UNIVERSITY OF CINCINNATI

Date: March 24, 2006

I, Victor S. Finomore Jr. **,**

hereby submit this work as part of the requirements for the degree of:

Master of Arts

in:

Psychology

It is entitled:

Effects of Feature Presence/Absence and Event Asynchrony on

Vigilance Performance and Perceived Mental Workload.

This work and its defense approved by:

Chair: Joel S. Warm

Gerry Matthews

Michael A. Riley

William N. Dember

Effects of Feature Presence/Absence and Event Asynchrony on Vigilance Performance and Perceived Mental Workload

A thesis submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

in part fulfillment of the
requirements for the degree of

MASTER OF ARTS

In the Department of Psychology
of the College of Arts and Science

2006

by

Victor S. Finomore Jr.

B.S. University of Dayton, 2003

Committee Chair: Dr. Joel S. Warm.

Abstract

The utility of a new measure of perceived mental workload in vigilance, The Multiple Resources Questionnaire (MRQ), was evaluated by comparing it against a standard measure in that area, the NASA-TLX, in sensitivity to the effects of factors theoretically predicted to affect task demand, event asynchrony and search asymmetry (detecting stimulus presence/absence). Contrary to expectation, the former had little impact upon performance but the latter did; detection probability was significantly greater when critical signals for detection were defined by stimulus presence than stimulus absence. This effect was echoed in higher workload scores for absence than presence when workload was measured by the NASA-TLX but not by the MRQ, indicating poorer sensitivity for the new instrument. On the other hand, the MRQ did identify resources utilized in the vigilance task that are not reflected in the standard measure. Therefore, the new scale could be a useful adjunct to the older one.

TABLE OF CONTENTS

Page

ABSTRACT ……………………………………………………………..… ii

TABLE OF CONTENTS……………………………………………..… iv

LIST OF TABLES…………………………………………………..... vi

LIST OF FIGURES……………………………………………….…. vii

CHAPTER 1

Introduction……………………………………………………..... 1

THE PROBLEM OF VIGILANCE…………………………………… 1
    Vigilance in an Automated World……………………….… 1
    Historical Roots……………………………………….. 2

THE PSYCHOPHYSICS OF VIGILANCE…………………….. 4
    Multidimensional factors……………………………………… 4
    The Monk (1984) suggestion: Search Asymmetry………. 5
    Event Asynchrony…………………………………...… 8

THE WORKLOAD OF SUSTAINED ATTENTION…………... 9
    The Arousal and Mindlessness Models………………….. 9
    Research with the NASA-TLX…………………………. 10
    The Multiple Resources Questionnaire………………..… 13

CHAPTER 2

Method……………………………………………………….. 16
Participants…………………………………………………… 16
Design……………………………………………...…...…….. 16
Apparatus…………………………………………………….. 16
Procedure…………………………………………………….. 18

Chapter 3

Results………………………………………………………... 20

VIGILANCE PERFORMANCE……………………………….. 20
    Detection Probability…………………………………… 20
    False Alarms……………………………………………. 21

WORKLOAD: NASA-TLX……………………………………. 22
    Global Scores…………………………………………… 22
    Weighted Workload Ratings……………………………. 22

WORKLOAD: MRQ………………………………………… 24
    MRQ Global Workload………………………………. 24
    Revised Global Scoring……………………………… 26
    Resource Profile……………………………………… 27
    Inter-Scale Correlations……………………………… 28

CHAPTER 4

Discussion……………………………………………..… 30
    Vigilance Performance………………………………..… 30
    Perceived Mental Workload…………………………. 33

REFERENCES……………………………………………..… 37

APPENDIX A: NASA- Task Load Index…………………………… 54

APPENDIX B: Multiple Resources Questionnaire…………..……… 59

APPENDIX C: Task Instructions……………….…………………… 62

APPENDIX D: Informed Consent Form…………………………..… 67

APPENDIX E: ANOVA Tables……………………………...… 70

E1. Analysis of Variance for correct detection scores………..… 71

E2. Analysis of Variance for false alarm scores……………..… 71

E3. Analysis of Variance for NASA-TLX global workload scores……………………………………………………... 71

E4. Analysis of Variance for NASA-TLX subscale scores…..… 72

E5. Analysis of Variance for MRQ global workload scores…… 72

E6. Analysis of Variance for modified MRQ global workload scores ……………………...…………………………..… 72

E7. Analysis of Variance for the modified MRQ resource dimension scores.……….. 73

LIST OF TABLES

Table Page

1. The 17 Dimensions of the Multiple Resources Questionnaire…...… 15
2. Mean percentages of correct detections for all combinations of task type, event schedule, and periods of watch……………………….… 20
3. Mean percentages of false alarms for all combinations of task type, event schedule, and periods of watch…………………………..… 21
4. Mean NASA-TLX global workload scores for all combinations of task type and event schedule…………………………………………..… 22
5. Mean NASA-TLX rating scores for all combinations of task type and event schedule……………………………………………………..… 23
6. Mean MRQ global workload scores for all combinations of task type and event schedule…………………………………………………..… 24
7. Percentage of observers using a non-zero rating for each of the MRQ resource dimensions…………………………………………….. 26
8. Mean MRQ global workload scores of the selected eight resource dimensions for all combinations of task type and event schedule….. 26
9. Mean rating scores of the selected MRQ dimensions for all combinations of task type and event schedule……………………… 27
10. Correlations between the eight resource dimensions of the MRQ and the six subscales of the NASA-TLX……………………………...… 29

LIST OF FIGURES

Figure Page

1. Examples of neutral events and critical signals in the feature presence/absence conditions…………………………………………... 18

Chapter 1

Introduction

The Problem of Vigilance

Vigilance in an Automated World. The study of vigilance or sustained attention focuses upon the ability of observers to detect and respond to unpredictable events over extended periods of time (Ballard, 1996; Davies & Parasuraman, 1982; Warm, 1984, 1993). This aspect of human performance is an important concern for human factors/ergonomic specialists due to the critical role that vigilance plays in many operational settings, especially those involving automated human-machine systems. Advancements in technology have transformed the role of workers from that of active controllers to system executives who monitor the functioning of machines that do the work for them and intervene only in the event of potential problems (Sheridan, 1970, 1980). Consequently, vigilance is a critical component of human performance in a diverse array of work environments including military surveillance, air-traffic control, transportation security, nuclear power plant regulation, industrial quality control, and long-distance driving (Hancock & Hart, 2002; Hartley, Arnold, Kobryn, & Macleod, 1989; Satchel, 1993; Warm, 1984, 1993). Vigilance also contributes to performance efficiency in medical settings, including x-ray and cytological screening and the inspection of anesthesia gauges during surgery (Gill, 1996; Warm & Dember, 1998; Weinger & Englund, 1990).

Although automation has reduced the information-processing load placed upon observers and has enhanced productivity (Parasuraman, 1987; Warm, 1993; Wiener, 1984, 1985), it appears to be a double-edge sword. Several studies have shown that accidents ranging in scale from minor to major are often the result of vigilance failure on the part of human operators (Molloy & Parasuraman, 1996). One solution to this dilemma would be to eliminate the need for the human component in automated systems. However, as Parasuraman (1987) has argued, a solution of that sort is not feasible because of the need for human operators to serve in a fail-safe capacity in the case of system malfunction. With this in mind, an understanding of the factors that influence vigilance performance and their underlying mechanisms is crucial for system integrity and public safety (Nickerson, 1992; Warm & Dember, 1998).

Historical Roots. The term "vigilance" was coined by Sir Henry Head (1923), who used it to refer to a state of maximum physiological and psychological readiness to react. Early research on this topic was conducted by Wyatt and Langdon (1932), who described time-related variations in the performance of inspectors examining cartridge cases for flaws prior to packaging. However, controlled laboratory research on sustained attention is generally considered to have begun during World War II when British airborne observers on patrol over the Bay of Biscay used "blips" on pulse-position radar displays to signify the presence of German U-boats on the surface of the sea below (Warm, 1984). This new and innovative technology should have given these observers a major advantage in combating the U-boat threat to the British war effort. However, after only about 30-min on watch, the well-trained and highly motivated observers began to miss the radar signals and the undetected U-boats were free to sink allied ships.

The Royal Air Force commissioned Norman Mackworth to study this problem. Toward that end, he created a simulated radar display called the "Clock Test" (Mackworth, 1948, 1950/1961). It consisted of a black pointer that moved along the circumference of a blank-face clock devoid of any scale markings to serve as reference points. The pointer rotated around the face of the clock in 0.3-inch jumps at a rate of one jump per second. Occasionally, it would make a double jump of 0.6 inches signaling a critical signal for the observer to detect by pressing a button. In Mackworth's experiments and *most of the others that have followed,* observers were tested individually for a prolonged and continuous period of time (2 hours in Mackworth's case) under conditions in which the signals to be detected were clearly perceivable when observers were alerted to them, but were not compelling changes in the operating environment. The critical signals for detection occurred infrequently and aperiodically, and the observers' responses had no effect on signal occurrence.

The "Clock Test" permitted Mackworth to confirm in a controlled laboratory environment field-generated suspicion that the quality of sustained attention is fragile, waning quickly over time. He found that signal detections declined from 85% to 75% within the first 30 minutes on task and continued to drop more gradually for the remainder of the 2-hour vigil. This drop in performance over time is known as the *vigilance decrement* or the *decrement function*. It has been replicated in many subsequent studies and is the most ubiquitous finding in vigilance research (Davies & Parasuraman, 1982; See, Howe, Warm, & Dember, 1995; Warm, 1984). Generally, the major portion of the decrement appears within the first 15-min of watch (Teichner, 1974), but it can appear within the first 5-min on task when conditions are highly demanding (Helton, Dember, Warm, & Matthews, 2000; Jerison, 1963; Neuchterlein, Parasuraman, & Jiang, 1983; Rose, Murphy, Byard, & Zikzad, 2002; Temple et al., 2000). The vigilance decrement has been found with experienced as well as with inexperienced observers, and it appears in operational environments as well as in laboratory settings (Baker, 1962; Colquhoun, 1967, 1977; Pigeau, Angus, O'Neill, & Mack, 1995; Schmidke, 1976). Studies of the psychophysical factors that determine the decrement and the overall level of signal detection in vigilance tasks and experiments focusing on the mental workload imposed by these tasks have played major roles in efforts to understand the nature of sustained attention (Johnson & Proctor, 2004; Matthews, Davies, Westerman, & Stammers, 2000; Warm, Dember, & Hancock, 1996; Warm & Jerison, 1984). These approaches are featured in this investigation.

The Psychophysics of Vigilance

Multidimensional factors. As Warm and Jerison (1984) have noted, signal detection in vigilance studies requires that energy from the monitored display be transformed and encoded by the observer's perceptual system. Thus, as in the study of other perceptual phenomena, research into the nature of sustained attention has profited from the precise determination of the stimulus characteristics that influence performance efficiency (Dember & Warm, 1979). A substantial database has emerged demonstrating that the quality of sustained attention depends upon a variety of stimulus dimensions, including the sensory modality and the salience of the signals to be detected as well as upon the temporal and spatial contexts in which they appear. Acoustic, tactile, and visual stimuli have been used in vigilance tasks, and the sensory channel in which stimuli are delivered has a major impact upon performance efficiency. In general, the overall speed and accuracy of signal detections tend to be greater for auditory than for either visual or tactual signals (Buckner & McGrath, 1963; Craig, Colquhoun, & Corcoran, 1967; Hawkes & Loeb, 1962), and the vigilance decrement tends to be less pronounced in the case of acoustic vigilance tasks than for their visual and tactual analogs (Hawkes & Loeb, 1961; Sipowicz & Baker, 1961; Ware, 1961). As is often the case in perceptual tasks, the likelihood of signal detection in vigilance increases with increments in signal intensity and duration (Adams, 1956; Baker, 1963; Guralnick, 1972; Loeb & Binford, 1963; Metzger, Warm, & Senter, 1974; See et al., 1995; Thurmond, Binford, & Loeb, 1970; Warm, Loeb, & Alluisi, 1970; Wiener, 1964). Moreover, as observers become more certain about when and where signals will appear through increments in signal frequency and temporal regularity and through reductions in uncertainty about the spatial locations in which signals will come into view, they can more effectively align attention with signal occurrences, and signal detections increase accordingly (Adams & Boulter, 1964; Baddeley & Colquhoun, 1969; Helton et al., 2005; Hollander, Warm, Matthews, Dember, & Parasuraman, 2002; Jenkins, 1958; Krulewitz & Warm, 1977; Kulp & Alluisi, 1967; Methot & Huitema, 1998; Milosevic, 1974; Warm, Dember, Murphy, & Dittmar, 1992; Warm, Epps, & Ferguson, 1974; Williges, 1971). A key additional psychophysical factor in vigilance is task complexity, defined by the number of displays to be monitored. In an initial study, Jerison (1963) demonstrated that the overall level of signal detection in vigilance declined as the number of items (Mackworth clocks) to be examined for critical signals increased. Later studies by Grubb, Warm, Dember, and Berch (1995) and by Miller, Warm, Dember, and Schumsky (1998), using simulated aircraft displays, confirmed this effect by showing that the probability of signal detection varies inversely with the number of instruments to be monitored.

