of cognitive state associated with task performance. These 11 scales constitute three secondary factors: task engagement, reflected by the Energetic Arousal, Concentration, and two motivation scales (Success and Intrinsic); distress, reflected by Tense Arousal, Hedonic Tone, and Control and Confidence; and worry, comprising Self-Focused Attention, Self-Esteem, Task-Related Cognitive Interference (thoughts concerning task performance), and Task-Irrelevant Cognitive Interference (thoughts unrelated to task performance). Coping strategies were assessed using the Coping Inventory for Task Stressors (CITS), a situational coping inventory designed to assess strategies for dealing with task-based stress (Matthews & Campbell, 1998). Specifically, the CITS identifies the degree to which a person engages in task-focused, emotion-focused, and avoidant coping.

### *Procedure*

Upon entering the laboratory, participants surrendered their wristwatches, pagers, and cell phones. Participants then received instructions and were provided with examples of neutral events and critical signals. Participants in the midpoint identification task were informed of the system dynamics (i.e., that under normal conditions the output was the average of the two inputs) but were not informed of the display features that could be used to aid performance on that task. However, most participants in the groups using the bar graph–common baseline and the polygon graph displays in the midpoint identification task indicated during debriefing that they used these features to discriminate signals from nonsignals. Participants in the dot–figure distance discrimination conditions were informed that the heights of the rectangles or shape of the polygon would change during the session but that they should ignore these changes and attend only to the distance of the dots from the rectangles or polygon. Similarly, participants assigned to the midpoint identification task were informed that occasionally one of the dots would move closer to the rectangle or polygon but that they should ignore these changes and attend only to the heights of the rectangles or polygon.

Participants completed the pre-version of the DSSQ after the task instructions and then performed a 12-min practice vigil divided into two continuous 6-min periods on watch. Participants received auditory feedback in the form of knowledge of results (KR) regarding the quality of their performance, in which a female voice controlled by the computer announced "correct" to indicate correct detections, "false alarm" to indicate errors of commission, and "miss" to indicate errors of commission. KR was not provided during the main vigil. The practice session was followed immediately by a 24-min vigil divided into four continuous 6-min periods on watch. After the vigil, participants completed a computerized version of the TLX, a posttask version of the DSSQ, and the CITS questionnaire. The order in which these measures were administered was counterbalanced.

To be retained in the study, participants had to detect 70% of critical signals and commit no more than 10% false alarms during the first period on watch. Twenty-seven participants did not meet these criteria and were replaced. Of those participants, 16 were in the bar graph–different baselines condition. *T* tests were computed to ensure that individuals in that group did not differ significantly in workload, stress, or coping strategy from those who did not meet the criteria. There were no statistically significant differences in these dependent measures between participants who did not meet the criteria and those who did. Thus, for the purposes of this study, the two groups differed significantly only in their ability to learn the task sufficiently to meet the performance criteria. Note that omitting the poor-performing participants from the control group actually *reduced* the probability of observing the hypothesized effects.

## RESULTS

### *Performance*

For each participant, proportions of correct detections and false alarms during the vigil were used to compute signal detection theory measures of $A'$ and $\beta_D''$. $A'$ is a nonparametric area measure of perceptual sensitivity that ranges from 0.5 to 1 and has been recommended for use in vigilance research (Craig, 1979). The nonparametric response bias index, $\beta_D''$, ranges from −1 to +1, with higher values indicating greater conservatism and zero indicating unbiased responding (Donaldson, 1992). See, Warm, Dember, and Howe (1997) established that among the nonparametric measures of response bias $\beta_D''$ was the most accurate for vigilance. The means and standard deviations for all conditions and dependent variables are in Tables 1–4. These data were analyzed via a 2 (task) by 3 (display) by 4 (period) mixed ANOVA with repeated measures on the last factor. For all analyses involving repeated measures, Box's epsilon was used to compute degrees of freedom to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004). The Bonferroni correction, with $\alpha = .05$, was used for each post hoc comparison. Mean $A'$ scores for each display are plotted as a function of period on watch in Figure 2a (midpoint identification task) and 2b (dot distance monitoring).

Statistically significant effects were observed for period on watch, $F(2, 206) = 18.05$, $p < .001$, $\omega^2 = .12$, task by display, $F(2, 90) = 4.71$, $p = .011$, $\omega^2 = .02$, and the three-way interaction between task, display, and period on watch, $F(5, 206) = 5.35$, $p < .001$, $\omega^2 = .01$. Separate display by period ANOVAs within each task revealed that in the midpoint identification task there were significant main effects for period, $F(2, 107) = 7.00$, $p = .001$, $\omega^2 = .06$, and for display, $F(2, 45) = 5.52$, $p = .007$, $\omega^2 = .09$. There was also a significant display × period interaction, $F(5, 107) = 4.08$, $p = .002$, $\omega^2 = .06$. Following recommendations by Maxwell and Delaney (2004), tests for the simple effects of display within each period were computed using separate error terms for each watch period rather than the pooled error term due to violation

