AN ABSTRACT OF THE DISSERTATION OF

Shannon P. Gyles for the degree of Doctor of Philosophy in Psychology presented on June 8, 2022.

Title: Mechanisms of Vigilance Loss in Sensory and Cognitive Tasks

Abstract approved:

---

Jason S. McCarley

When observers monitor for infrequent signals for extended durations, they generally experience a decline in detections over time. This decline is termed the *vigilance decrement*. Current theories of vigilance attribute the decrement to three potential mechanisms: conservative shifts in response bias, losses of sensitivity, and an increased rate of attentional lapses over time. Understanding which mechanisms contribute to the losses of vigilance is necessary to mitigate the decrement in applied settings. Unfortunately, much of the existing literature examining vigilance performance relies on measures that are not suited for distinguishing between all three proposed mechanisms. Using novel methods of analysis, the present project examined the extent to which bias shifts, sensitivity losses, and attentional lapses contributed to the vigilance decrement across a range of vigilance tasks. Study 1 (Chapter 3) used psychometric curves to analyze changes in response bias, sensitivity, and lapse rate in two online vigilance tasks. Data showed that the decrement was largely driven by an attentional lapses and conservative shifts in bias over time, with inconclusive evidence for a sensitivity loss. Study 2 (Chapter 4) presents a generative process model to simulate cognitive mechanisms directly and tests the adequacy of the model by reanalyzing data previously fitted with psychometric curves. Results provide converging evidence that the decrement was driven by attentional lapses and shifts in bias. Study 3 (Chapter 5) uses the generative model to assess vigilance performance within a cognitive vigilance task. Vigilance was relatively stable in the cognitive task and data gave strong evidence that the decrement, albeit small, was driven by an increase in attentional lapses. Together, findings provide strong evidence that vigilance decrement is driven by attentional lapses, followed by conservative shifts in bias. Relatively weak evidence for sensitivity loss. Suggests interventions that target lapses and response criteria most effective for minimizing vigilance losses in applied settings.

Mechanisms of Vigilance Loss in Sensory and Cognitive Tasks

by
Shannon P. Gyles

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 8, 2022
Commencement June 2022

Doctor of Philosophy dissertation of Shannon P. Gyles presented on June 8, 2022.

APPROVED:

Major Professor, representing Psychology

Director of the School of Psychological Science

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Shannon Gyles, Author

TABLE OF CONTENTS

LIST OF FIGURES

# Chapter 1: General Introduction

Discriminating rare signals from background noise for long durations requires *vigilance*—a state of readiness to perceive and respond to stimuli (N. H. Mackworth, 1948). Maintaining vigilance is a fundamental part of many routine and safety-critical activities, from sonar monitoring and air-traffic control to quality control and long-distance driving. Though these tasks vary in complexity and consequence, each can be conceptualized as a discrimination between two states of the world: one in which a signal is present among noise, and one in which a signal is absent. As time on watch progresses, monitors often fail to notice signals they are otherwise capable of detecting. For example, early researchers noted that inspection workers (Wyatt & Langdon, 1932), ship lookouts (Ditchburn, 1943), and radar operators (Lindsley & Anderson, 1944) detected progressively fewer targets throughout their shifts. This decline in detections over time is called the *vigilance decrement*.

Broadly, the objective of the current thesis is to identify the mechanisms that contribute to the vigilance decrement. Although this is not a new endeavor, previous work has used methods unsuitable for distinguishing between criterion shifts, sensitivity losses, and attentional lapses, leaving the question of what causes vigilance decrements largely unresolved. An additional point of uncertainty is the nature of performance differences between *sensory* vigilance tasks, in which observers judge perceptual features of stimuli, and *cognitive* vigilance tasks, in which observers judge the meaning or value of symbolic stimuli (Deaton & Parasuraman, 1993).