The Monk (1984) Suggestion: Search Asymmetry. Along with the traditional perceptual factors outlined above, Monk (1984) has suggested that a psychophysical approach to vigilance might profit from an examination of stimulus dimensions that affect the efficiency of performance in visual search tasks. This suggestion has been verified in regard to a well-established finding in search tasks known as *search asymmetry* (Quinlan, 2003; Treisman & Gormican, 1988) - the finding that target detection is more rapid when searching for a distinguishing feature in an array of stimuli as opposed to searching for the absence of that feature. When searching for presence, the critical feature appears to be so salient that it seems to “pop out” of the display. According to the feature integration model proposed by Treisman and Gormican (1988), searching for feature-presence is guided by parallel, preattentive processing, while more deliberate, serial processing is needed for determining feature-absence.

Studies by Schoenfeld and Scerbo (1997,1999) and Hollander et al. (2004) have extended the psychophysics of vigilance by incorporating the presence/absence distinction inherent in the search asymmetry effect into performance in sustained attention tasks. Treisman and Gormican's (1988) notion that searching for the absence of a feature is more capacity-demanding than searching for its presence led Schoenfeld and Scerbo (1997,1999) to examine task complexity in term of stimulus presence/absence. They predicted that the complexity effect would be stronger when observers were required to detect feature absence than presence. Consistent with that prediction, they found that in the absence condition, signal detection declined as the size of the stimulus array to be monitored was increased from two to five elements. Increasing array size, however, had no effect on performance when observers were asked to monitor for feature-presence.

As described by Warm (1993), vigilance experiments frequently employ dynamic displays in which critical signals for detection appear within an ensemble of recurrent nonsignal events. For example, observers may be asked to detect occasional "brighter" flashes of light in a background of dimmer flashes, occasional longer lines in a background of shorter lines, or occasional longer duration acoustic pulses in a background of shorter duration pulses. Although the background events may be neutral in the sense that they require no overt response from observers, they are far from neutral in their influence on the quality of sustained attention. The frequency of background events (the "background event rate") is a critical determinant of performance efficiency. The accuracy of signal detection varies inversely with event rate, and the vigilance decrement and the effects of signal amplitude tend to be more pronounced in the context of a fast as compared to a slow event rate (Galinsky, Rosa, Warm, & Dember, 1993; Galinsky, Warm, Dember, Weiler, & Scerbo, 1990; Jerison & Pickett, 1964; Krulewitz & Warm, 1977; Lanzetta, Dember, Warm, & Berch, 1987; Loeb, & Binford, 1968; Metzger et al., 1974; Moore & Gross, 1973; Parasuraman, 1979; Taub & Osborne, 1968; Todkill & Humphreys, 1994).

The event rate effect has been accounted for in terms of an information-processing model in which it is assumed that fast event rates drain more resource capacity than slow event rates because the need to make frequent and rapid signal/noise discriminations is greater under conditions of a fast as compared to a slow event rate (Davies & Parasuraman, 1982; Parasuraman, Warm, & Dember, 1987; Warm & Dember, 1998). Support for an account of this sort comes from experiments by Bowers (1982) and Parasuraman (1985) in which a secondary-task procedure, often used in attention studies to assess resource demands (Matthews et al., 2000; Wickens & Hollands, 2000), was employed. In the Bowers (1982) and Parasuraman (1985) studies, observers were asked to perform a vigilance task and also a probe-detection task under instructions that their primary responsibility was to do well on the former and that performance on the latter was secondary. In both studies, probe detection was significantly slower in the presence of a fast as compared to a slow primary event rate. These findings and the belief that searching for feature-absence is more capacity-demanding than searching for feature-presence (Treisman & Gelade, 1980; Treisman & Gormican, 1988), led Hollander and his associates (2004) to predict that the negative effects of increments in event rate would be more pronounced when the critical signals for detection were defined by feature absence than presence. Consistent with that expectation, detection probability declined significantly as event rate was increased from 6 events/min to 24 events/min in the absence condition while event rate had no effect upon performance in the presence condition.

Event Asynchrony. One goal for the present study was to examine the implications of the feature presence/absence distinction for still another aspect of the background event ensemble–the temporal structure of the schedule of background events in which critical signals for detection are embedded. It is typical in vigilance studies for the background events to appear in a temporally synchronous manner such as once every 10 sec at a slow event rate of 6 events/min or once every 2.5 sec at a faster event rate of 24 events/min. Background events can also be scheduled to appear in an asynchronous or temporally irregular manner. In the example above, the mean inter-event intervals could be 10 sec at an event rate of 6 events/min or 2.5 sec at an event rate of 24 events /min but in both cases, these means could represent the central tendency of a range of inter-event intervals. Under the synchronous conditions, observers can generate veridical expectations about when an event requiring inspection will appear. Therefore, they do not have to continually monitor the display. When the schedule of background events is asynchronous, however, observers are not certain when an event requiring inspection will appear and must, therefore, monitor the display continually. Consequently, it might be anticipated that a vigilance task featuring an asynchronous schedule of background events will consume more information-processing resources than one featuring synchronous background events, leading to poorer performance in the asynchronous background event condition. In accord with that anticipation, studies by Richter, Senter, and Warm (1981) and Scerbo and his associates (Scerbo, Warm, & Fisk, 1987; Scerbo, Warm, Doettling, Parasuraman, & Fisk, 1986) have shown that an asynchronous schedule degrades signal detection in comparison to a synchronous schedule. Given the Treisman and Gormican (1988) proposal about differential capacity-demand in detecting feature absence and presence, and the finding of Hollander et al. (2004) that the degrading effects of event rate are amplified in the stimulus absence case, it might also be anticipated that the effects of event asynchrony would be more pronounced when observers must monitor for the absence than for the presence of a feature. The present study was designed, also, to test that possibility.

The Workload of Sustained Attention.

The Arousal and Mindlessness Models. From the description of the characteristics of the typical vigilance experiment given at the outset of this chapter, one could gain the impression that vigilance tasks are tedious and understimulating assignments that impose little workload upon observers. An impression of that sort has formed the basis of two theories of vigilance performance. One of these is the long-standing *arousal* or *activation model* that accounts for the vigilance decrement in terms of the lack of stimulation necessary to maintain alertness. According to that model, the repetitive and monotonous aspects of vigilance tasks reduce the level of stimulation needed by elements of the central nervous system—the ascending reticular formation, the locus ceroleus, and the diffuse thalamic projection system—necessary to succor wakefulness and alertness. As a result, the brain becomes less responsive to stimulation, and performance efficiency declines (Aston-Jones, 1985; Frankmann & Adams, 1962; Heilman, 1995; Loeb & Alluisi, 1977; Nachreiner & Hanecke, 1992; Proctor & Van Zandt, 1994; Welford, 1968). A more recent view, also based upon understimulation, is the *mindlessness* model suggested by Robertson and his colleagues (Manly, Robertson,& Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997) in which the repetitive and tedious nature of vigilance tasks is considered to lead observers to withdraw attentional effort from the task and approach their assignment in a thoughtless, routinized manner. As noted by Dickman (2002) and by Helton et al. (2005), this approach reflects an endogenous modulation of attention rather than a decline in wakefulness and vigor accompanying lowered arousal. Both of these models point to a potential paradox of automation: Although designed to reduce workload, such a reduction in the case of vigilance may place observers at a functional disadvantage due to understimulation.

Warm, Dember, and Hancock (1996) have indicated, however, that the view of vigilance tasks as under-arousing stemmed from a rather superficial task analysis; it was not based upon the actual degree of underload inherent in these tasks. Warm and his associates (1996) attempted to provide such evidence through measurements of the perceived mental workload or the information-processing load and/or resource demands imposed by such tasks (Eggemeier, 1988; Gopher & Braune, 1984; O'Donnell & Eggemeier, 1986) and found evidence to indicate that the cost of mental operations in vigilance is not at all consistent with the understimulation notion that gave rise to the arousal and mindlessness models.

Research with the NASA-TLX. In their effort to assess the workload of sustained attention, Warm and his associates (1996) made use of the NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988). This instrument is considered to be one of the most effective measures of perceived mental workload currently available (Farmer & Brownson, 2003; Hill, Iavecchia, Byers, Zaklad, & Christ, 1992; Lysaght et al., 1989; Nygren, 1991; Proctor & Van Zandt, 1994; Wickens & Hollands, 2000). It provides a reliable measure of overall or global workload on a scale from 0 to 100 (test-retest reliability = .83) and also identifies the relative contributions of six sources of workload. Three of these sources reflect the demands that tasks place upon operators—Mental, Temporal, and Physical Demand—while the remainder characterize the interaction between observers and the tasks that confront them—Performance, Effort, and Frustration. The NASA-TLX is reproduced in Appendix A.

It is important to emphasize that the NASA-TLX is essentially a subjective scale, and Natsoulas (1967) has pointed out that there is always some question as to whether any form of self-report accurately reflects respondents' true perceptual experiences. He suggested that this problem might be overcome by linking perceptual reports to psychophysical factors known to influence task difficulty. With this suggestion in mind, Warm and his associates (1996) attacked the problem of measuring perceived mental workload in vigilance by linking workload ratings secured by the NASA-TLX to psychophysical factors that degrade performance efficiency on these types of tasks. In an extensive series of studies, they found that the decline in signal detections over time is accompanied by a linear increase in overall workload, that overall workload increases as signal salience decreases, and that overall workload is positively related to increments in observers' uncertainty as to the location of critical signal appearances and to increments in event rate. In all of these studies, the global workload scores fell within the upper end of the NASA-TLX scale. As such, they exceeded those typically observed with other types of laboratory tasks, such as memory search, mental arithmetic, grammatical reasoning, choice-reaction time, and simple tracking, and were similar to those observed with a motion-based flight simulator (Hancock, 1988; Hancock, Rodenburg, Mathews, & Vercruyssen, 1988; Hart & Staveland, 1988; Liu & Wickens, 1987; Sanderson & Woods, 1987). In addition, these studies also revealed a consistent workload signature among the NASA-TLX subscales in which Mental Demand and Frustration were the primary components of the workload associated with vigilance tasks. The results obtained in the studies described by Warm et al. (1996) have been replicated in several other experiments using the NASA-TLX (Deaton & Parasuraman, 1993; Dittmar, Warm, Dember, & Ricks, 1993; Grier et al., 2003; Grubb et al., 1995; Helton et al., 2005; Matthews, 1996; Miller et al.,1998; Parsons et al., 2000; Scerbo, Greenwald, & Sawin, 1993; Szalma et al., 2004; Temple et al., 2000; Warm, Dember, & Parasuraman, 1991). Of special interest in regard to the current investigation, Schoenfeld and Scerbo (1997, 1999) and Hollander and his associates (2004) have also found workload to be greater when the critical signals for detection are defined by signal absence than by signal presence.

In a recent review of the problems inherent in the need to sustain attention, Johnson and Proctor (2004) have noted that research with the NASA-TLX has important theoretical implications. They affirm that the finding of high information-processing demand in vigilance tasks challenges arousal theory and supports a view, such as that proposed by Parasuraman and his associates (Davies & Parasuraman, 1982; Parasuraman & Davies, 1977; Parasuraman, 1984; Parasuraman, Warm, & Dember, 1987; Warm & Dember, 1998), that the workload imposed by vigilance tasks reflects the impact of focused mental effort and a drain on information-processing resources. A similar argument regarding the mindlessness model has been made by Grier et al. (2003) and by Helton et al. (2005).

A critical component of the argument that research with the NASA-TLX supports a resource model of vigilance is the belief that the high workload scores on the scale arise from direct costs associated with the vigilance task itself. However, Scerbo (1998) and Sawin and Scerbo (1995) have argued that before the direct cost belief can be accepted, it is necessary to eliminate an indirect cost possibility in which the high workload associated with vigilance tasks arises not from the information-processing demand imposed by those tasks, but from observers' efforts to overcome the tedium and boredom also associated with such tasks. Hitchcock, Dember, Warm, Moroney, and See (1999) and Alikonis, Warm, Matthews, Dember, and Kellaris (2002) have carried out experiments employing converging operations to do just that. Hitchcock and his associates (1999) sought to disengage workload and boredom by providing observers with accurate cues as to the imminent arrival of critical signals. In this way, they expected the information-processing demand of the vigilance task to be reduced because observers would only need to inspect the vigilance display when cued to signal arrival. On the other hand, since the tedious and repetitive nature of the task environment remained unchanged under the cueing condition, boredom was expected to be unaffected by the cueing manipulation. As anticipated, cueing significantly reduced the perceived mental workload of the vigilance task but had no effect upon boredom. Alikonis and her associates (2002) sought to disengage workload and boredom in another way. Taking advantage of the fact that music has been found to be effective in modifying observers' moods and emotions (Hargreaves & North, 1999; Lewis, Dember, Schefft, & Radenhausen, 1995), Alikonis and her associates (2002) asked observers to listen to a pleasant musical selection while performing a vigilance task. They expected that the music would reduce the boredom of the task. However, since the musical background afforded observers no aid in regard to signal/noise discriminations, it was not expected to reduce the workload induced by the vigilance task. Consistent with these expectations, Alikonis and her associates reported that boredom was reduced but workload was unaffected by the musical background. Clearly, vigilance research with the NASA-TLX has withstood the indirect cost challenge and has, therefore, been of significant value in identifying the presence of high information-processing demand in a task that had heretofore incorrectly been considered a quintessential example of task underload.