**TABLE 1.** Means (*SDs*) for perceptual sensitivity and response bias scores ($N = 96$)

**Perceptual sensitivity (A')**

| Task/display condition | Period 1 | Period 2 | Period 3 | Period 4 | Overall |
|---|---|---|---|---|---|
| **Dot figure distance task** | | | | | |
| Bar graph–different baselines | 0.964 (0.025) | 0.958 (0.028) | 0.958 (0.028) | 0.950 (0.046) | 0.958 (0.033) |
| Bar graph–common baseline | 0.971 (0.014) | 0.958 (0.036) | 0.948 (0.036) | 0.940 (0.046) | 0.955 (0.037) |
| Polygon graph | 0.974 (0.015) | 0.962 (0.025) | 0.954 (0.025) | 0.915 (0.053) | 0.953 (0.039) |
| **Midpoint identification task** | | | | | |
| Bar graph–different baselines | 0.974 (0.018) | 0.957 (0.030) | 0.926 (0.052) | 0.913 (0.068) | 0.943 (0.052) |
| Bar graph–common baseline | 0.972 (0.025) | 0.971 (0.026) | 0.978 (0.019) | 0.964 (0.052) | 0.972 (0.033) |
| Polygon graph | 0.981 (0.016) | 0.984 (0.033) | 0.966 (0.051) | 0.972 (0.040) | 0.974 (0.037) |
| Overall | 0.973 (0.020) | 0.963 (0.030) | 0.955 (0.040) | 0.942 (0.055) | 0.959 (0.040) |

**Response bias ($\beta_D$")**

| Task/display condition | Period 1 | Period 2 | Period 3 | Period 4 | Overall |
|---|---|---|---|---|---|
| **Dot figure distance task** | | | | | |
| Bar graph–different baselines | 0.751 (0.223) | 0.842 (0.141) | 0.830 (0.232) | 0.895 (0.099) | 0.830 (0.186) |
| Bar graph–common baseline | 0.771 (0.196) | 0.860 (0.155) | 0.895 (0.129) | 0.861 (0.233) | 0.847 (0.185) |
| Polygon graph | 0.734 (0.367) | 0.851 (0.201) | 0.864 (0.254) | 0.950 (0.058) | 0.851 (0.253) |
| **Midpoint identification task** | | | | | |
| Bar graph–different baselines | 0.553 (0.394) | 0.775 (0.276) | 0.894 (0.153) | 0.880 (0.129) | 0.776 (0.289) |
| Bar graph–common baseline | 0.803 (0.200) | 0.860 (0.166) | 0.846 (0.091) | 0.833 (0.160) | 0.836 (0.157) |
| Polygon graph | 0.825 (0.145) | 0.864 (0.071) | 0.855 (0.080) | 0.807 (0.180) | 0.838 (0.126) |
| Overall | 0.740 (0.278) | 0.842 (0.177) | 0.864 (0.167) | 0.871 (0.157) | 0.830 (0.207) |

**TABLE 2.** Means (*SDs*) for NASA-TLX Scores ($N = 96$)

**NASA-TLX scale**

| Task/display condition | GWL | MD | TD | PW | E | F |
|---|---|---|---|---|---|---|
| **Dot figure distance task** | | | | | | |
| Bar graph–different baselines | 65.20 (14.11) | 247.25 (133.95) | 185.31 (119.80) | 65.69 (47.48) | 245.31 (106.87) | 178.75 (123.69) |
| Bar graph–common baseline | 63.93x (18.16) | 274.69 (119.27) | 238.44 (126.88) | 85.50 (74.96) | 161.44 (119.03) | 187.31 (149.34) |
| Polygon graph | 54.52 (20.28) | 177.81 (143.69) | 231.56 (142.85) | 93.75 (57.43) | 156.88 (123.92) | 140.00 (135.55) |
| **Midpoint identification task** | | | | | | |
| Bar graph–different baselines | 62.52 (20.73) | 290.00 (133.69) | 202.50 (151.25) | 127.19 (120.19) | 133.44 (97.96) | 132.81 (118.01) |
| Bar graph–common baseline | 49.08 (19.76) | 165.31 (101.13) | 152.50 (111.46) | 89.69 (91.13) | 136.88 (104.29) | 146.25 (136.83) |
| Polygon graph | 48.62 (21.91) | 167.81 (163.56) | 142.13 (119.40) | 54.69 (54.33) | 128.06 (104.38) | 210.31 (182.02) |

Note. *E = weighted effort; F = weighted frustration; GWL = global workload; MD = weighted mental demand; PW = weighted performance workload; TD = weighted temporal demand; TLX = Task Load Index.*