The studies included in this dissertation aim to address these issues, furthering our understanding of the factors that make human monitors vulnerable to declines in detection performance over time. This understanding is necessary to effectively mitigate the occurrence of vigilance decrements in applied tasks that require sustained attention, including security monitoring, baggage screening, driving, medical screening, and air traffic control, since different theories of vigilance loss suggest conflicting interventions. For example, the effects of overly conservative response criteria might be reduced by adjusting training procedures, increasing the base rate of signals (e.g., as in threat image projection used in airport baggage screening), or payoffs for hits and false alarms to encourage monitors to adopt more liberal response criteria. Alternatively, sensitivity losses might be mitigated by minimizing task load—if attentional resources are depleted—or increasing task engagement—if attention is reallocated toward task unrelated thoughts. If vigilance decrements are driven by attentional lapses, methods of monitoring and recapturing attention may prove fruitful.

Before presenting empirical work examining mechanisms of vigilance loss, Chapter 2 provides a review of existing literature. The review describes patterns of vigilance performance, classification of vigilance tasks, measurement of the vigilance decrement, and theories of vigilance loss. Chapter 3 presents two studies that examine changes in three potential mechanisms of vigilance loss—bias shifts, sensitivity losses, and attentional lapses—over time. Study 1 used psychometric curves to analyze the extent to which each mechanism contributes to the vigilance decrement in an online, sensory vigilance task. Study 2 is a close, pre-registered replication.

Chapter 4 presents an alternative, process modeling approach for analyzing mechanisms of vigilance loss. We demonstrate the utility of the model by reanalyzing data from the pre-registered online task from Chapter 3, Study 2, as well as data from a pre-registered lab-based sensory task. Chapter 5 uses the cognitive process model tested in Chapter 4 to examine the extent to which changes in response bias, sensitivity, and attentional lapses contribute to the vigilance decrement in a cognitive vigilance task. Chapter 6 presents a general discussion of findings.

# Chapter 2: Literature Review

The systematic study of vigilance was prompted by the Royal Air Force's need during World War II to determine the length of time over which radar operators could accurately detect enemy submarines. To simulate the key features of a watch-keeping task, N. H. Mackworth (1948) devised the clock test—a task in which operators were to observe a ticking clock hand and detect instances in which it made a jump of twice the usual size. Signals were infrequent, difficult to perceive, and interspersed among frequent noise stimuli. Mackworth observed that the incidence of missed signals was relatively low during the first thirty minutes of the task but increased significantly after that, remaining high for the rest of the watch period.

In the two decades following Mackworth's clock test, vigilance was assessed across a range of laboratory-based watch-keeping tasks. Two key issues emerged. First, vigilance performance was inconsistent across tasks. Secondly, no clear explanation of the vigilance decrement was forthcoming.

## Inconsistent Performance Across Vigilance Tasks

A glance at the results of just about any vigilance study will reveal large individual differences in measures of performance, including correct detections (Ware et al., 1962) and reaction times (McCormack, 1959). Performance indices vary with characteristics such as the monitor's age (Parasuraman & Giambra, 1991), cognitive ability (Shaw et al., 2010), working memory (Caggiano & Parasuraman, 2004), and task unrelated images and thoughts (Grodsky & Giambra, 1990). Researchers have also examined individual differences in the slope of decline over time (Bakan, 1955;

N. H. Mackworth, 1950), the magnitude of practice and fatigue effects, and the extent to which noise harms performance (Broadbent, 1954).

Typically, individual differences in performance are consistent within tasks. For instance, Jenkins (1958) compared observers' performance on a visual vigilance task completed in the morning to performance on the same task completed in the afternoon, finding high correlations in the percentage of signals detected from time 1 to time 2. Ware et al. (1961) found similar results when they examined the reliability of monitors' performance in an auditory vigilance task over five successive days. The observed consistency of performance within tasks presumably reflects stable individual differences in the ability to maintain vigilance over prolonged watch periods (Jenkins, 1958). This prompted speculation about the possibility of a common vigilance factor (H. J. Jerison & Wing, 1961; Tyler et al., 1972).