The Multiple Resources Questionnaire. To date, all of the research on the perceived mental workload of vigilance tasks has been carried out with the NASA-TLX. Boles and Adair (2001a) have suggested that as useful as it has been, this instrument is nevertheless subject to a potentially important limitation—it is overly restrictive with respect to the mental processes that it represents. More specifically, the NASA-TLX treats resources in the way that they were conceptualized in Kahneman's (1973) original derivation of resource theory, as a pool of undifferentiated information-processing entities that could be parceled out to one or more tasks as needed. However, multi-tasking studies in short-term attention and in vigilance have demonstrated that certain pairings of tasks or task components produce greater dual-task processing deficits than others, indicating that attention may not be unitary; it is or is not subject to interference when processing simultaneous events depending upon the demands that are made upon similar resource pools (Azuma, Prinz, & Koch, 2004; Boles, 2001; Caggiano & Parasuraman, 2004; Kantowitz & Knight, 1976; Navon & Gopher, 1979; Wickens, 1980; Wickens & Hollands, 2000).

In an effort to address this concern, Boles and Adair (2001a) have offered a new instrument, The Multiple Resources Questionnaire (MRQ), in which observers are presented with a set of multiple mental processes based upon a combination of dimensions drawn from Wickens' Multiple Resource Theory (Wickens, 1984, 1991, 1992) and from factor-analytic studies carried out by Boles (1998) and Boles and Law (1998). The MRQ consists of the 17 resource dimensions listed in Table 1. Fifteen of the dimensions reflect encoding/central processing resources; the remaining two are response resources. Using a scale from 0 (no usage) to 4 (extreme usage), observers are asked to rate the extent to which a task they just performed utilized each dimension. The complete MRQ is reproduced in Appendix B.

Research with the MRQ has indicated that interrater reliability approximates 0.9 when ratings are aggregated over at least eight individuals (Boles & Adair, 2001a) and that the instrument is able to uncover different key resource dimensions in tasks involving different skills such as reading bar graphs, determining the spatial position of a line, word interpretation, and medical imaging (Boles & Adair, 2001b; Klein, Riley, Warm, & Matthews, 2005). In addition, the MRQ has been successful in predicting the interference between tasks based upon shard resource dimensions (Boles & Adair, 2001b; Boles, Phillips, Bursk, & Perdelwitz, 2004; Phillips & Boles, 2004). Accordingly, a second goal for this study was to provide the initial application of the MRQ to vigilance by contrasting the sensitivity of the MRQ and the NASA-TLX to variations in psychophysical demand brought about by the feature presence/absence and event asynchrony dimensions and to identify the specific MRQ resources that are involved in the vigilance task featured in this study.

Table 1
The 17 Dimensions of the Multiple Resources Questionnaire

- Auditory Emotional Process
- Auditory Linguistic Process
- Facial Figural Process
- Facial Motive Process
- Manual Process
- Short Term Memory Process
- Spatial Attentive Process
- Spatial Categorical Process
- Spatial Concentrative Process
- Spatial Emergent Process
- Spatial Positional Process
- Spatial Quantitative Process
- Tactile Figural Process
- Visual Lexical Process
- Visual Phonetic Process
- Visual Temporal Process
- Vocal Process

In summary, the goals for the present study were twofold. One was to extend the integration of the search asymmetry effect to vigilance performance by determining if feature presence/absence can serve as a moderator variable for the event asynchrony component of the background event ensemble in which critical signals for detection are embedded, as it does for the event rate component of that key contextual element in the psychophysics of vigilance. The second goal focused upon the perceived mental workload of sustained attention by comparing the sensitivity of the newly developed MRQ with the standard measure of workload used in vigilance, the NASA-TLX, to the effects of factors theoretically predicted to affect task demand, event asynchrony and monitoring for stimulus presence/absence.

Chapter 2

Method

Participants

Eighty undergraduate students (40 men and 40 women) from introductory psychology classes at the University of Cincinnati served as observers for course credit. They ranged in age from 18 - 44 years, with a mean age of 21 years. All observers had normal or corrected-to-normal vision.

Design

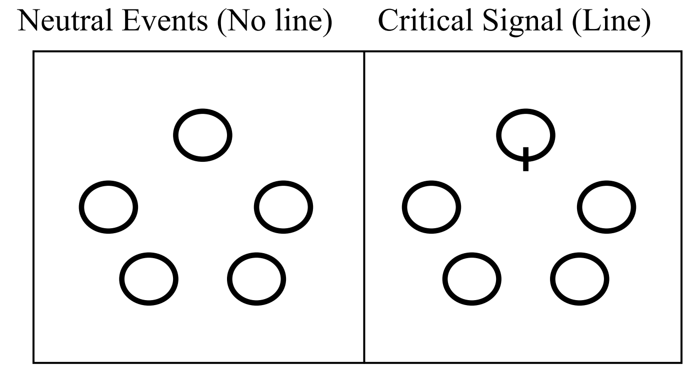
Twenty observers (10 men and 10 women) were assigned at random (using a block randomization procedure) to each of four experimental conditions defined by the factorial combination of two task types (presence/absence) and two background event schedules (synchronous/asynchronous) with the restriction that the conditions were equated for sex.

Apparatus

In all experimental conditions, observers participated in a continuous 40-min vigil divided into four 10-min periods of watch during which they monitored an array of five open circles (14 mm diameter) outlined by a 1mm black line (transluminance = 0.11 cd/m^2) that appeared on the white background (transluminance = 64.32 cd/m^2) of a video display terminal (VDT). The circles were positioned 75 mm from the center of the VDT at the 3, 5, 7, 9, and 12 o'clock locations. Examples of the feature presence and absence conditions are illustrated in Figure 1.

The critical signal for detection in the presence condition was the appearance of a vertical 4 × 1 mm black line (transluminance = 0 .11 cd/m^2) intersecting the 6 o'clock position within *one* of the five circles in the display. In the absence condition, the vertical 6 o'clock line was present in all circles *but one*. The Michaelson contrast ratio (Coren, Ward, & Enns, 2004) of the circles' contours/target lines to their background was 99.6%. Displays were updated once every four sec in the synchronous condition, resulting in an event rate of 15 events/min. An identical event rate was maintained in the asynchronous condition. However, in that condition, updates appeared randomly for each observer at intervals of 1.8, 2.2, 3.0, 4.0, or 9.0 sec, with a mean *inter-event interval* of four sec. All stimuli were exposed for 0.25 sec. For each observer in each condition, the intervals between *critical signals* varied at random over a range of 20 – 120 sec, with the restrictions that signals came into view on an average of once/min during each period of watch (signal probability = .067) and the signals appeared equally often on each of the five circles comprising the vigilance display during each period of watch. In all conditions, observers signified their detection of critical signals by pressing the spacebar on a computer keyboard. Responses occurring within 1.5 sec after the appearance of critical signals were recorded as correct detections; responses to non-signal events were classified as false alarms. Pilot work ensured that signals in the presence and absence conditions were equally detectable under alerted conditions. Complete task instructions can be found in Appendix C.

Feature-Presence Case



Feature-Absence Case

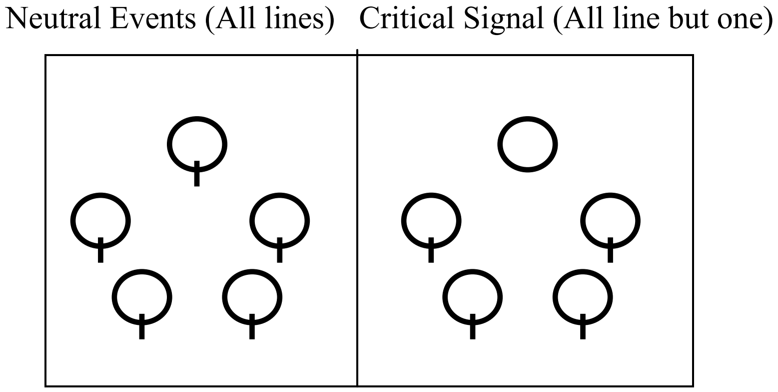


Figure 1. Examples of neutral events and critical signals in the feature presence/absence conditions.

Procedure

Observers were tested individually in a 2.0 × 1.9 × 1.9 m Industrial Acoustic Sound Chamber. Ambient illumination in the chamber (5.12 cd/m^2) was provided by a 25-watt light bulb housed in a parabolic reflector located above and behind the observer and angled to reduce glare on the VDT. The VDT was positioned on a table at eye-level 55 cm from and directly in front of the seated observer. Stimulus presentation and response recording were orchestrated by a Dell personal computer (Dimension 2400) running *SuperLab* software (Cedrus, version 2.0).

Upon reporting for the experiment, observers received a verbal briefing about the task they were to perform and completed an informed-consent form (see Appendix D). They were then given two 5-min practice trials that duplicated the forthcoming vigilance task. A computerized male voice provided feedback as to correct detections, misses, and false alarms during practice. To be retained in the study, observers had to detect at least 80 percent of all critical signals during the second practice trial and to commit no more than 10 percent false alarms. All observers met these criteria. The mean percentage of correct responses across all combinations of task type and background event schedule in the second practice trial was 92%. Feedback was not available during the main vigil. Computerized versions of the NASA-TLX and the MRQ were administered immediately upon completion of the main vigil. The order of administration of the scales was balanced across sexes in each experimental condition. Observers surrendered their timepieces, cellphones, and pagers upon entering the laboratory and had no knowledge of the length of the experimental session other than it would not exceed 90 min.

Chapter 3

Results

Vigilance Performance

Detection Probability. Mean percentages of correct detections and their associated standard errors are presented in Table 2 for all combinations of task type, event schedule, and periods of watch.

Table 2
Mean percentages of correct detections for all combinations of task type, event schedule,and periods of watch. Standard errors are in parentheses.

		Periods (10 minutes)				
Task Type	Event Schedule	1	2	3	4	Mean
Presence	Synchronous	95.0 (4.0)	93.0 (2.3)	89.0 (5.1)	89.5 (3.2)	91.6
	Asynchronous	97.0 (1.6)	95.0 (1.9)	92.0 (1.9)	88.5 (4.7)	93.1
Absence	Synchronous	79.0 (4.4)	71.5 (5.4)	70.5 (5.2)	66.5 (4.8)	71.9
	Asynchronous	84.5 (2.9)	72.0 (4.8)	79.0 (4.2)	71.5 (4.5)	76.8
	Mean	88.9	82.9	82.6	79.0	83.4

Examination of Table 2 will reveal that the probability of correct detections was greater in the feature-presence condition (M = 92.4%) than in the feature-absence condition (M = 74.4%), and that signal detections were slightly more frequent in the asynchronous (M = 84.9%) than in the synchronous condition (M = 81.7%). In addition, it is evident in the table that the likelihood of signal detection decreased consistently with time on task. A 2 (task type) × 2(event schedule) × 4 (periods of watch) mixed-analysis of variance (ANOVA) was performed on the arcsines of the percentage scores. The transformation was used to normalize the percentage data (Kirk, 1995). Statistically significant main effects were found for task type, F (1, 76) = 43.20, $p < .001$, and periods of watch, F (2.60, 197.45) = 8.56, $p < .001$. The difference between the two event schedules was not significant and all of the interactions in the analysis lacked significance, $p > .05$ in all cases. In this and all subsequent ANOVAs, Box's epsilon was used to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004). Complete summaries of this and all subsequent ANOVAs are presented in Appendix E.

False Alarms. Mean percentages of false alarms and their associated standard errors are presented in Table 3 for all combinations of task type, event schedule, and periods of watch.

Table 3
Mean percentages of false alarms for all combinations of task type, event schedule and periods of watch. Standard errors are in parentheses.