**TABLE 3.** Means (*SD*s) for DSSQ Pre–Post Task Change Scores (*N* = 96)

| | DSSQ scale | | |
|---|---|---|---|
| Task/display condition | Task engagement | Distress | Worry |
| **Dot figure distance task** | | | |
| Bar graph–different baselines | –0.46 (0.73) | 1.13 (1.15) | –0.61 (0.81) |
| Bar graph–common baseline | –0.74 (0.73) | 1.13 (1.40) | –0.37 (0.91) |
| Polygon graph | –0.92 (1.01) | 0.66 (0.87) | –0.47 (0.78) |
| **Midpoint identification task** | | | |
| Bar graph–different baselines | –1.02 (0.65) | 1.18 (0.88) | –0.37 (0.76) |
| Bar graph–common baseline | –0.80 (0.60) | 0.51 (1.26) | –0.34 (0.99) |
| Polygon graph | –0.79 (0.65) | 0.59 (0.92) | –0.37 (0.76) |

Note. *DSSQ = Dundee Stress State Questionnaire.*

**TABLE 4.** Means (*SD*s) for Coping Strategy Scores (*N* = 96)

| | Coping scale | | |
|---|---|---|---|
| Task/display condition | TC | EC | AC |
| **Dot figure distance task** | | | |
| Bar graph–different baselines | 16.81 (5.32) | 6.81 (4.20) | 5.00 (4.72) |
| Bar graph–common baseline | 16.06 (5.81) | 5.78 (3.19) | 6.06 (6.04) |
| Polygon graph | 15.81 (5.02) | 4.06 (3.07) | 5.06 (3.77) |
| **Midpoint identification task** | | | |
| Bar graph–different baselines | 16.63 (5.17) | 7.56 (5.73) | 7.63 (5.29) |
| Bar graph–common baseline | 17.19 (4.51) | 8.88 (4.76) | 7.94 (3.96) |
| Polygon graph | 16.81 (4.28) | 6.31 (6.02) | 7.63 (6.97) |

Note. *AC = avoidant coping; EC = emotion-focused coping; TC = task-focused coping.*

of sphericity, $\chi^2(5) = 18.67$, $p = .002$. Significant differences between displays were observed only in period 3, $F(2, 45) = 6.34$, $p = .004$, $\omega^2 = .18$, and in period 4, $F(2, 45) = 5.48$, $p = .007$, $\omega^2 = .16$ Post hoc tests for period 3 indicated that relative to the bar graph–different baselines group, participants in the bar graph–common baseline condition (Cohen's $d = 1.34$) and the polygon graph condition ($d = 0.78$) achieved significantly higher A′ scores. A similar pattern was observed for period 4. Relative to the bar graph–different baselines condition, participants in the bar graph–common baseline ($d = 0.84$) and polygon graph condition ($d = 1.06$) achieved significantly higher sensitivity scores during that period on watch. Sensitivity scores for the two latter conditions did not differ significantly from one another in either period of watch ($p > .40$ in each case).

In the dot distance discrimination task, sensitivity declined significantly over time, $F(2, 96) = 11.81$, $p < .001$, $\omega^2 = .14$ (Figure 2b), but the three display types did not differ significantly from one another in sensitivity ($p = .75$). There was a trend toward a greater decrement in the polygon graph display condition than in the other two groups, but the

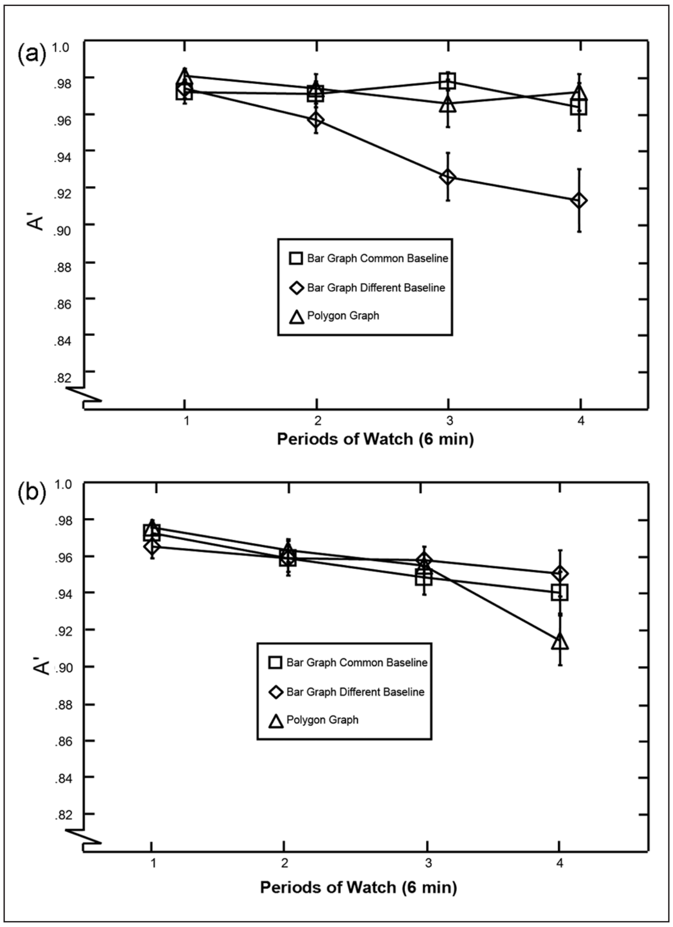


**FIGURE 2.** Sensitivity as a function of periods of watch for the three display formats (a) in the midpoint identification averaging task and (b) in the dot–figure distance monitoring task. Error bars are standard errors