Performance across tasks is less consistent. Buckner and collaborators (1960) assessed the performance of 54 sonar operators, twice a day, four days a week, for four weeks, finding that correlations for detection rate, detection latency, and error rates were high across sessions, but low across visual and auditory modalities. Similarly, McGrath (1961) assessed the reliability of performance measures across 16 visual and 16 auditory vigilance tasks, reporting coefficients of .89 and .72 for the percentage of signals detected within visual and auditory tasks, respectively. However, correlations were low when comparing performance on tasks across modalities. Within modalities, detection rates are also highly correlated across tasks that match type and difficulty of signal discriminations (Baker, 1963a; Buckner & McGrath, 1963; Gruber 1964; McGrath et al., 1960; McGrath, 1961; Pope &

McKechnie, 1963), suggesting that these factors may be important determinants of vigilance performance (Parasuraman & Davies, 1977).

Levine et al. (1973) identified two primary perceptual abilities—perceptual speed and flexibility of closure—by which monitoring tasks could be classified. Perceptual speed describes the ability to rapidly compare sensory stimuli for identity or similarity, as when detecting long flashes of light among a series of shorter flashes (Eason et al., 1965). Flexibility of closure describes the ability to detect a previously specified target among distractors (Baddeley & Colquhoun, 1969). Levine and collaborators analyzed the effects of signal rate, sensory mode, and knowledge of results on vigilance performance as a function of ability category, revealing that the shape of the decrement function was dependent on the abilities involved. At low (<1/min) and moderate (1-2/min) signal rates, detections declined gradually over time in perceptual speed tasks but declined sharply then plateaued in flexibility of closure tasks.

Building on the abilities classification above (Levine et al., 1973), Parasuraman (1976) examined the consistency of individual differences across four visual vigilance tasks. Tasks were categorized according to whether they required perceptual speed or flexibility of closure, and whether the signal to be discriminated was a decrease in the intensity of a flashing light or a decrease in the separation between two lines. Three groups of participants (n = 10) completed selected pairs of tasks (see Figure 1).

Note that Parasuraman (1979) later adopted the terms successive and simultaneous to describe what he considered to be the key difference between tasks involving perceptual speed versus flexibility of closure—the type of signal discrimination. Successive discriminations, common in "perceptual speed" tasks, require observers to detect a change in some feature (e.g., flash duration) of a stimulus when a standard stimulus is not present for comparison. In contrast, simultaneous discriminations, featured in tasks requiring flexibility of closure, require observers to detect a target among distractors when signal and noise stimuli are presented simultaneously (Parasuraman, 1979).

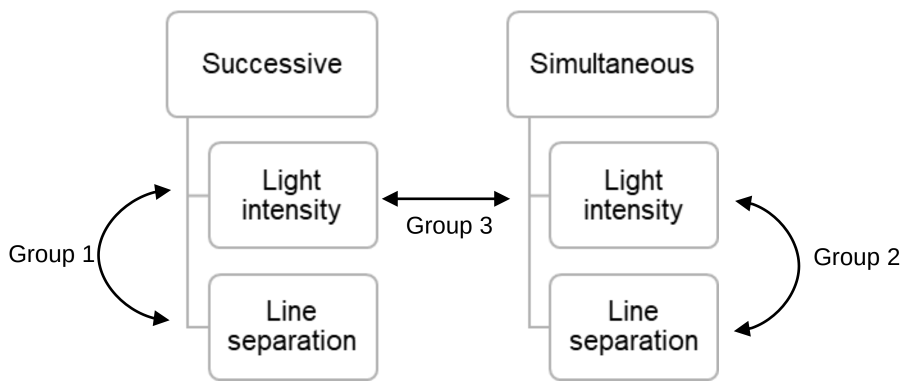


*Figure 1.* Task pairs performed by participants in Parasuraman (1976).

Each task lasted 45 minutes, presented stimuli at a rate of 15 events/minute, and presented signals at a rate of 1/minute. Performance was analyzed as a function of 15-minute time blocks, revealing that correct detections decreased significantly throughout the course of the task in each group. Three measures of performance—correct detections, false alarms, and sensitivity—were highly correlated across tasks within groups 1 and 2, suggesting that performance was consistent across tasks matched on discrimination type, irrespective of signal type. These measures were uncorrelated for group 3's tasks, which matched signal type (a decrease in light intensity) but differed with respect to discrimination type (simultaneous vs. successive).