	Event	Periods (10 minutes)				
Task Type	Schedule	1	2	3	4	Mean
Presence	Synchronous	0.5 (0.4)	0.9 (0.6)	0.9 (0.7)	0.8 (0.7)	0.8
	Asynchronous	0.9 (0.4)	0.9 (0.7)	1.0 (0.5)	1.2 (0.7)	1.0
Absence	Synchronous	4.9 (1.6)	2.8 (1.0)	2.4 (0.8)	2.9 (1.0)	3.3
	Asynchronous	3.3 (0.9)	2.3 (0.7)	2.3 (0.7)	2.8 (1.0)	2.7
	Mean	2.4	1.7	1.7	1.9	2.0

Inspection of Table 3 will reveal that false alarms were more frequent in the absence (M = 3.0%) than in the presence (M = 0.9%) condition, that false alarms were slightly more likely in the synchronous (M = 2.0%) than in the asynchronous (M = 1.8%) condition, and that the frequency of false alarms generally declined from the first to the last period of watch. A 2 (task type) × 2 (event schedule) × 4 (periods of watch) mixed-ANOVA on the arcsines of the percentage scores revealed that the difference in the frequency of false alarms between the two task types was statistically significant, $F(1, 76) = 17.40$, $p < .001$. All of the other sources of variance in the analysis lacked significance, $p > .05$ in each case.

Workload: NASA-TLX

Global scores. Following the procedure outlined by Hart and Staveland (1988), global workload and weighted rating subscale scores were determined for each observer using the ratings given to each subscale and a paired comparison procedure in which the subscales were judged against each other for their importance in contributing to the workload of the task at hand. Mean NASA-TLX global workload scores and their associated standard errors for all combinations of task type and event schedule are presented in Table 4.

Table 4
Mean NASA-TLX global workload scores for all combinations of task type and event schedule. Standard errors are in parentheses.

	Event Schedule		
Task Type	Synchronous	Asychronous	Mean
Presence	44.4 (4.4)	46.5 (4.3)	45.5
Absence	66.4 (2.5)	61.0 (2.5)	63.7
Mean	55.4	53.8	54.6

It is evident in Table 4 that observers found their vigilance assignment to be demanding since the global workload scores generally fell near or above the mid-level of the NASA-TLX scale (50). It is also evident in the table that there was a substantial disparity in perceived mental workload between the two task types and a moderate difference between the two event schedules. An ANOVA of the data of Table 4 revealed that perceived mental workload was significantly greater in the absence than in the presence condition, F (1, 76) = 27.02, $p < .001$. The main effect for event schedule and the Task Type × Event Schedule interaction were not significant sources of variation in this analysis, $p > .05$ in these cases.

Weighted Workload Ratings. Mean weighted ratings on the subscales of the NASA-TLX are presented in Table 5. It can be seen in the table that Mental Demand ($M = 218.5$) contributed most to workload and that Physical Demand ($M = 36.7$) contributed least.

Table 5
Mean NASA-TLX rating scores for all combinations of task type and event schedule. Standard errors are in parentheses.

Task Type	Event Schedule	NASA-TLX Subscales MD	PD	TD	P	E	F	Mean
Presence	Synchronous	201.0	11.5	85.5	72.5	133.8	160.0	110.7
		(37.6)	(3.6)	(17.1)	(16.0)	(25.3)	(36.3)	
	Asynchronous	196.5	56.8	86.5	84.5	173.0	99.8	116.2
		(28.9)	(24.3)	(20.9)	(12.3)	(27)	(30.2)	
Absence	Synchronous	282.8	46.5	188.5	82.8	178.8	216.5	166.0
		(35.8)	(25.4)	(30.1)	(14.5)	(24.6)	(30.4)	
	Asynchronous	193.5	31.8	192.5	142.0	154.8	203	152.9
		(26.7)	(18.1)	(25.1)	(28.1)	(25.8)	(42.5)	
	Mean	218.5	36.7	138.3	95.5	160.1	169.9	136.5

Note: MD= Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration.

The data of Table 5 were subjected to a 2 (task type) × 2 (event schedule) × 5 (subscales) mixed-ANOVA. Due to the paired-comparison method used in determining the dimensional weightings (Hart & Staveland, 1988), the Physical Demand subscale was dropped from the ANOVA in order to meet the independence assumption of the statistical procedure (Kirk, 1995). Consistent with results obtained with the global workload data, the overall mean weighted rating score for the absence condition (M = 183.5) was significantly greater than that for the presence condition (M = 129.3), F (1, 76) = 24.97, p < .001. In addition, the main effect for subscales was also significant, F (3.13, 237.73) = 9.83, p < .001. All of the remaining sources of variance in the analysis lacked significance, p > .05 in each case. Supplementary Tukey-tests with alpha set at .05 confirmed the dominance of Mental Demand in contributing to workload in this study; the mean of that subscale was significantly greater than that of each of the other subscales. The means of the Frustration and Effort subscales, which did not differ significantly from each other, were significantly greater then that for the Performance subscale, while the mean of the Temporal Demand subscale did not differ significantly from any of these three. All in all, it would appear that next to Mental Demand, Frustration and Effort were the major contributors to workload in this investigation

Workload: MRQ

MRQ Global Workload: Like the NASA-TLX, the MRQ provides two indices of workload for an observer: (1) a global index defined as the sum of the observer's ratings across all of the scale's resource dimensions and (2) a profile of resource contribution to workload defined by the absolute value of the rating given to each resource dimension (Boles & Adair, 2001a). Mean global workload scores based upon the 17 resource dimensions of the MRQ and their associated standard errors are presented for all combinations of task type and event schedule in Table 6. In this case, individual global workload scores could range from 0-68.

Table 6
Mean MRQ global workload scores for all combinations of task type and event schedule. Standard errors are in parentheses.

	Event Schedule		
Task Type	Synchronous	Asychronous	Mean
Presence	23.6 (2.6)	25.7 (2.8)	24.7
Absence	29.0 (2.1)	22.3 (2.2)	25.7
Mean	26.3	24	25.2

It is immediately evident in the table that the global workload scores generally fell below the mid-level of the MRQ scale (34). Therefore, the substantial level of task demand reflected in the NASA-TLX scale was not duplicated in the MRQ measure. Moreover, unlike the NASA-TLX, the MRQ scale did not reflect differential levels of workload associated with task type since all of the main effects and the interaction in an ANOVA of the data of Table 6 failed to reach statistical significance, $p > .05$ in each case.

The relatively low workload scores and the absence of task-type differences in the MRQ results may be due, at least in part, to the “no-usage” (zero) value employed in the rating scale. More specifically, Boles and Adair (2001a) have pointed out that including resources that are rated as having “no usage” in calculating the global score might distort the workload picture reflected in the MRQ by masking the utilization magnitude of the resources that did contribute to the performance of the task in question. Consequently, they suggest that it would be desirable in calculating a global score to first delete items from the questionnaire that did not meet a “greater than zero usage” standard. One approach to such a standard would be to include only those resource dimensions in which significantly more than 50% of the observers gave ratings other than zero. Toward that end, the percentage of observers using a rating greater than zero was determined for each of the 17 resource dimensions of the scale and the values were tested against 50% by means of *t*-tests using a significance level of .05 and the Bonferroni correction for the number of tests made. As can be seen in Table 7, eight of the 17 dimensions, those denoted by a star, met this criterion. Accordingly, these dimensions were considered to be the mental processes tapped by the vigilance tasks used in this study and were the dimensions included in a revised global scoring of the MRQ and the subsequent resource profile.

Table 7

Percentage of observers using a non-zero rating for each of the MRQ resource dimensions

Subscales	Percentage non-zero	
Vocal	20	
Tactile Figural	31	
Facial Figural	34	
Auditory Linguistic	39	
Auditory Emotional	42	
Visual Phonetic	43	
Visual Lexical	46	
Facial Motive	61	
Spatial Quantitative	64	
STM	70	*
Spatial Concentrative	75	*
Spatial Positional	81	*
Spatial Emergent	83	*
Visual Temporal	83	*
Spatial Categorical	89	*
Manual Process	93	*
Spatial Attentive	96	*

Note: * Item met inclusion standard

Revised Global Scoring. Mean global workload scores based upon the eight resource dimensions of the MRQ that met the usage criterion and their associated standard errors are presented for all combinations of task type and event schedule in Table 8. In this case, individual global scores could range from 0-32.

Table 8

Mean MRQ global workload scores of the selected eight resource dimensions for all combinations of task type and event schedule. Standard errors are in parentheses.

	Event Schedule		
Task Type	Synchronous	Asychronous	Mean
Presence	17.2 (1.5)	17.4 (1.4)	17.3
Absence	19.6 (1.1)	16.1 (1.6)	17.9
Mean	18.4	16.8	17.6

Perusal of the table will reveal that when only the eight resource dimensions meeting the usage criterion were considered, global workload scores generally fell above the midpoint (16) of the revised scale, a result that duplicated the substantial level of task demand reflected in the NASA-TLX scale. On the other hand, the scale still did not reveal differential levels of workload to be associated with either the task type or event schedule dimensions since the main effects of these factors and their interaction failed to reach statistical significance in an ANOVA of the revised global workload data, *p* > .05 in each case.

Resource Profile. The revised global analysis identified the resources that were engaged by the vigilance tasks in this study but it did not differentiate among the relative contributions made by those resources. To answer to that question, comparisons were made of the ratings given to each of the eight resource dimensions. Mean ratings and their associated standard errors are presented in Table 9 for each dimension under all combinations of task type and event schedule.

Table 9

Mean rating scores of the selected MRQ dimentions for all combinations of task type and event schedule. Standard error are in parentheses.

Task Type	Event Schedule	MP	ST	SA	SC	S	SE	SP	VT	Mean
Presence	Synchronous	2.6	1.5	3.5	2.0	1.4	1.9	2.2	2.2	2.2
		(0.2)	(0.3)	(0.2)	(0.3)	(0.3)	(0.4)	(0.3)	(0.3)	
	Asynchronous	2.0	1.7	3.3	2.3	1.8	2.4	2.0	2.2	2.2
		(0.2)	(0.3)	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	
Absence	Synchronous	2.3	2.3	3.5	2.7	1.8	2.7	2.1	2.3	2.5
		(0.2)	(0.3)	(0.2)	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	
	Asynchronous	2.0	1.2	3.2	1.8	1.7	2.6	2.2	1.5	2.0
		(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.3)	(0.3)	
	Mean	2.2	1.7	3.4	2.2	1.7	2.4	2.1	2.1	2.2

Note: MP= Manual Process, ST = STM, SA = Spatial Attentive, SC = Spatial Categorical, S = Spatial Concentrative, SE = Spatial Emergent, SP = Spatial Positional, VT = Visual Temporal

An ANOVA of the data of Table 9 revealed a significant main effect for resource dimensions, F (5.80, 440.50) = 19.95, p <.001. None of the other sources of variance in the analysis were significant, p > .05 in all cases. Supplementary Tukey tests with alpha set at .05 revealed that the Spatial Attentive (SA) dimension was the dominant mental process tapped in this study; the mean for that dimension significantly exceeded that for each of the other dimensions. The Spatial Emergent (SE) dimension also appeared to play an important role; the mean for that dimension significantly exceeded that for the Short Term Memory (ST) and Spatial Concentrative (S) dimensions. All other comparisons in the supplementary tests were not significant.

Inter-Scale Correlations. Given that the MRQ is a new measure of perceived mental workload in vigilance, it is important to determine its co-variation with the current standard in that area, the NASA-TLX. Accordingly, correlations were determined between global values on the adjusted MRQ scale and global and subscale values on the NASA-TLX and between the eight resource dimensions on the adjusted MRQ scale and the subscales of the NASA-TLX. To compensate for differences in ranges of values, the relevant scores for each observer were converted to *z-scores* based upon the overall means and standard deviations of the distributions involved. The correlation between the global workload scores of both instruments was not significant, r (78) = .16, p > .05. Correlations between the global adjusted MRQ and each of the subscales of the NASA-TLX were: Mental Demand r = .29; Physical Demand r = .02; Temporal Demand r = .03; Performance r = .08; Effort r = .13; Frustration r = -.16. The only significant value at the .05 level was that for Mental Demand, and that value was modest. Correlations between the eight resource dimensions of the adjusted MRQ and the subscales of the NASA-TLX are displayed in Table 10. Seven of the correlations reached significance at the .05 level.

They involved the Mental Demand, Physical Demand, Performance, Effort, and Frustration subscales of the NASA-TLX and the Spatial Concentrative, Spatial Emergent, Spatial Positional, Spatial Categorical, and Short Term Memory dimensions of the MRQ. Two aspects of this correlational array are notable: the majority of the significant interscale correlations involved the Mental Demand subscale of the NASA-TLX and, like the correlation between Mental Demand and the global score of the adjusted MRQ, they were modest in magnitude. Using the z′ procedure for averaging correlations outlined by McNemar (1969), the mean value of the subscale correlations was .24.