period × display interaction was not statistically significant when the correction for sphericity was applied, $F(4, 96) = 2.12$, $p = .079$, $\omega^2 = .03$. Inspection of Figure 2b indicates that this trend probably resulted from sensitivity differences between display format conditions during the final period on watch. The pairwise effect sizes for that period for the polygon graph display and the bar graph displays were $d = 0.50$ for the bar graph–common baseline condition and $d = 0.70$ for the bar graph–different baselines condition. These represent medium and large effect sizes, respectively (Cohen, 1988).

Response bias scores for participants in each display condition are plotted as a function of period of watch in Figure 3a and Figure 3b for the midpoint identification and dot distance monitoring tasks, respectively. An ANOVA indicated a main effect for period, such that conservatism increased with time on watch, $F(2, 198) = 14.77$, $p < .001$, $\omega^2 = .10$, and a significant three-way interaction between task, display, and period on watch, $F(4, 198) = 3.44$, $p = .008$, $\omega^2 = .006$. All other sources of variance failed to reach statistical significance ($p > .05$ in each case). Separate display by period ANOVAs within each task indicated a significant effect for period on watch in the midpoint identification task, $F(2, 76) = 6.44$, $p = .004$, $\omega^2 = .08$, and a significant display by period interaction, $F(3, 76) = 4.34$, $p = .005$, $\omega^2 = .09$. The main effect for display was not significant ($p = .29$). As in the case for sensitivity, tests for the simple effects of display within each period were computed using separate error terms due to violation of sphericity, $\chi^2(5) = 45.51$, $p < .001$. Statistically significant differences between displays were observed only for the first period, $F(2, 45) = 5.10$, $p = .01$, $\omega^2 = .14$. Post hoc tests indicated that relative to the control condition, participants in the bar graph–common baseline ($d = 0.80$) and polygon graph ($d = 0.92$) conditions were significantly more conservative in responding. Response bias scores for the two latter conditions did not differ significantly from one another ($p > .05$; $d = .12$). In the distance monitoring task the only statistically significant effect was for period, such that participants became more conservative over time, $F(2, 108) = 9.49$, $p < .001$, $\omega^2 = .12$.

### *Workload*

#### *GLOBAL WORKLOAD*

A 2 (task) × 3 (display) ANOVA on global workload scores revealed a statistically significant effect for task type, $F(1, 90) = 3.91$, $p = .051$, $\omega^2 = .03$; the midpoint identification task imposed less global workload than the distance monitoring task ($d = 0.39$; see Table 2). There was also a main effect for display, $F(2, 90) = 3.27$, $p = .043$, $\omega^2 = .04$. The interaction between task and display format was not significant ($p = .43$). Post hoc comparisons revealed that the polygon graph display imposed significantly less workload than the bar graph–different baselines display ($d = 0.63$). The global workload associated with the bar graph–common baseline display was not significantly different from the scores associated with the bar graph–different baselines group ($p = .13$;

$d$ = 0.39) or from scores associated with the polygon graph condition ($p$ > .31; $d$ = 0.24).

In addition to global workload, ratings on the weighted subscales of the TLX were also collected (see Table 2). An ANOVA revealed significant effects for scales, $F(3, 304) = 16.83$, $p < .001$, $\omega^2 = .12$, task, $F(1, 90) = 5.14$, $p = .026$, $\omega^2 = .01$, and a significant task × display × scales interaction, $F(7, 304) = 2.28$, $p = .030$, $\omega^2 = .03$. The three-way interaction was explored via separate task by display ANOVAs for each scale.

*MENTAL AND TEMPORAL DEMAND*

For the Mental Demand scale the only significant effect was for display, $F(2, 90) = 4.09$, $p = .020$, $\omega^2 = .06$. Post hoc tests revealed that the polygon graph display ($M$ = 172.8) imposed significantly less mental demand than the control group ($M$ = 268.6; $d$ = 0.67) The mental demand associated with the bar graph–common baseline display ($M$ = 220.0) did not differ significantly from that associated with either the bar graph–different baseline ($d$ = 0.38) or the polygon graph display (Cohen's $d$ = 0.34). Note that although these differences were not statistically significant, the effect size magnitudes are in the small to medium range (Cohen, 1988). For the Temporal Demand scale the only significant effect was for task, $F(1, 90) = 3.99$, $p = .049$, $\omega^2 = .03$, with the midpoint identification task ($M$ = 165.7) inducing lower temporal demand than the distance monitoring task ($M$ = 218.4; $d$ = 0.41).