In a follow-up experiment, another three groups of subjects (n = 10 per group) completed selected pairs of tasks across visual and auditory modalities (Parasuraman & Davies, 1977). Performance was correlated for tasks matched on discrimination type but not modality, but uncorrelated for tasks that differed in both discrimination type and modality. Together, the two experiments indicate that that vigilance performance is not entirely task-specific, as inconsistencies in the earlier literature may have suggested, nor is it correlated across all tasks as expected if driven by a common vigilance factor. Instead, performance appears to depend on task characteristics, particularly, whether a task requires the observer to make successive or simultaneous discriminations.

**Signal Detection Analysis of the Vigilance Decrement**

The vigilance decrement is typically indexed by a decline in the percentage of correctly detected signals over time. Early theories of vigilance assumed that the decrement was caused by a decline in observer's perceptual ability to distinguish signals from noise (Frankman & Adams, 1962; J. F. Mackworth, 1968b). The development of signal detection theory (SDT; Green & Swets, 1966; Swets et al., 1961; Tanner & Swets, 1954), however, introduced the possibility that perceptual losses alone might not explain vigilance failures.

SDT models an observer's ability to discriminate between signal and noise events in the presence of noise or uncertainty.  The most basic form of signal detection task, the yes-no task, requires the observer to report each trial whether a signal is present or not. The combination of two possible states of the world (signal present/signal absent) and two possible responses (signal/noise) produces four stimulus-response events, shown in Figure 2. When the true state of the world is signal, signal responses are correct and are called hits. Conversely, when the true state of the world is noise, signal responses are incorrect and are called false alarms (Green & Swets, 1966). The proportions of signal trials judged as signal and noise are termed hit rate (HR) and false-alarm rate (FAR), respectively, and are used to calculate measures of sensitivity and bias. The miss rate and correct rejection rate are the complements of hit and false alarm rates, respectively, making the latter sufficient to fully describe signal detection performance.

| | Signal present | Signal absent |
|---|---|---|
| Respond signal | Hit | False alarm |
| Respond noise | Miss | Correct rejection |

*Figure 2.* Stimulus-response events for signal detection judgments.

The SDT model assumes that upon observing a stimulus, the observer encodes evidence to decide whether the stimulus represents signal or noise, and represents the evidence as a univariate decision variable (T. D. Wickens, 2002). Signal and noise events correspond to separate distributions of evidence values. On average, signals produce larger evidence values than noise. However, unless the distinction between the two alternatives is unambiguous, the evidence distributions will overlap, and noise trials will sometimes produce larger evidence values than signal trials. The decision-maker's ability to distinguish between signal and noise, as determined by the overlap between the distributions, is termed sensitivity. Assuming that signal and noise distributions are normal and of equal variance, sensitivity can be measured with the statistic d',

$$d' = z(HR) - z(FAR)$$

where HR = p(signal response | signal) and FAR = p(signal response | noise).

Following encoding, SDT assumes that observers arrive at discrete judgments by comparing the decision value to a criterion value (Green & Swets, 1966). Evidence values that exceed the criterion are transformed into signal judgments and values that fall below the criterion are transformed into noise judgments. The position of the criterion reflects the likelihood of making a particular response. For instance, an unbiased criterion indicates that the observer has no preference for signal or noise judgments, and the proportions of misses and false alarms are equal. This is ideal when base rates of signal and noise are equal, and the payoff schedule for signal and noise responses are symmetrical. Alternatively, a conservative criterion is biased towards noise judgments, which may be ideal when signals are rare or false alarms are costly, and a liberal criterion is biased towards signal judgments, which may be ideal when signals are frequent, or the value of a hit outweighs the costs of a false alarm. The distance from the observer's criterion to the unbiased position provides a measure of response bias, denoted c (T. D. Wickens, 2002).

Applying signal detection analyses to vigilance data allows researchers to attribute changes in observers' hits and false alarms to underlying changes in sensitivity and response bias. Figure 3 illustrates how changes in an observer's sensitivity and bias affect their hit and false alarm rates. In general, as the observer's criterion shifts conservatively (from the left to right panels), both hits and false alarms decrease, and as sensitivity decreases (from the top to bottom panels), hits decrease, and false alarms increase. Therefore, changes in sensitivity and bias are inferred from the pattern of tradeoffs between hits and false alarms.