Table 10

Correlations between the eight resource dimensions of the MRQ and the six subscales of the NASA-TLX

NASA-TLX Subscales	MRQ Resource Dimensions							
	Manual Process	Short Term Memory	Spatial Attentive	Spatial Categorical	Spatial Concentrative	Spatial Emergent	Spatial Positional	Visual Temporal
MD	.12	.14	.20	.27*	.21	.23*	.23*	.07
PD	-.01	.16	-.19	-.07	.24*	-.08	-.09	.10
TD	.03	.12	-.06	.14	-.07	.01	-.06	.00
P	-.07	.04	-.18	.05	.24*	.07	.18	.02
E	.02	.23*	.02	.07	.10	.08	.17	-.06
F	-.08	-.17	-.02	-.04	-.19	.03	-.27*	-.06

Note: MD = Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration.

* = significant correlation

Chapter 4

Discussion

Vigilance Performance. The present study was designed, in part, to examine the role of feature presence/absence as a moderator variable for the degrading effects of event asynchrony on vigilance performance. Based upon the feature integration model in which detecting feature-absence is considered to be more capacity-demanding than detecting feature-presence (Treisman & Gelade, 1980; Treisman & Gormican, 1988) and Hollander et al.'s (2004) finding that stimulus presence/absence moderates the effects of variations in event rate, it was anticipated that the negative effects of event asynchrony would be more pronounced when observers were required to monitor for feature absence than presence. However, the anticipated Event Schedule × Task Type interaction did not materialize. Indeed, contrary to prior studies (Richter, et al., 1981; Scerbo et al., 1987; Scerbo et al., 1986), event asynchrony had no significant effect upon performance efficiency in the present case.

One possibility for the lack of an event-asynchrony effect may be that the level of variability in the event schedule used in this study was more constrained than in past experiments. This is not the case, however. The inter-event intervals employed in this study covered a range of 1.8 to 9 sec, while those in earlier studies encompassed ranges of 1 to 3 sec (Richter et al., 1981) or 0.6 to 3 sec (Scerbo et al., 1987; Scerbo et al., 1986), ranges that were narrower than that employed here.

A more likely possibility is that the lack of an event asynchrony effect may be related to the successive/simultaneous discrimination categories employed by Parasuraman and Davies to portray the manner in which observers identify critical signals for detection in vigilance tasks (Davies & Parasuraman, 1982; Parasuraman & Davies, 1977). Successive discriminations represent absolute judgments, in which observers must compare current input against a standard retained in recent memory in order to separate signals from noise. The latter represent comparative judgments, in which all the information needed for signal detection is present in the stimuli themselves with little involvement of recent memory for the signal characteristic. Considerable evidence is available to indicate that successive discriminations are more capacity-demanding than are simultaneous discriminations (Parasuraman et al., 1987; Warm & Dember, 1998). The feature presence/absence tasks employed in this study involved simultaneous-type discriminations since observers were required to detect either the presence or the absence of a line in a concurrently observable array of circles. To date, all of the studies that have been successful in demonstrating the event-asynchrony effect have used successive-type discriminations (Richter et al., 1981; Scerbo et al., 1987; Scerbo et al., 1986). The only report of a nonsignificant effect for event asynchrony has been made by Scerbo and his associates (1987), who also found a null effect for this factor when simultaneous-type discriminations were involved. Since task factors that adversely affect detection efficiency in vigilance situations generally combine in their effects on performance (Matthews et al., 2000; Warm & Jerison, 1984), it is possible that the need to maintain stimulus representation in memory may be a key factor in driving the event-asynchrony effect. In this sense, the information-processing demand imposed upon observers by *the event-asynchrony* component of the background event ensemble in which critical signals appear may be less potent than that of *the event rate component.* Indeed, the event rate component exerts its effects on signal detection regardless of whether simultaneous or successive type discriminations are involved, although fast event rates are more degrading in the context of the latter than the former (Lanzetta et al., 1987). Moreover, the degrading effects of event rate are modified by the need to monitor for feature presence or absence (Hollander et al., 2004).

In contrast to the null effect for event asynchrony, stimulus presence/absence had a substantial impact upon both performance efficiency and perceived mental workload in this study. Consistent with the earlier reports by Schoenfeld and Scerbo (1997, 1999) and Hollander et al. (2004), performance efficiency was significantly poorer when observers were required to monitor for stimulus absence than presence. Along that line, it is noteworthy that signals were detected significantly less frequently and false alarms occurred significantly more often in the absence than in the presence condition, a result suggesting a loss in perceptual sensitivity in the absence condition (Macmillan & Creelman, 2005). Additionally, as in the earlier studies by Schoenfeld and Scerbo (1997, 1999) and Hollander et al. (2004), global workload as measured by the NASA-TLX was also higher when critical signals for detection were defined by feature absence than presence. Clearly, feature presence/absence is a critical psychophysical parameter in the study of sustained attention.

As noted earlier, interest in the feature presence/absence dimension evolved from Monk's (1984) suggestion that factors affecting visual search may also have importance for vigilance performance. The arrow of benefit in the relation between search and vigilance may not be one-sided; it may point both ways. One of the key elements of Treisman and Gormican's (1988) feature integration theory of visual information-processing was the assumption that detecting feature-presence in an array of elements is not only less capacity-demanding than detecting feature-absence, it is preattentive in character and incurs little or no information-processing cost. However, an emerging view in the search literature is that the alignment of feature detection with preattentive processing may no longer be tenable (Pashler, 1998; Quinlan, 2003). The present results support that view in several ways, beginning with the finding of a similar vigilance decrement in both the feature-absence and the feature-presence conditions. To the extent that the decrement reflects the consumption of information-processing resources that are not replenished over time (Davies & Parasuraman, 1982; Parasuraman & Davies, 1977; Warm & Dember, 1998), the finding of a decrement in the feature-presence condition implies that some information-processing cost must be associated with feature detection. This interpretation is bolstered by the findings that although the overall workload in the presence condition was less than that in the absence condition, it still fell at the mid-level of the NASA-TLX scale (M = 45.5) and that Mental Demand was the dominant contributor to workload in both the feature-presence and feature-absence conditions. Further evidence confirming the conclusion that feature detection is not preattentive comes from the MRQ scale which revealed that several encoding/central processing resources were involved in the performance of both the presence and absence versions of the vigilance tasks used in this study and that the Spatial Attentive dimension was the dominant mental process tapped by both of these conditions. Results of this sort undermine the notion that detecting feature-presence in an array of stimulus elements is preattentive.

Perceived Mental Workload. In addition to a focus upon psychophysical issues per se, the present study was also designed to provide the initial test of the applicability of the MRQ to the workload imposed by sustained attention tasks. As described by O’Donnell and Eggemeier (1986), key dimensions in the evaluation of workload scales are sensitivity—the scale’s ability to reflect changes in task difficulty or resource demand—and diagnosticity—the scale’s ability to discriminate the sources of demand. With regard to the former, the present results indicate that the MRQ does not fare well in comparison to the current standard measure of the workload of sustained attention, the NASA-TLX. Contrary to the picture portrayed by that scale of vigilance as a highly demanding work assignment, the mean global workload scores based upon the 17 resource dimensions of the MRQ depicted the workload of vigilance as much more moderate, with scores falling well below the mid-point of the MRQ scale. In accord with Boles and Adair's (2001a) suggestion, the MRQ was revised to include only those dimensions in which significantly more than 50% of the observers gave ratings other than zero. When that was done, the revised MRQ yielded results that were comparable to the NASA-TLX in that the global scores now fell above the mid-point of the scale. Nevertheless, even with that modification, the MRQ failed to detect a global workload difference in the feature presence/absence conditions, a difference that would be anticipated on the basis of the theoretical view that feature-absence is more capacity-demanding than detecting feature-presence (Treisman & Gormican, 1988). That anticipated difference was detected in both the global and weighted scoring systems employed with the NASA-TLX in this study and also in the earlier studies by Schoenfeld and Scerbo (1997; 1999) and Hollander et al. (2004) which utilized the NASA-TLX scale. What might account for the lack of sensitivity of the MRQ? Since workload ratings for each subscale of the NASA-TLX can range from 0 to 100 while those for each dimension of the modified MRQ are scored from 0-4, "range restriction" (Nunnally & Bernstein, 1994) may minimize the ability of the MRQ to reflect task induced workload differences.

Although the MRQ was found to be lacking in sensitivity in this study, the instrument did point to the involvement of a set of mental resources that observers drew upon in carrying out their vigilance assignment. These resources can have theoretical and practical importance for an understanding of vigilance performance. Testimony to the theoretical importance has already been given above in regard to the preattentiveness issue in feature detection. On the practical side, a key performance problem is the determination of conditions in which the quality of sustained attention will suffer from interference by other tasks (Caggiano & Parasuraman, 2004). Because of the set of resources it has uncovered, the MRQ may be useful in that regard, as it has been in predicting interference in a variety of multi-tasking situations (Boles & Adair, 2001b; Boles et al., 2004; Phillips & Boles, 2004).

In addition to issues of sensitivity and diagnosticity, the adequacy of workload scales, like that of other psychological measures, needs to be considered in terms of the validity of the instruments. With regard to the MRQ, the finding that Spatial Concentrative, Spatial Positional, Spatial Emergent, Visual Temporal, Spatial Categorical, and Spatial Attentive resources played significant roles in this study fits well with the nature of a vigilance assignment in which observers were required to focus their attention continuously for 40 min on an array of circles in which critical signals for detection could appear in unpredictable spatial locations at unpredictable times. Consequently, the present results speak well for the content validity (Anastasi & Urbina, 1997; Nunnally & Bernstein, 1994) of the MRQ scale. Along that line, it is important to point out that the profile of resource demands in vigilance provided by the present results is not necessarily invariant. Other resources involving auditory and short-term memory dimensions may become more prominent when acoustic stimuli and successive-type discriminations characterize the vigilance tasks that confront observers. Similarly, lexical and phonetic resources may be a factor when the discriminations involved in the vigilance task are more cognitive than, as in the present case, sensory in nature (e.g., See et al., 1995). Additional research is necessary to explore these possibilities.

Along with content validity, the construct validity of the MRQ is also a key factor to consider by examining the correlations between the new instrument and the standard instrument used to measure the workload of sustained attention, the NASA-TLX (Anastasi & Urbina, 1997;

Kerlinger & Howard, 2000). The picture with regard to construct validity is complex. The global workload scores of the two instruments did not correlate significantly. On the other hand, the global workload score of the adjusted MRQ, as well as several of the MRQ resource dimensions, did show significant, although modest, covariation with the Mental Demand subscale of the NASA-TLX. The finding that scores on the MRQ were most closely associated with the Mental Demand subscale of the NASA-TLX but that the shared variance in these measures was small is consistent with the fact that Mental Demand is the dimension of the older scale that focuses most closely on thinking, decision making, and searching (see Appendix A), activities that are also emphasized in the MRQ, and with Boles and Adair's (2001a) assertion that the MRQ was designed to measure mental dimensions not incorporated in the NASA-TLX. Perhaps the most meaningful conclusion at this point is that while the different scales do tap some common elements, they also describe different facets of the workload imposed by vigilance tasks and each can therefore be seen as contributing something different to our knowledge of how observers sustain attention.

REFERENCES

Adams, J. A. (1956). Vigilance in detection of low-intensity visual stimuli. *Journal of Experimental Psychology, 52*, 204-208.

Adams, J. A., & Boulter, L. R. (1964). Spatial and temporal uncertainty as determinants of vigilance performance. *Journal of Experimental Psychology, 67*, 127-131.

Alikonis, C. R., Warm, J. S., Matthews, G., Dember, W. N., & Kellaris, J. J. (2002, March). *Effects of music on the workload and boredom of sustained attention.* Paper presented at the meeting of the Southern Society for Philosophy and Psychology, Nashville, TN.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Aston-Jones, G. (1985). Behavioral functions of locus coeruleus derived from cellular attributes. *Physiological Psychology, 13*, 118-126.

Azuma, R., Prinz, W., & Koch, I. (2004). Dual-task slowing and the effects of cross-task compatibility. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 57*, 693-713.

Baddeley, A. D., & Colquhoun, W. P. (1969). Signal probability and vigilance: A reappraisal of the "signal rate" effect. *British Journal of Psychology, 60*, 165-178.

Baker, C. H. (1962). *Man and radar displays*. New York: Macmillan.

Baker, C. H. (1963). Signal duration as factor in vigilance tasks. *Science, 141*, 1196-1197.

Ballard, J. C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology, 18*, 843-863.

Boles, D. B. (1998, October). *Multiple lateralized spatial processes.* Paper presented at the meeting of the Psychonomic Society, Dallas, TX.

Boles, D. B. (2001). Multiple resources. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (pp. 271-275). New York: Taylor & Francis.

Boles, D. B., & Adair, L. P. (2001a). The multiple resources questionnaire (MRQ). *Proceedings of the Human Factors and Ergonomics Society,45,*1790-1794.

Boles, D. B., & Adair, L. P. (2001b). Validity of the multiple resources questionnaire (MRQ). *Proceedings of the Human Factors and Ergonomics Society,45,*1795-1799.