*PERCEIVED PERFORMANCE*

For the Performance subscale there were no significant main effects ($p$ > .52 in each case), but there was a significant task × display interaction, $F(2, 90) = 3.31$, $p = .041$, $\omega^2 = .05$. Tests for the simple effects of display within each task revealed a marginally significant display effect in the midpoint identification task, $F(2, 45) = 2.46$, $p = .097$, $\omega^2 = .06$. Post hoc tests indicated a marginally significant ($p$ = .079) difference between the control group ($M$ = 127.2) and the polygon graph condition ($M$ = 54.7; $d$ = 0.78). Note that although this effect was statistically nonsignificant, the effect size is large (Cohen, 1988). In contrast, the effect size for the bar graph common baseline display relative to the control group was in the small to medium range, $d$ = 0.35, and the effect size for the two displays with easily perceivable features was in the medium range, $d$ = 0.47. In the distance monitoring task there were no significant differences between displays ($p$ = .42).

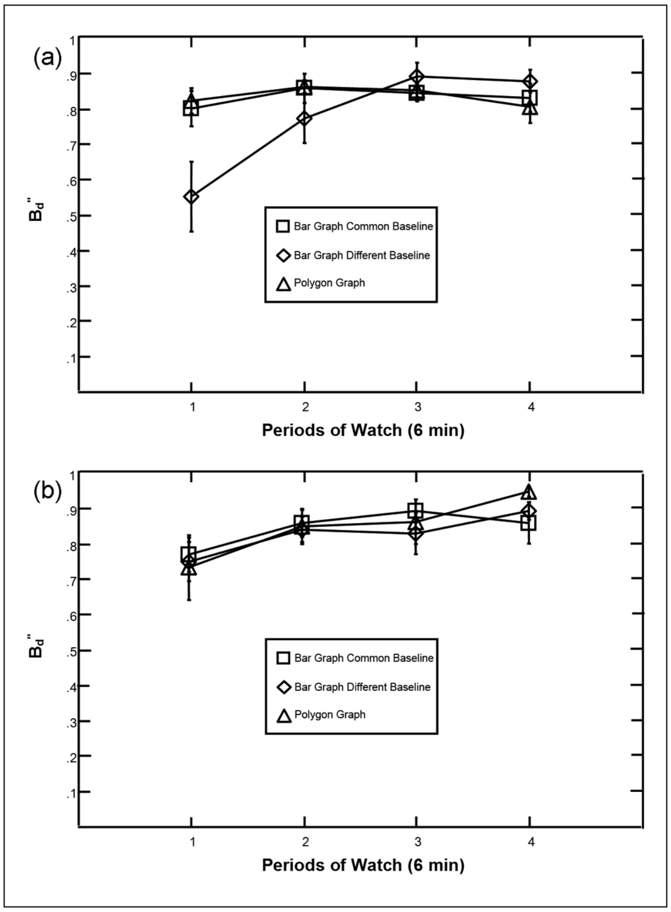


**FIGURE 3.** Response bias as a function of periods of watch for the three display formats (a) in the midpoint identification averaging task and (b) in the dot–figure distance monitoring task. Error bars are standard errors

*EFFORT AND FRUSTRATION*

On the Effort scale the only significant effect was for task, $F(1, 90) = 6.04$, $p = .02$, $\omega^2 = .05$. Participants who performed the midpoint identification task ($M$ = 132.8) reported less effort than their cohorts in the distance monitoring groups ($M$ = 187.9; Cohen's $d$ = 0.49). No statistically significant effects were observed for the Frustration subscale ($p$ > .18 in each case).

*Stress*

Participants' responses on the pre-DSSQ and post-DSSQ were used to calculate scores for each of the secondary factors, distress, task engagement, and worry (Matthews et al., 2002; see Table 3). A 2 (task) by 3 (display) ANOVA was computed for pre–post change scores for each secondary factor. No statistically significant effects were observed for the three dimensions of stress ($p > .15$ in each case). *T* tests were computed in order to determine whether there were global pre–post vigil changes in stress state and to verify that the absence of significant effects related to display or task type was not due to insensitivity of the DSSQ to the demands of the vigil per se. Scores on all three scales were significantly different from zero. Thus, across all experimental conditions there was a pre–post increase in distress ($M = .86$), $t(95) = 7.65$, $p < .001$, and declines in task engagement ($M = -.79$), $t(95) = -10.42$, $p < .001$, and worry ($M = -.42$), $t(95) = -5.03$, $p < .001$.

### *Coping Strategy*

Scores from the CITS were used to calculate the degree to which participants reported that they adopted three coping strategies: task-focused, emotion-focused, and avoidant coping (see Table 4). A 2 (task) × 3 (display) ANOVA was computed for each coping strategy. No statistically significant effects were observed for task-focused coping ($p > .58$ in each case). A significant effect for task was observed for the emotion-focused scale, $F(1, 90) = 4.61$, $p = .035$, $\omega^2 = .04$, and for avoidant coping, $F(1, 90) = 4.83$, $p = .030$, $\omega^2 = .04$. Participants who performed the midpoint identification task reported using more emotion-focused ($M = 7.58$, $SD = 5.51$) and avoidant ($M = 7.73$, SD = 5.42) coping strategies than those in the distance monitoring task (emotion coping: $M = 5.55$, $SD = 3.63$; $d = 0.44$ avoidant coping: $M = 5.38$, $SD = 4.85$, $d = 0.46$). All other effects failed to reach statistical significance ($p > .12$ in each case).