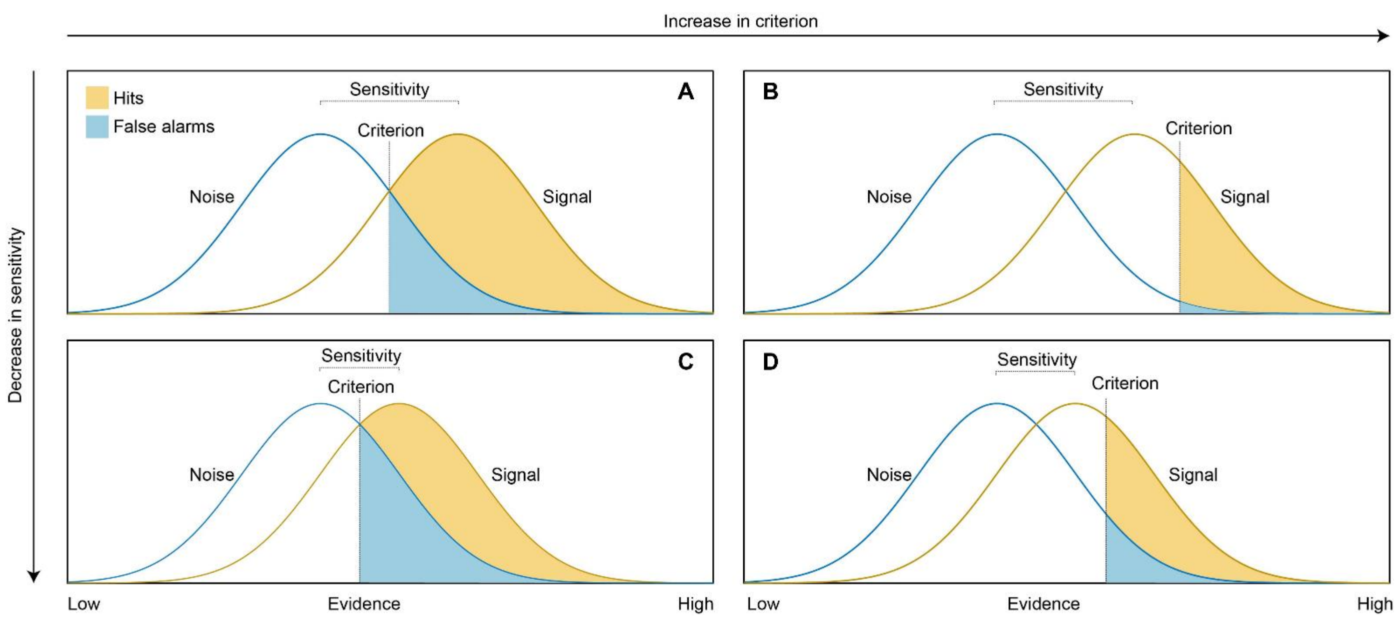


*Figure 3*. How changes in bias and sensitivity affect hits and false alarms. In Panel A, the observer's sensitivity and unbiased criterion produce a roughly 2:1 ratio of hits to false alarms. In Panel B, the observer adopts a more conservative criterion, resulting in a reduced HR and a reduced FAR. In Panel C, the observer retains the unbiased criterion, but has worse sensitivity. Panel D depicts a concurrent criterion increase

and sensitivity decrease, resulting in fewer hits and fewer false alarms than in Panel A.

Following the adoption of SDT by vigilance researchers, many studies conducted during the 1960's reported that observers adopted progressively more conservative criteria over time on task (e.g., Broadbent & Gregory, 1963; Colquhoun & Baddeley, 1967; Loeb & Binford, 1964). Reviews concluded that the vigilance decrement primarily results from changes of bias, rather than declines in sensitivity (Broadbent, 1971; Mackworth, 1970; Swets & Kristofferson, 1970).