Boles, D. B., & Law, M. B. (1998). A simultaneous task comparison of differentiated and undifferentiated hemispheric resource theories. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 204-215.

Boles, D. B., Phillips, J. B., Bursk, J. H., & Perdelwitz, J. R. (2004). Application of the multiple resources questionnaire (MRQ) to a complex gaming environment. *Proceedings of the Human Factors and Ergonomics Society, 48, 1968*-1972.

Bowers, J. C. (1982). *Effects of the rate of repetitive stimulation on probe detection in a vigilance task*. Unpublished master's thesis. University of Cincinnati, Cincinnati OH.

Buckner, D. N., & McGrath, J. J. (1963). A comparison of performance on single and dual sensory mode vigilance tasks. In D. N. Buckner & J. J. McGrath (Eds.), *Vigilance: A symposium*. New York: McGraw-Hill.

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review, 11*, 932-937.

Cedrus (1999). Superlab (Version 2.0) [Computer Software]. San Pedro, CA.

Colquhoun, W. P. (1967). Sonar target detection as a decision process. *Journal of Applied Psychology, 51*, 165-172.

Colquhoun, W. P. (1977). Simultaneous monitoring of a number of auditory sonar outputs. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 163-188). New York: Plenum Press.

Coren, S., Ward, L. M., & Enns, J. T. (2004). *Sensation and perception* (6th ed.). Hoboken, NJ: Wiley.

Craig, A., Colquhoun, W. P., & Corcoran, D. W. J. (1967). Combining evidence presented simultaneously to the eye and the ear: A comparison of some predictive models. *Perception & Psychophysics, 19*, 473-484.

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

Deaton, J. E., & Parasuraman, R. (1993). Sensory and cognitive vigilance: Effects of age on performance and subjective workload. *Human Performance, 6*, 71-97.

Dember, W. N., & Warm, J. S. (1979). *Psychology of perception* (2nd ed.). New York: Holt, Rinehart, & Winston.

Dickman, S. J. (2002). Dimensions of arousal: Wakefulness and vigor. *Human Factors, 44*, 429-442.

Dittmar, M. L., Warm, J. S., Dember, W. N., & Ricks, D. F. (1993). Sex differences in vigilance performance and perceived workload. *Journal of General Psychology, 120*, 309-322.

Eggemeier, F. T. (1988). Properties of workload assessment techniques. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload.* (pp. 41-62). Oxford, UK: North-Holland.

Farmer, E., & Brownson, A. (2003). *Review of workload measurement, analysis and interpretation methods*. Brussels: European Organization for the Safety of Air Navigation (Eurocontrol): Report for Eurocontrol Integra Program.

Frankmann, J. P., & Adams, J. A. (1962). Theories of vigilance. *Psychological Bulletin, 59*, 257-272.

Galinsky, T. L., Rosa, R. R., Warm, J. S., & Dember, W. N. (1993). Psychophysical determinants of stress in sustained attention. *Human Factors, 35*, 603-614.

Galinsky, T. L., Warm, J. S., Dember, W. N., Weiler, E. M., & Scerbo, M. (1990). Sensory alternation and vigilance performance: The role of pathway inhibition. *Human Factors, 32*, 717-728.

Gill, G. W. (1996). Vigilance in cytoscreening: Looking without seeing. *Advance for Medical Laboratory Professionals, 8*, 14-15.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors, 26,* 519-532.

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., et al. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors, 45*, 349-359.

Grubb, P. L., Warm, J. S., Dember, W. N., & Berch, D. B. (1995). Effects of multiple signal discrimination on vigilance performance and perceived workload. *Proceedings of the Human Factors and Ergonomics Society, 39, 1360*-1364.

Guralnick, M. J. (1972). Observing responses and decision processes in vigilance. *Journal of Experimental Psychology, 93*, 239-244.

Hancock, P. A. (1988). The effects of gender and time of day upon the subjective estimate of mental workload during the performance of a simple task. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 239-250). Amsterdam: North-Holland.

Hancock, P. A., & Hart, G. (2002). Defeating terrorism: What can human factors/ergonomics offer? *Ergonomics and Design, 10*, 6-16.

Hancock, P. A., Rodenburg, C. J., Mathews, W. D., & Vercruyssen, M. (1988). Estimation of duration and mental workload at differing times of the day by males and females. *Proceedings of the Human Factors and Ergonomics Society, 32,* 857-861.

Hargreaves, D. J., & North, A. C. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of Music, 27*, 71-83.

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). Oxford, UK: North-Holland.

Hartley, L. R., Arnold, P. K., Kobryn, H., & MacLeod, C. (1989). Vigilance, visual search, and attention in an agricultural task. *Applied Ergonomics, 20*, 9-16.

Hawkes, G. R., & Loeb, M. (1961). Vigilance for cutaneous and auditory signals. *Journal of Auditory Research, 4*, 272-284.

Hawkes, G. R., & Loeb, M. (1962). Vigilance for cutaneous and auditory stimuli as a function of intersignal interval and signal strength. *Journal of Psychology, 53*, 211-218.

Head, H. (1923). The conception of nervous and mental energy. II. Vigilance: A physiological state of the nervous system. *British Journal of Psychology, 14*, 126-147.

Heilman, K. M. (1995). Attentional asymmetries. In R. J. Davidson & K. Hugdahl (Eds.), *Brain asymmetry.* (pp. 217-234). Cambridge, MA: The MIT Press.

Helton, W. S., Dember, W. N., Warm, J. S., & Matthews, G. (2000). Optimism, pessimism, and false failure feedback: Effects on vigilance performance. *Current Psychology, 18*, 311-325.

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., et al. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology, 96*(2), 249-261.

Hill, S. G., Iavecchia, H. P., Byers, A.C., Zaklad, A.L., & Christ, R.E. (1992). Comparison of four subjective workload rating scales. H*uman Factors, 34*, 429-439.

Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors, 41*, 365-372.

Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., & Parasuraman, R. (2002). Effects of signal salience and periodicity: Cerebral hemovelocity in vigilance. *Proceedings of the Human Factors and Ergonomics Society, 46,* 1654-1658 .

Hollander, T. D., Warm, J. S., Matthews, G., Shockley, K., Dember, W. D., Weiler, E. M., et al. (2004). Feature presence/absence modifies the event rate effect and cerebral hemovelocity in vigilance performance. *Proceedings of the Human Factors and Ergonomics Society, 48,* 1943-1947.

Jenkins, H. M. (1958). The effect of signal-rate on performance in visual monitoring. *American Journal of Psychology, 71*, 647-661.

Jerison, H. J. (1963). On the decrement function in human vigilance. In D. N. Buckner & J. J. McGrath (Eds.), *Vigilance: A symposium* (pp. 199-212). New York: McGraw-Hill.

Jerison, H. J., & Pickett, R. M. (1964). Vigilance: The importance of the elicited observing rate. *Science, 143*, 970-971.

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice*. Thousand Oaks, CA: Sage.

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

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping timesharing: II. Auditory secondary task. *Acta Psychologica, 40*, 343-362.

Kerlinger, F. N., & Howard, B. L. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth: Harcourt College Publishers.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove: Brooks/Cole.

Klein, M. I., Riley, M. A., Warm, J. S., & Matthews, G. (2005). Perceived mental workload in an endoscopic surgery simulator. *Proceedings of the Human Factors and Ergonomics Society, 49,*1014-1018.

Krulewitz, J. E., & Warm, J. S. (1977). The event rate context in vigilance: Relation to signal probability and expectancy. *Bulletin of the Psychonomic Society, 10*, 429-432.

Kulp, R. A., & Alluisi, E. A. (1967). Effects of stimulus-response uncertainty on watchkeeping performance and choice reactions. *Perception & Psychophysics, 18*, 245-249.

Lanzetta, T. M., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of task type and stimulus homogeneity on the event rate function in sustained attention. *Human Factors, 29*, 625-633.

Lewis, L. M., Dember, W. N., Schefft, B. K., & Radenhausen, R. A. (1995). Can experimentally induced mood affect optimism and pessimism scores? *Current Psychology: Developmental, Learning, Personality, Social, 14*, 29-41.

Liu, Y., & Wickens, C. D. (1987). The effects of processing code, response modality task, and difficulty on dual task performance and subjective workload in a manual system. *Proceedings of the Human Factors and Ergonomics Society, 31,* 847-851.

Loeb, M., & Alluisi, E. A. (1977). An update of findings regarding vigilance and a reconsideration of underlying mechanisms. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 719-814). New York: Plenum.

Loeb, M., & Binford, J.R. (1968). Variation in performance on auditory and visual monitoring tasks as a function of signal and stimulus frequencies. *Perception & Psychophysics, 4*, 361-367.

Loeb, M., & Binford, J. R. (1963). Some factors influencing the effective auditory intensive difference limen. *Journal of the Acoustical Society of America, 35*, 884-891.

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., Linton, P. M., Wierwille, W. W., et al. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies*. Technical Report 851. U.S. Army Research Institute for the Behavioral and Social Sciences. Alexandria, VA.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 6-21.

Mackworth, N. H. (1950/1961). *Researches on the measurement of human performance*. London: HM Stationary Office.

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

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37,* 661-670.

Matthews, G. (1996). Signal probability effects on high-workload vigilance tasks. *Psychonomic Bulletin & Review, 3*, 339-343.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance: Cognition, stress, and individual differences*. East Sussex, UK: Psychology Press.

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

McNemar, Q. (1969). *Psychological Statistics* (4th ed.). New York: Wiley.

Methot, L.L. & Huitema, B.E. (1998). Effects of signal probability on individual differences in vigilance. *Human Factors, 40,* 102-110.

Metzger, K. R., Warm, J. S., & Senter, R. J. (1974). Effects of background event rate and critical signal amplitude on vigilance performance. *Perceptual and Motor Skills, 38*, 1175-1181.

Miller, L. C., Warm, J. S., Dember, W. N., & Schumsky, D. A. (1998). Sustained attention and feature-integrative displays. *Proceedings of the Human Factors and Ergonomics Society, 42,* 1585-1589.

Milosevic, S. (1974). Effects of time and space uncertainty on a vigilance task. *Perception & Psychophysics, 15*, 331-334.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors, 38*, 311-322.

Monk, T. H. (1984). Search. In J. S. Warm (Ed.), *Sustained attention in human performance*. Chichester, UK: Wiley.

Moore, S. F., & Gross, S. J. (1973). Influence of critical signal regularity, stimulus event matrix and cognitive style on vigilance performance. *Journal of Experimental Psychology, 99*, 137-139.

Nachreiner, F., & Hanecke, K. (1992). *Vigilance*. In A.P. Smith & D.M. Jones (Eds.) Handbook of human performance (pp 262-288). San Diego, CA: Academic Press, Inc.

Natsoulas, T. (1967). What are perceptual reports about? *Psychological Bulletin, 67*, 249-272.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*, 214-255.

Nickerson, R. S. (1992). *Looking ahead: Human factors challenges in a changing world*. Mahwah, NJ: Erlbaum.

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science, 220*, 327-329.

Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill, Inc.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 33*, 17-33.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance, vol. 2: Cognitive processes and performance.* (pp. 41-1-42-49). New York: Wiley.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science, 205*, 924-927.

Parasuraman, R. (1984). Sustained Attention in detection and discrimination. In R. Parasuraman & D.R. Davies (Eds). *Varieties of attention* (pp. 243-271). Orlando: Academic Press.

Parasuraman, R. (1985). Sustained attention: A multifactorial approach. In M. I. Posner & O. S. Marin (Eds.), *Attention and performance XI* (pp. 493-512). New York: Erlbaum.

Parasuraman, R. (1987). Human computer monitoring. *Human Factors, 29*, 695-706.

Parasuraman, R., & Davies, D. R. (1977). A taxonomic analysis of vigilance. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 559-574). New York: Plenum.

Parasuraman, R., & Davies, D. R. (1984). *Varieties of attention*. San Diego, CA: Academic Press.

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In L. S. Mark, J. S.Warm, & R. L. Huston (Eds.), *Ergonomics and human factors: Recent research* (pp. 11-32). New York: Springer-Verlag.

Parsons, K. S., Warm, J. S., Matthews, G., Dember, W. N., Galinsky, T. L., & Hitchcock, E. M. (2000, November). *Changes in signal probability and vigilance performance*. Paper presented at the meeting of the Psychonomic Society, New Orleans, LA.

Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Phillips, J. B., & Boles, D. B. (2004). Multiple resources questionnaire and workload profile: Application of competing models to subjective workload measurement. *Proceedings of the Human Factors and Ergonomics Society, 48, 1963-1967.*

Pigeau, R. A., Agnes, R. G., O'Neil, P., & Mack, I. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors, 37*, 622-634.

Proctor, R. W., & Van Zandt (1994). *Human factors in simple and complex systems*. Needham Heights, MA: Allyn & Bacon.

Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin, 129*, 643-673.

Richter, D. O., Senter, R. J., & Warm, J. S. (1981). Effects of the rate and regularity of background events on sustained attention. *Bulletin of the Psychonomic Society, 18*, 207-210.

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.

Rose, C. L., Murphy, L. B., Byard, L., & Zikzad, K. (2002). The role of the big five personality factors in vigilance performance and workload. *European Journal of Personality, 16*, 185-200.

Sanderson, P. M., & Woods, M. D. (1987). *Subjective mental workload and locus of control* (Report No. EPL-87-03). Urbana, IL: Engineering Psychology Research Laboratory, Department of Mechanical and Industrial Engineering.

Satchel, P. M. (1993). *Cockpit monitoring and alerting systems*. Brookfield, VT: Ashgate.

Sawin, D. A., & Scerbo, M. W. (1995). Effects of instruction type and boredom proneness in vigilance: Implications for boredom and workload. *Human Factors, 37*, 752-765.

Scerbo, M., Warm, J. S., Doettling, V. S., Parasuraman, R., & Fisk, A. D. (1987). Event asynchrony and task demands in sustained attention. In L. S. Mark, J. S.Warm & R. L.

Huston (Eds.), *Ergonomics and human factors/recent research* (pp. 33-39). New York: Springer-Verlag.

Scerbo, M. W. (1998). 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.

Scerbo, M. W., Greenwald, C. Q., & Sawin, D. A. (1993). The effects of subject-controlled pacing and task type on sustained attention and subjective workload. *Journal of General Psychology, 120*, 293-307.

Scerbo, M. W., Warm, J. S., & Fisk, A. D. (1986). Event asynchrony and signal regularity in sustained attention. *Current Psychological Research & Reviews, 5*, 335-343.

Schmidke, H. (1976). Vigilance. In E. Simpson & P. C. Weiser (Eds.), *Psychological and physiological correlates of work and fatigue*. Springfield, IL: Thomas.

Schoenfeld, V. S., & Scerbo, M. W. (1997). Search differences for the presence and absence of features in sustained attention. *Proceedings of the Human Factors and Ergonomics Society, 41,* 1288-1292.

Schoenfeld, V. S., & Scerbo, M. W. (1999). The effects of search differences for the presence and absence of features on vigilance performance and mental workload. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 177-182). Mahwah, NJ: Erlbaum.

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

Sheridan, T. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics SSC-6*, 140-145.

Sheridan, T. (1980). Supervisory control. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1243-1268). New York: Wiley.

Sipowicz, R. R., & Baker, R. A. (1961). Effects of intelligence on vigilance: A replication. *Perceptual and Motor Skills, 13*, 398.

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., et al. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors, 46*, 219-233.

Taub, H. A., & Osborne, F. H. (1968). Effects of signal and stimulus rates on vigilance performance. *Journal of Applied Psychology, 52*, 133-138.

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time on watch. *Human Factors, 16*, 339-353.

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.

Thurmod, J. B., Binford, J. R., & Loeb, M. (1970). Effects of signal-to-noise variability over repeated sessions in an auditory vigilance task. *Perception & Psychophysics, 7*, 100-102.

Todkill, A. L., & Humphreys, M. S. (1994). Stimulus comparison strategies and task demands in successive discrimination. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 47*, 761-779.

Treisman, A.M., & Gelade, G. (1980). A feature integration theory of attention. C*ognitive Psychology, 12,* 97-136.

Treisman, A. M., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*, 15-48.

Ware, J. R. (1961). Effects of intelligence on signal detection in visual and auditory monitoring. *Perceptual and Motor Skills, 13*, 99-102.

Warm, J. S. (1984). An introduction to vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 1-14). Chichester, UK: Wiley.

Warm, J. S. (1993). Vigilance and target detection. In C. D. Wickens & B. Huey (Eds.), *Teams in transition: Workload, stress, and human factors* (pp. 139-170). Washington, DC: National Research Council.

Warm, J. S., & Dember, W. N. (1998). Tests of a vigilance taxonomy. In R. R. Hoffman, M. F. Sherick & J.S.Warm (Eds.), *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 87-112). Washington, DC: American Psychological Association.

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* (183-200). Mahwah, NJ: Erlbaum.

Warm, J. S., Dember, W. N., Murphy, A. Z., & Dittmar, M. L. (1992). Sensing and decision-making components of the signal-regularity effect in vigilance performance. *Bulletin of the Psychonomic Society, 30*, 297-300.

Warm, J. S., Dember, W. N., & Parasuraman, R. (1991). Effects of olfactory stimulation on performance and stress in a visual sustained attention task. *Journal of the Society of Cosmetic Chemists, 42*, 199-210.

Warm, J. S., Epps, B. D., & Ferguson, E. P. (1974). Effects of knowledge of results and signal regularity on vigilance performance. *Bulletin of the Psychonomic Society, 4*, 272-274.

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

Warm, J. S., Loeb, M., & Alluisi, E. A. (1970). Variations in watchkeeping performance as a function of the rate and duration of visual signals. *Perception & Psychophysics, 7*, 97-99.

Weinger, M. B., & Englund, C. E. (1990). Ergonomics and human factors affecting anesthetic vigilance and monitoring performance in the operating room environment. *Anesthesiology, 73*, 995-1021.

Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance viii*. Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

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

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Wiener, E. L. (1964). Transfer of training in monitoring: Signal amplitude. *Perceptual and Motor Skills, 18*, 104.

Wiener, E. L. (1984). Vigilance and inspection. In J. S.Warm (Ed.), *Sustained attention in human performance*(207-246). Chichester, UK: Wiley.

Wiener, E. L. (1985). Beyond the sterile cockpit. *Human Factors, 27*, 75-90.

Williges, R. C. (1971). The role of payoffs and signal ratios in criterion changes during a monitoring task. *Human Factors, 13*, 261-267.

Wyatt, S., & Langdon, J. N. (1932). *Inspection processes in industry*. London: HM Stationery Office.

Appendix A:

NASA- Task Load Index

INSTRUCTIONS: RATINGS

We are not only interested in assessing your performance but also the experiences you had during the different task conditions. Right now we are going to describe the technique that will be used to examine your experiences. In the most general sense we are examining the "Workload" you experienced. Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another. Physical components of workload are relatively easy to conceptualize and evaluate. However, the mental components of workload may be more difficult to measure.

Since workload is something that is experienced individually by each person, there are no effective “rulers” that can be used to estimate the workload of different activities. One way to find out about workload is to ask people to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually rather than lumping them into a single global evaluation of overall workload. This set of six rating scales was developed for you to use in evaluating your experiences during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales in the table, please ask the experimenter about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

After performing the task, six rating scales will be displayed. You will evaluate the task by marking each scale at the point which matches your experience. Each line has two endpoint descriptors that describe the scale. Note that "Performance" goes from “good” on the left to “bad” on the right. This order has been confusing for some people. Left-click the mouse in the desired location. Please consider your responses carefully in distinguishing among the task conditions. Consider each scale individually. Your ratings will play an important role in the evaluation being conducted, thus, your active participation is essential to the success of this experiment, and is greatly appreciated.

RATING SCALE DEFINITIONS

Title	Endpoints	Descriptions
MENTAL DEMAND	***Low/High***	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	***Low/High***	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	***Low/High***	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	***Good/Poor***	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	***Low/High***	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION	***Low/High***	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

MENTAL DEMAND

Low **High**

PHYSICAL DEMAND

Low **High**

TEMPORAL DEMAND

Low **High**

PERFORMANCE

Good **Poor**

EFFORT

Low **High**

FRUSTRATION

Low **High**

INSTRUCTIONS: SOURCES-OF-WORKLOAD EVALUATION

Throughout this experiment the rating scales are used to assess your experiences in the different task conditions. Scales of this sort are extremely useful, but their utility suffers from the tendency people have to interpret them in individual ways. For example, some people feel that mental or temporal demands are the essential aspects of workload regardless of the effort they expended or the performance they achieved. Others feel that if they performed well the workload must have been, low, and vice versa. Yet others feel that effort or feelings of frustration are the most important factors in workload; and so on. The results of previous studies have already found every conceivable pattern of values. In addition, the factors that create levels of workload differ depending on the task. For example, some tasks might be difficult because they must be completed very quickly. Others may seem easy or hard because of the intensity of mental or physical effort required. Yet others feel difficult because they cannot be performed well, no matter how much effort is expended.

The evaluation you are about to perform is a technique that has been developed by NASA to assess the relative importance of six factors in determining how much workload you experienced. The procedure is simple: You will be presented with a series of pairs of rating scale titles (for example, Effort vs. Mental Demands) and asked to choose which of the items was more important to your experience of workload in the task(s) that you just performed. Each pair of scale titles will appear separately on the screen.

Left-click the mouse to select the Scale Title that represents the more important contributor to workload for the Specific task you performed in this experiment.

After you have finished the entire series we will be able to use the pattern of your choices to create a weighted combination of the ratings from that task into a summary workload score. Please consider your choices carefully and make them consistent with how you used the rating scales during the particular task you were asked to evaluate. Don't think that there is any correct pattern; we are only interested in your opinions. If you have any questions, please ask them now.

Appendix B:

Multiple Resources Questionnaire

The Purpose of this questionnaire is to characterize the nature of the mental processes used in the task with which you have become familiar. Please read each carefully so that you understand the nature of the process. Then rate the task on the extent to which it uses each process, using the following scale.

No Usage	Light Usage	Moderate Usage	Heavy Usage	Extreme Usage
0	1	2	3	4

Important:
All parts of a process definition should be satisfied for it to be satisfied for it to be judged as having been used. For example, recognizing geometric figures presented visually should not lead you to judge that the "Tactile Figural process" was used, just because figures were involved. For that process to be used, figures would need to be processed tactilely (i.e., using the sense of touch).
Please judge the task as a whole, average over the time you performed it. If a certain process was used at one point in the task and not another, your rating should not reflect "peak usage" but should instead reflect average usage over the entire length of the task.

Auditory Emotional process – Required judgments of emotion (e.g., tone of voice of musical mood) presented through the sense of hearing. ____

Auditory Linguistic process – Required recognition of words, syllables, or other verbal parts of speech presented through the sense of hearing. ____

Facial Figural process – Required recognition of faces, or of the emotions shown on faces, presented through the sense of vision. ____

Facial Motive process – Required movement of your own face muscles, unconnected to speech or the expressing of emotion. ____

Manual process – Required movement of the arms, hands, and/or fingers. ____

Short Term Memory process – Required remembering of information for a period of time ranging from a couple of seconds to half a minute. ____

Spatial Attentive process – Required focusing of attention on a location, using the sense of vision. ____

Spatial Categorical process – Required judgment of simple left-versus-right or up-versus-down relationship, without consideration of precise location, using the sense of vision.____

Spatial Concentrative process – Required judgment of how tightly spaced are numerous visual objects or forms. ____

Spatial Emergent process – Required "picking out" of a form or object from highly cluttered or confusing background, using the sense of vision. ____

Spatial Positional process – Required recognition of precise location as differing from other locations, using the sense of vision. ____

Spatial Quantitative process – Required judgment of numerical quantity based on a nonverbal, nondigital representation (for example, bar graphs or small clusters of items), using the sense of vision. ____

Tactile Figural process – Required recognition or judgment of shapes (figures), using the sense of touch. ____

Visual lexical process – Required recognition of words, letters, or digits, using the sense of vision. ____

Visual Phonetic process – Required detailed analysis of the sound of words, letters, or digits, presented using the sense of vision. ____

Visual Temporal process – Required judgment of time intervals, or of the timing of events, using the sense of vision. _____

Vocal process – Required use of your voice. _____

Appendix C:

Task Instructions

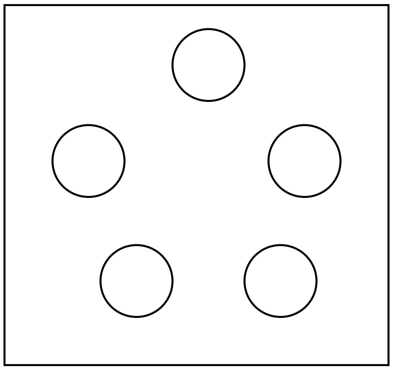
Instructions to participants in the *Feature Presence* Condition.
Instruction set 1.

During this experiment, your task will be to monitor the computer screen for the presence of a circle with a line.

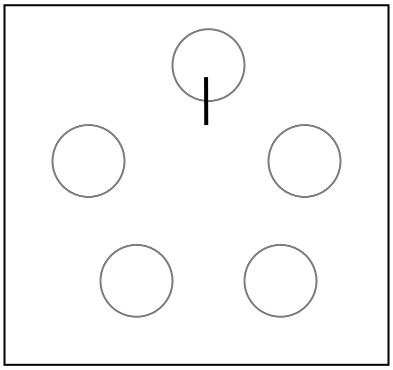
Throughout the task, you will see sets of 5 circles in the middle of the screen. They will flash very fast so it is important to pay close attention! Normally, the five circles will be empty which requires no response from you. Occasionally, however, one of the 5 circles will contain a line. Hence, your job is to respond by pressing the spacebar whenever one of the circles contains a line.