## DISCUSSION

The purpose of this investigation was to examine the effect of using configural displays as a manipulation of signal salience on the performance, workload, and stress of vigilance. The results indicate that displays with easily perceived configural features can reduce the vigilance decrement in tasks that require midpoint identification, tasks for which these displays have been demonstrated to have their most beneficial effect (Bennett & Flach, 1992). In the dot distance monitoring task the three display formats yielded statistically equivalent performance, although there was a decline in perceptual sensitivity during the final period on watch for participants in the polygon graph condition. Recall that the effect sizes associated with the final period on watch were $d = 0.70$ for the bar graph–different baseline and $d = 0.50$ for the bar graph–common baseline condition. These effects correspond to a decline in performance of 26% and 19%, respectively. Nevertheless, differences between displays were substantially smaller in the dot distance monitoring task than in the midpoint identification task, suggesting that the facilitative effects of configural displays containing easily perceivable features on vigilance are task dependent and occur only when they increase signal salience.

Display differences in response bias scores also depended on task type. For a task requiring midpoint identification, the effect of display varied as a function of period on watch. Earlier in the watch conservatism was higher in the bar graph–common baseline and polygon graph display conditions than in the bar graph–different baselines display condition, but during the last three periods participants in the control group achieved levels of conservatism similar to those of the other two groups. The configural displays may have facilitated stable responding earlier in the vigil because the features increase awareness of the low signal frequency, permitting participants to adjust their responding accordingly (Craig, 1978). That is, the easier discrimination associated with the configural displays imposed more rapid criterion adjustment because participants may have become aware of the low critical signal probability early in the first period. In contrast, participants in the dot distance discrimination conditions did not have the benefit of increased signal salience, and so their awareness of the relative infrequency of signals was slower to emerge.

### *Workload*

Perceived workload effects were influenced by display format and by task type, but, contrary to prediction, these two factors did not interact. The polygon graph display reduced the global workload associated with vigilance for both tasks. The relative simplicity of this display may have contributed to these effects, as it was the Mental Demand subscale that was most strongly associated with reduced workload. In regard to task effects, the midpoint identification task imposed less global workload than the dot–figure distance monitoring task. Analysis of the subscales revealed that temporal demand was the major contributor to task differences. It may be that the requirement to scan three different elements in the display induced time pressure in participants performing the distance monitoring task, whereas those in the other task may have been too preoccupied with the necessary mental calculations (despite the presence of easily perceived emergent features) for midpoint identification to attend to the temporal component of the task. That is, differences in the distortion of time perception under stress may have contributed to task differences in both perceived workload and stress (see Hancock & Weaver, 2005). Such an interpretation must be tentative, however, and await further studies that systematically manipulate the spatial and temporal properties of the tasks (see Hancock & Warm, 1989; Teo & Szalma, 2010).

*Performance–Workload Associations and Dissociations*

The results for perceptual sensitivity and perceived workload indicate a complex pattern of associations and dissociations between workload and performance, depending on task and display type. Associations are cases in which increases in task load (and performance impairment) and increases in workload co-occur, whereas dissociations are cases in which decreases in task load (and performance improvement) are linked with increases in workload (or vice versa; see Hancock, 1996; Parasuraman & Hancock, 2001; Oron-Gilad, Szalma, Stafford, & Hancock, 2008). In this experiment, the performance–workload links across display formats were complex and depended on task type. The polygon graph display reduced workload for both tasks but improved perceptual sensitivity only in the midpoint identification task, and it impaired performance during the final period on watch of the dot distance monitoring task. This pattern indicates an association between workload and performance for that display in the midpoint identification task but dissociation between workload and performance in the distance monitoring task. The bar graph common baseline display showed a different pattern: Workload was not reduced in either task, but perceptual sensitivity was increased in the midpoint identification task. This represents a workload insensitivity with respect to the midpoint identification task (workload was not affected but performance was improved; Hancock, 1996) and association with respect to the distance monitoring task.

During the final period, however, performance in the polygon graph condition declined relative to those of the other two conditions. The lower perceived workload reported by participants in that condition is therefore more striking because the NASA-TLX relies on participants' recall of their experience of the task, and in that condition the most recent memory would be of the final period on watch, in which performance declined. Declines in both workload and performance (dissociation) are diagnostic of withdrawal of effort from a task (Hancock, 1996). This interpretation is supported by the finding that participants in the polygon graph conditions reported less mental demand and effort than their cohorts who monitored the other displays.