Some studies, however, did report decreases in sensitivity concurrent with changes in bias. More specifically, sensitivity decrements occurred in tasks that used visual signals and high event rates, requiring continuous attention to the display. For example, J.F. Mackworth (1968) found declines in *d'* in tests requiring continuous observation, but constant d' in discrete tests. Loeb and Binford (1968) and Smith and Barany (1970) found a linear decreases in *d'* as event rate increased. Interestingly, changes in *d'* were not observed for auditory signals, regardless of event rate, suggesting that performance might also depend on sensory modality.

To further evaluate the effects of event rate and sensory modality on sensitivity and bias, Swets (1977) analyzed the results of twelve studies published between 1969 and 1971. All twelve experiments showed an increasingly strict criterion over time, (Colquhoun & Edwards, 1970; Deaton et al., 1971; Guralnick, 1972; Hatfield & Soderquist, 1970; Johnston et al., 1969; Loeb & Binford, 1970; McCann, 1969; Milosêvić, 1969; Thurmond et al., 1970; Williges; 1969, 1971, 1973). Four of the studies showed concurrent decreases in sensitivity (M. Deaton et al., 1971; Guralnick, 19720101; Hatfield & Soderquist, 1970; Williges, 1973); the other eight studies found no change in sensitivity. In contrast to previous findings (Loeb & Binford, 1968; J. F. Mackworth, 1968a), sensitivity decrements in Swets' analysis were not exclusive to visual tasks and were not ubiquitous among tasks with "high" event rates. Swets suggested that the effect of event rate on sensitivity may vary from task to task. A later study by Guralnick (1972) found an interaction between event rate and signal discriminability, whereby increasing the event rate reduced sensitivity, but only when signals were not readily detectable.

The application of signal detection theory to vigilance data provided researchers with a method of assessing vigilance performance that isolates sensitivity from bias. Analyses using this method, though sometimes yielding inconsistent results, showed that declines in correct detections were mostly caused by observers adopting progressively more conservative response criteria over time (Broadbent, 1971; Parasuraman, 1976; Swets, 1973; Swets & Kristofferson, 1970). Data also suggested that sensitivity decrements were more likely to occur in tasks with higher event rates, but the relationship between event rate and sensitivity was inconsistent.

## Emergence of a Vigilance Taxonomy

Seeking to more concretely identify the task characteristics that produce sensitivity decrements, Parasuraman and Davies (1977) evaluated 27 vigilance studies reporting signal detection analyses. As in the experiments on intra- and inter-modal correlations (Parasuraman, 1976; Parasuraman & Davies, 1977), they classified tasks according to type of discrimination (successive, simultaneous) and sense modality (auditory, visual). They also classified tasks according to source complexity (single source, multi-source) and stimulus event rate (low, high). Based on Loeb and Binford's (1968) finding that sensitivity declined at an event rate of 24/min, Parasuraman and Davies defined event rates ≤24/min as low and >24/min as high. Parasuraman and Davies observed sensitivity decrements in 13 of the 27 studies they examined. The task classification analysis revealed that 1) individual differences were only consistent for tasks requiring the same ability, and 2) sensitivity decrements only occurred in tasks that combined successive discrimination with high event rates. Parasuraman and Davies presented these results as a preliminary taxonomy of vigilance, in which sensitivity decrements are a function of discrimination type and event rate.

To further examine the effects of task dimensions on sensitivity losses, Parasuraman (1979) manipulated discrimination type and event rate in two experiments. In the first experiment, forty participants performed a 45-minute auditory vigilance task under one of four conditions produced by the combination of two discrimination types (successive, simultaneous) and two event rates (15 events/min, 30 events/min). Again, a vigilance decrement was observed in all conditions, but sensitivity declined only for the successive-high event rate task.

A second experiment compared performance across a successive discrimination task and two simultaneous discrimination tasks that differed in noise levels. Sensitivity declined in the successive task, but not in either of the simultaneous tasks, supporting the hypothesis that event rate and memory load, rather than noise, drive sensitivity decrements. These findings provided support for the proposed taxonomy of vigilance performance (Parasuraman & Davies, 1977), prompting Parasuraman (1979) to assert that vigilance decrements are primarily driven by declines in sensitivity, with little change in bias, when tasks require successive discriminations and employ a high event rate, but are driven only by increases in response criterion under alternative conditions.