But you should not press the spacebar indiscriminately – respond only when you detect a circle that contains a line

Press the spacebar to view examples of the stimuli



Neutral Signal



Critical Signal
Line can appear in any circle

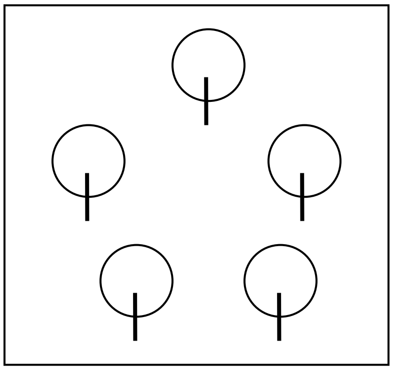
Instructions to participants in the *Feature Absence* Condition.
Instruction set 1.

During this experiment, your task will be to monitor the computer screen for the absence of a circle with a line.

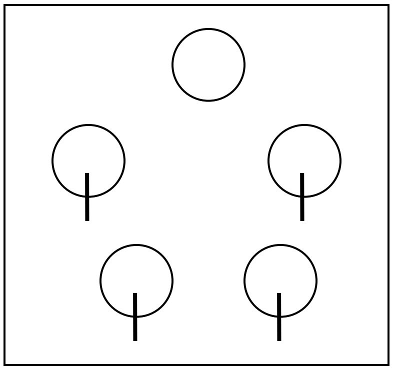
Throughout the task, you will see sets of 5 circles in the middle of the screen. They will flash very fast so it is important to pay close attention! Normally, the five circles will contain a line which requires no response from you. Occasionally, however, one of the 5 circles will be empty. Hence, your job is to respond by pressing the spacebar whenever one of the circles does not contains a line.

But you should not press the spacebar indiscriminately – respond only when you detect a circle that does not contain a line.

Press the spacebar to view examples of the stimuli



Neutral Signal



Critical Signal
Line can be absence in
any circle

Instructions to participants in the *Feature Presence* Condition.
Instruction set 2.

A short practice session will now follow in order to acquaint you with the display and what is required from you. During this practice you will receive feedback about your performance.

- The computer will say "Hit" if you correctly press the spacebar when there is a circle that contains a line in the display.
- The computer will say "Miss" if you accidentally fail to press the spacebar when there is a circle that contains a line in the display.
- If you press the spacebar and do not hear a responds that means you made a False Alarm.
- If you do not hear a responds and you did not press the spacebar that means you made a Correct Decision not to respond.

Do not press the spacebar indiscriminately, respond only when there is a circle that contains a line in the display

Also, do not respond when the stimuli are on the screen- wait until they have flashed off the screen before pressing the spacebar

Instructions to participants in the *Feature Absence* Condition.
Instruction set 2.

A short practice session will now follow in order to acquaint you with the display and what is required from you. During this practice you will receive feedback about your performance.

- The computer will say "Hit" if you correctly press the spacebar when there is a circle that does not contains a line in the display.
- The computer will say "Miss" if you accidentally fail to press the spacebar when there is a circle that does not contains a line in the display.
- If you press the spacebar and do not hear a responds that means you made a False Alarm.
- If you do not hear a responds and you did not press the spacebar that means you made a Correct Decision not to respond.

Do not press the spacebar indiscriminately, respond only when there is a circle that does not contains a line in the display

Also, do not respond when the stimuli are on the screen- wait until they have flashed off the screen before pressing the spacebar

Instructions to participants in the *Feature Presence* Condition.
Instruction set 3.

Another practice session similar to the first one will now follow. Once again, when all of the circles in the display do not contain a line, you do not need to make a response at all. When one of the circles in the display does contain a line, respond by pressing the spacebar. Also, do not press the spacebar indiscriminately, respond only when one of the circles in the display contains a line.

Remember also the importance of not responding while the display is on the screen; wait until after it has flashed off the screen before pressing the spacebar.

Instructions to participants in the *Feature Absence* Condition.
Instruction set 3.

Another practice session similar to the first one will now follow. Once again, when all of the circles in the display contain a line, you do not need to make a response at all. When one of the circles in the display does not contain a line, respond by pressing the spacebar. Also, do not press the spacebar indiscriminately, respond only when one of the circles in the display does not contains a line.

Remember also the importance of not responding while the display is on the screen; wait until after it has flashed off the screen before pressing the spacebar.

Instructions to participants in the *Feature Presence* Condition.
Instruction set 4.

The Experimental session will now follow. During the experiment all of the stimuli that you will see will be the same as those during the practice sessions. However, there will be no feedback during this session

Remember, whenever you see a circle that contains a line, you need to respond by pressing the spacebar. Also, just as before you should not press the spacebar indiscriminately, respond only when you see a circle that contains a line

You may press the spacebar to begin as soon as the experimenter leaves the room

Instructions to participants in the *Feature Absence* Condition.
Instruction set 4.

The Experimental session will now follow. During the experiment all of the stimuli that you will see will be the same as those during the practice sessions. However, there will be no feedback during this session

Remember, whenever you see a circle that does not contain a line, you need to respond by pressing the spacebar. Also, just as before you should not press the spacebar indiscriminately, respond only when you see a circle that does not contains a line

You may press the spacebar to begin as soon as the experimenter leaves the room

Appendix D:

Informed Consent Form

University of Cincinnati
Consent to Participate in the Research Study
Department of Psychology
Victor Finomore: Phone: 513-255-0055, E-Mail: Finomovs@email.uc.edu

Title of the Study: Effects of Feature Presence/Absence and Event Asynchrony on Vigilance Performance and Workload

Introduction: Before agreeing to participate in this study, it is important that the following explanation of the proposed procedures be read and understood. It describes the purpose, procedures, risk, and benefits of the study. It also describes the rights you have as a participant in this study, which includes your right to withdrawal from the study at any time. It is important to understand that no guarantee or assurance can be made as to the results of the study.

Purpose and Procedure: The purpose of this investigation is to study the effects of different levels of task demands have on vigilance performance, perceived mental workload, and stress.

During the session, you will be seated alone in a quiet room. Your job will be to monitor the computer screen in front of you and to respond to particular changes that occur (these will be explained to you in the instructions given by the researcher).

You will also be asked to complete one questionnaire before the task and three questionnaires after the task. The first one asks several questions asking how strongly you agree with different statements that relate to your current mood, motivations, thinking style, thinking content, and other general questions about yourself. Two of the post-task questionnaires ask you to give your rating of how demanding the task was related to specific dimensions. The final questionnaire is similar to the first pre-task questionnaire.

The entire session will take one hour and twenty minutes and you will receive two hours of credit for your participation. You will be one of approximately 80 participants taking part in this study.

Risk/Discomforts: There are no risks involved however, some participants have reported short-term eyestrain and fatigue, which are similar to what a person might experience when working in front of a computer for an extended period.

Benefits: If you are participating in this study to satisfy a requirement for an introductory psychology class, participation will earn 2 hours of experimental credit towards the completion of this course (you will earn one hour of experimental credit for each hour of participation). All other requirements pertaining to the completion of this course must be discussed with the instructor. The task you will be performing is analogous to those that confront air-traffic controllers, airplane pilots, airport baggage inspectors, people engaged in nuclear power plants regulation, and industrial quality control inspectors, and physicians monitoring patient data during surgery. You may receive no direct benefit from you participation in this study, but you participation will provide information about how various factors can influence performance in these types of tasks. This information may have potentially important implications for worker comfort and performance as well as the safety of the public.

Confidentiality: Every effort will be made to maintain the confidentiality of your study records. Agents of the University of Cincinnati will be allowed to inspect sections of the research records related to this study. The data from the study may be published; however, you will not be identified by name or any other traceable information. All data will be coded with an identification number to ensure confidentiality. There are no identifying questions in the experiment that can lead anyone to link the test packet back to you. The records are kept in a secure location and computer, which is only assessable to the research team

Right to Refuse or Withdraw: Your participation in this study is voluntary. You may refuse to participate, or may discontinue your participation at any time without forfeiting your participation credit. The investigator has the right to dismiss you from the study at any time. Your withdrawal from the study may be for reasons related solely to you (not following study-related directions from the investigator, etc.) or because the entire study has been terminated.

Offer to Answer Questions: if you have any other questions about this study, you may call Victor Finomore at 513-255-0055 or Dr. Joel Warm at 513-556-5533. If you have any questions about your rights as a research participant, you may call Dr. Margaret Miller, Chair of the Institutional Review Board- Social and Behavioral Sciences, at 513-558-5784.

Legal Rights: Nothing in this consent form waives any legal right you may have nor does it release the investigator, the sponsor, the institution, or its agents from liability or negligence.

I HAVE READ THE INFORMATION PROVIDED ABOVE. I VOLUNTARILY AGREE TO PARTICIPATE IN THIS STUDY. I MAY REQUEST TO RECEIVE A COPY OF THIS CONSENT FORM FOR MY INFORMATION.

________________________________ Date:____________

Participant (Print)

____

Participant (Signature)

________________________________ Date:____________

Investigator (Signature)

This study has been reviewed and approved by the Committee on Human Research of the University of Cincinnati.

Appendix E:

ANOVA Tables

Table E1.

Analysis of Variance for correct detection scores (Arcsin transformed)

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	29.590	43.197	<.001
Event Schedule (B)	1	-	0.422	0.616	0.435
A x B	1	-	0.210	0.307	0.582
Error	76	-	0.685		
Within Subjects					
Periods (C)	3	2.598	1.463	8.556	<.001
A x C	3	2.598	0.119	0.696	0.557
B x C	3	2.598	0.016	0.094	0.962
A x B x C	3	2.598	0.116	0.678	0.566
Error	228	197.448	0.171		

Note: df_{adj} = degrees of freedom obtained when Box's ε was used to correct for violations of sphericity. Box's ε = .866

Table E2.

Analysis of Variance for false alarm scores (Arcsin transformed)

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	2.401	17.399	<.001
Event Schedule (B)	1	-	0.006	0.043	0.837
A x B	1	-	0.073	0.529	0.470
Error	76	-	0.138		
Within Subjects					
Periods (C)	3	2.541	0.023	2.091	0.094
A x C	3	2.541	0.026	2.364	0.068
B x C	3	2.541	0.006	0.545	0.644
A x B x C	3	2.541	0.019	1.727	0.160
Error	228	193.116	0.011		

Note: df_{adj} = degrees of freedom obtained when Box's ε was used to correct for violations of sphericity. Box's ε = .847

Table E3.

Analysis of Variance for NASA-TLX global workload scores

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	6771.200	27.020	<.001
Event Schedule (B)	1	-	45.000	0.180	0.673
A x B	1	-	273.800	1.093	0.299
Error	76	-	250.603		

Table E4.

Analysis of Variance for NASA-TLX subscale scores

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	293764.000	24.968	<.001
Event Schedule (B)	1	-	5776.000	0.491	0.669
A x B	1	-	2601.000	0.221	0.298
Error	76	-	11765.461		
Within Subjects					
Subscales (C)	4	3.128	161756.187	9.829	<.001
A x C	4	3.128	27444.313	1.668	0.082
B x C	4	3.128	23008.812	1.398	0.253
A x B x C	4	3.128	18862.563	1.146	0.345
Error	304	237.728	16457.278		

Note: df_{adj} = degrees of freedom obtained when Box's ε was used to correct for violations of sphericity. Box's ε = .782

Table E5.

Analysis of Variance for MRQ global workload scores

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	21.012	0.178	0.674
Event Schedule (B)	1	-	108.112	0.918	0.341
A x B	1	-	391.612	3.326	0.072
Error	76	-	117.753		

Table E6.

Analysis of Variance for modified MRQ global workload scores

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	6.050	0.150	0.699
Event Schedule (B)	1	-	54.450	1.352	0.249
A x B	1	-	72.200	1.793	0.185
Error	76	-	40.278		

Table E7.

Analysis of Variance for the modified MRQ resource dimension scores

Source	df	df_{adj}	MS	F	p
Between Subjects					
Task Type (A)	1	-	0.756	0.150	0.699
Event Schedule (B)	1	-	6.806	1.352	0.249
A x B	1	-	9.025	1.792	0.185
Error	76	-	5.035		
Within Subjects					
Subscales (C)	7	5.796	22.264	19.950	<.001
A x C	7	5.796	1.253	1.123	0.347
B x C	7	5.796	1.417	1.270	0.263
A x B x C	7	5.796	1.821	1.632	0.124
Error	532	440.496	1.116		

Note: df_{adj} = degrees of freedom obtained when Box's ε was used to correct for violations of sphericity. Box's ε = .828