In the midpoint identification task the workload insensitivity of the bar graph–common baseline display (workload was not significantly different from the control group but performance was better) may result from similar working memory demands imposed by the two bar graph displays (Yeh & Wickens, 1988). The emergent feature of the bar graph–common baseline display may result in better performance than a bar graph display with different baselines, but it does not relieve the working memory demands associated with a midpoint identification task. However, note that although the difference in workload between the two bar graph conditions was not statistically significant, the effect size was in the medium range ($d = 0.39$). The workload insensitivity for the bar graph–common baseline condition may therefore be a statistical artifact. Nevertheless, relative to the bar graph different baselines display, the effect sizes for the bar graph–common baseline ($d = 0.39$) and polygon display ($d = 0.63$) were substantially different from one another, suggesting that these two displays are not equivalent in their impact on workload. One possibility is that scanning the elements of the bar graph displays required more effort as a result of greater spatial separation, whereas scanning the polygon display required less effort because the elements were closer together (cf. Buttigieg & Sanderson, 1991). It is possible that the scanning imperative in the bar graph displays required more effort than was needed to scan the single polygon but that the emergent feature in the bar graph–common baseline display was sufficiently salient to improve performance.

### *Stress*

The stress of sustained attention was not influenced by either task or display type, although overall pre–post changes on the scales of the DSSQ indicated that the vigil was stressful. The latter finding indicates that the absence of significant task or display differences is not likely to be a floor effect; rather, although configural displays with salient features improve vigilance and, in the case of the polygon display, reduce perceived workload, they do not necessarily reduce the subjective costs of performance maintenance (cf. Hockey, 1997; see also Helton & Warm, 2008). However, previous research found that increasing signal salience decreased self-reports of stress in vigilance tasks (Helton et al., 2009; Temple et al., 2000). The pattern of results in this study suggests that there is more to the stress of vigilance tasks than display or discrimination characteristics per se. Hancock (1998) and Scerbo (1998a) suggested that vigilance is aversive because the task is imposed by either an experimenter or a manager in situations where the operator cannot easily quit.

### *Coping*

It was expected that the use of the highly salient features in configural displays might facilitate the adoption of task-focused coping strategies to deal with task demands rather than emotion-focused or avoidant coping methods. Alternatively, if the coping strategy participants use to deal with the stress of vigilance is similar across display formats, then configurality might be unrelated to coping strategy. Although results conformed to the latter position—there were no significant display differences in coping style—there were task-based differences, such that participants in the midpoint identification tasks were more likely to engage in emotion-focused or avoidant coping. Recall that the midpoint identification task induced less global workload than the distance monitoring task. It may be that in the midpoint identification task, the lower workload increased the boredom often experienced in vigilance (see Scerbo, 1998b), leading the participants to engage in coping methods that distracted them from the situation in attempts to compensate for their boredom. These results may indicate a cost to the performance benefits of high signal salience in terms of the coping strategies adopted to deal with the boredom of the easier task. In the bar graph–different baselines condition, the use of emotion-focused and avoidant coping may have resulted from a reduction in allocation of resources to the task due to appraisals by participants of their capacity to achieve task goals.

The coping results differ from those of Matthews and Campbell (1998), who observed a positive correlation between emotion-focused coping and perceived workload. In the current study higher levels of emotion-focused coping in the midpoint identification task were associated with reductions in perceived workload. Matthews and Campbell (1998) also observed correlations between stress state and coping strategy (e.g., emotion-focused coping related to greater tense arousal [distress], whereas task-focused coping was related to greater motivation [task engagement]), a relationship not observed here. Participants in the two task conditions reported similar stress levels, but participants in the midpoint identification task reported using emotion-focused and avoidant coping strategies to a greater extent than their cohorts in the distance monitoring task conditions. These results suggest that the relationships between coping strategy and perceived workload and stress reported by Matthews and Campbell (1998) may be task dependent.

## *Conclusions*

The results of this study indicate that increasing the salience of signals in a vigilance task using configural displays with easily perceived features can improve performance and reduce the decrement and that these effects are equivalent for displays with emergent features (separable configural displays) or physical contours (object displays). However, performance differences (or their absence) between displays may not entirely capture the information processing demands of tasks, and concluding that configural displays with salient emergent features are functionally equivalent to those with physical contours may mask the subjective cost of performance and important differences in how workload and performance are linked. Although it improves performance and reduces the workload of vigilance, the use of emergent features in object displays may not reduce the stress of sustained attention or encourage the use of task-focused coping strategies. The results of the present study reveal that there are hidden costs (latent performance decrements; Hockey, 1997) to vigilance performance, in terms of cognitive state, even when highly salient configural displays are used. These hidden costs may place operators charged with monitoring tasks at risk for performance failure. Designers who use configural displays to improve performance should therefore assess the workload and stress associated with their use in vigilance tasks.

###### NOTE

Address correspondence about this article to James L. Szalma, Performance Research Laboratory, Department of Psychology, Building 99, Suite 320, Room 351, University of Central Florida, P.O. Box 161390, Orlando, FL 32816-1390 (e-mail: James.Szalma@ucf.edu).

###### REFERENCES

Bennett, K. B., & Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors, 34,* 513–533.