Taken at face value, these experiments seem to provide evidence that sensitivity decrements occur only in successive, high event rate tasks. Although the extremely small sample sizes and high familywise error rates produced by multiple comparisons might weaken confidence in these early studies, many aspects of the taxonomy are supported by subsequent work.

Parasuraman (1979) attributed the observed sensitivity decrements to the increased information processing demands—specifically, memory and temporal demands—placed on the observer in high event rate, successive discrimination tasks. This explanation can be understood within the context of resource models of attention, which conceptualize attention as the allocation of information processing resources to a task (Kahneman, 1973; Norman & Bobrow, 1975; Schneider & Shiffrin, 1977; C. D. Wickens, 2002). Information processing resources (e.g., processing effort, memory capacity, and communication channels; Norman & Bobrow, 1975) are drawn from a limited capacity ‘pool’ of resources that is taxed by cognitive activity.

A task’s demand on resources depends on parameters such as the sensory quality of stimuli, response complexity, and the subject’s level of skill (Navon & Gopher, 1979). Subjects allocate resources to a task to achieve the desired level of performance, with performance suffering if the demand on resources exceeds the supply (Kahneman, 1973; Kantowitz & Knight, 1974; Moray, 1967; Norman & Bobrow, 1975). Sustained attention tasks, in general, are said to place particularly high demands on information processing resources; observers need not only pay attention for prolonged periods, but must also put forth effort to maintain attention despite low signal rates (De Waard, 2002; Hancock, 1989).

Parasuraman’s resource depletion account of the vigilance decrement proposed that tasks with high event rates and successive discriminations are more demanding than tasks with low event rates and simultaneous discriminations, and therefore tax resources to a greater extent. When the limited supply of cognitive resources is used faster than it is replenished, fewer resources are available to the vigilance task, and the observer’s ability to discriminate signal and noise declines over time. Tasks with high event rates may be more demanding than those with low event rates due to the higher rate of processing and time pressure they impose. Research shows that time constraints increase perceived mental workload (Hertzum & Holmegaard, 2013), which reduces the residual capacity available to the task (Block et al., 2010; Hancock & Weaver, 2005). Successive discrimination tasks may be more demanding than simultaneous discrimination tasks due to the memory load they impose. When stimulus events are presented successively, the observer must hold a representation of the standards stimulus in memory to compare to the current stimulus each trial (Caggiano & Parasuraman, 2004). Maintaining information in memory requires attention (Engle, 2002, 2018) and has been shown to be capacity-demanding (Baddeley & Hitch, 1974).

In contrast, tasks with simultaneous discriminations or low event rates are less demanding and do not overburden resources. Decreases in detection rates are therefore attributed entirely to conservative changes in response bias (Broadbent & Gregory, 1963; Broadbent, 1971, 1971; Egan et al., 1961; Hatfield & Loeb, 1968; Loeb & Binford, 1964; Parasuraman, 1976; Swets, 1973; Swets & Kristofferson, 1970). The optimal criterion placement depends on base rates of signals and noise events (Jerison et al., 1965; Williges, 1971) and, to a lesser extent, payoffs for hits and false alarms (Levine, 1966; Williges, 1971). Assuming symmetrical payoffs, a criterion biased toward the more frequent stimulus event will maximize payoffs. Thus, in vigilance tasks containing rare signals, this reduced willingness to report 'signal' over time represents a trend toward optimal decision behavior (Broadbent & Gregory, 1963; Broadbent & Gregory, 1965; Levine, 1966; Williges, 1969).

Sixteen years after Davies and Parasuraman published their vigilance taxonomy, See et al. (1995) conducted a meta-analysis of 42 studies that specifically examined sensitivity in vigilance tasks. The studies' 138 conditions were classified according to type of discrimination, event rate, whether the tasks were sensory or cognitive, and several additional features that could potentially increase task demands (e.g., signal regularity, spatial uncertainty, and vigil length). The meta-analysis revealed sensitivity decrements with effect sizes greater than 0.2 in 78% of the conditions and confirmed that sensitivity decrements mainly occurred in tasks with successive discriminations and high event rates.