Bennett, K. B., & Flach, J. M. (2011). *Display and interface design: Subtle science, exact art*. Boca Raton, FL: CRC Press.

Buttigieg, M. A., & Sanderson, P. M. (1991). Emergent features in visual display design for two types of failure detection tasks. *Human Factors, 33,* 631–651.

Carswell, C. M., & Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics, 30,* 511–527.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Craig, A. (1978). Is the vigilance decrement simply a response towards probability matching? *Human Factors, 20,* 441–446.

Craig, A. (1979). Nonparametric measures of sensory efficiency for sustained monitoring tasks. *Human Factors, 21,* 69–78.

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

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121,* 275–277.

Groemer, G., Gruber, V., Bishop, S., Peham, D., Wolf, L., & Hogl, B. (2010). Human performance data in a high workload environment during the simulated Mars expedition "AustroMars." *Acta Astronautica, 66*(5/6), 780–787.

Hancock, P. A. (1996). Effects of control order, augmented feedback, input device, and practice on tracking performance and perceived workload. *Ergonomics, 39,* 1146–1162.

Hancock, P. A. (1998). The price of freedom. *Proceedings of the Human Factors and Ergonomics Society, 42,* 1577–1578.

Hancock, P. A., & Szalma, J. L. (2003). Operator stress and display design. *Ergonomics in Design, 11,* 13–18.

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

Hancock, P. A., & Weaver, J. L. (2005). On time distortion under stress. *Theoretical Issues in Ergonomics Science, 6,* 193–211.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, The Netherlands: North-Holland.

Helton, W. S., Matthews, G., & Warm, J. S. (2009). Stress state mediation between environmental variables and performance: The case of noise and vigilance. *Acta Psychologica, 130,* 204–213.

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica, 129,* 18–25.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive–energetical framework. *Biological Psychology, 45,* 73–93.

Hwang, S. L., Yau, Y. J., Lin, Y. T., Chen, J. H., Huang, T. H., Yenn, T. C., & Hsu, C. C. (2008). Predicting work performance in nuclear power plants. *Safety Science, 46,* 1115–1124.

Matthews, G., & Campbell, S. E. (1998). Task-induced stress and individual differences in coping. *Proceedings of the Human Factors and Ergonomics Society, 42,* 821–825.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion, 2,* 315–340.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Erlbaum.

Molloy, R., & Parasuraman, R. (1994). Automation-induced monitoring efficiency: The role of display integration and redundant color coding. In R. Parasuraman & M. Mouloua (Eds.), *Human performance in automated systems: Current research and trends* (pp. 224–228). Hillsdale, NJ: Erlbaum.

Oron-Gilad, T., Szalma, J. L., Stafford, S. C., & Hancock, P. A. (2008). The workload and performance relationship in the real world: A study of police officers in a field shooting exercise. *International Journal of Occupational Safety and Ergonomics, 14,* 119–131.

Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of mental workload. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 305–320). Mahwah, NJ: Erlbaum.

Peebles, D. (2008). The effect of emergent features on judgments of quantity in configural and separable displays. *Journal of Experimental Psychology: Applied, 14*(2), 85–100.

Scerbo, M. W. (1998a). Fifty years of vigilance research: Where are we now? *Proceedings of the Human Factors and Ergonomics Society, 42,* 1575–1576.

Scerbo, M. W. (1998b). What's so boring about vigilance? In R. R. Hoffman, M. F. Sherrick, & J. S. Warm (Eds.), *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 145–166). Washington, DC: American Psychological Association.

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.

See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors, 39,* 14–29.

Szalma, J. L., Hancock, P. A., Dember, W. N., & Warm, J. S. (2006). Training for vigilance: The effect of KR format and dispositional optimism and pessimism on performance and stress. *British Journal of Psychology, 97,* 115–135.

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors, 42,* 183–194.

Teo, G. W. L., & Szalma, J. L. (2010). The effect of spatial and temporal task characteristics on performance, workload, and stress. *Proceedings of the Human Factors and Ergonomics Society, 54,* 1699–1703.

Warm, J. S., & Jerison, H. J. (1984). The psychophysics of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 15–59). Chichester, UK: Wiley.

Warm, J. S., Matthews, G., & Finomore, V. S. (2008). Vigilance, workload, and stress. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (pp. 115–141). Hampshire, UK: Ashgate.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50,* 433–441.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology, 25,* 229–235.

Yeh, Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors, 30,* 111–120.

## Library Recommendation Form

THE
# American Journal
OF
# Psychology

**To:** Librarian / Library Acquisition Committee

**From:** ______________________________

**Position:** ____________________ **Dept:** ____________

**Email:** ____________________

**Phone:** __________________

I recommend that the library subscribe to *The American Journal of Psychology*. Please include this journal in your next serials review meeting with my recommendation to subscribe.

**Signature:** ______________________________

**Date:** __________

For subscription information please contact the University of Illinois Press online at http://www.press.uillinois.edu/journals/ajp.html