Aside from the studies included in the meta-analysis, most of the research examining discrimination type and event rate has instead focused on their effects on detection rates in general, rather than on changes in sensitivity and bias. For example, researchers have found that successive tasks produce larger vigilance decrements than simultaneous tasks (Dember et al., 1985; Lanzetta et al., 1985; Parasuraman et al., 1984; Parasuraman & Mouloua, 1987), and that the magnitude of the vigilance decrement is proportional to the memory load imposed by a task (Helton & Warm, 2008).

Similarly, high event rate tasks consistently produce larger vigilance decrements than low event rate tasks (Claypoole et al., 2019; Davies & Parasuraman, 1982; Galinsky et al., 1993; Guralnick, 1972; Jerison & Pickett, 1964; Lanzetta et al., 1987; Meuter & Lacherez, 2016; Mouloua & Parasuraman, 1995; Parasuraman, 1979; Parasuraman & Giambra, 1991; Rose et al., 2001; Smith et al., 2002; Warm & Jerison, 1984; Yadav et al., 2015). In fact, although Parasuraman and Davies (1977) originally dichotomized event rate as high or low, later studies have shown that detection rate is inversely linearly related to event rate (Galinsky et al., 1993; Warm & Jerison, 1984).

Together, these findings provide convincing evidence that successive discriminations and high event rates produce greater decrements in vigilance performance than their simultaneous and low event rate counterparts. Further, they show that increases in memory demand and event rate negatively impact performance. These results are consistent with Parasuraman's claim that vigilance tasks produce sensitivity decrements because the increased processing load depletes attentional resources and, as such, they are largely interpreted as evidence for the resource depletion account of the vigilance decrement. The results do not, however, directly support the claim that decrements in these conditions are driven by losses of sensitivity.

Other studies have reported findings inconsistent with the taxonomy altogether. Although Parasuraman and Davies (1977) argued that simultaneous discriminations only produce changes in response criteria, sensitivity decrements have been observed in simultaneous tasks with poor signal discriminability (Dittmar et al., 1985.; Parasuraman, 1985; Parasuraman & Mouloua, 1987; Scerbo et al., 1987; Warm et al., 1987). For example, Nuechterlein et al. (1983) manipulated the degree of stimulus degradation in a simultaneous, high-event rate task (which, according to the taxonomy, should not produce a sensitivity loss). They found that sensitivity declined within 5 minutes in the highly degraded condition, but not in the moderately degraded or undegraded) conditions. A later study (Parasuraman & Mouloua, 1987) found that only successive judgments showed sensitivity decrements when stimulus discriminability was moderate or high, but that successive and simultaneous tasks both produced sensitivity decrements when stimulus discriminability was poor.

The taxonomy also specifies that tasks with low event rates (i.e., fewer than 24/min) only produce changes in criteria, not sensitivity. Yet, researchers have found sensitivity decrements in many tasks with event rates below the taxonomy's cutoff (Beh, 1989; Eilers et al., 1988; Joshi, 1985; Mackworth, 1970; Tomporowski & Simpson, 1990; Williams, 1986; Williges, 1971)

**Revisions to the Taxonomy**

Although the taxonomy is supported by findings of sensitivity losses in successive, high event rate tasks, sensitivity losses in both simultaneous and low event rate tasks suggest that the original taxonomy was incomplete. Researchers have proposed two major revisions, discussed below.

**Overall Task Load.** In response to findings that sensitivity losses sometimes occurred in simultaneous or low event rate tasks, researchers, including Parasuraman himself (Parasuraman & Mouloua, 1987), suggested that overall task demand may be a more important determinant of sensitivity decrements than the specific combination of successive discriminations and high event rates (Lanzetta et al., 1987; Nuechterlein et al., 1983). It is generally accepted now that sensitivity decrements may result from any combination of variables that sufficiently increases the demand for attentional resources (See et al., 1995; Smit et al., 2004). See et al. (1995) suggested that total task demand may be indexed by the average level of sensitivity achieved throughout the vigil. More demanding tasks—those with lower initial sensitivity—suffer more with time on task (Helton & Russell, 2011; See et al., 1995).

**Sensory versus Cognitive Vigilance.** See et al. (1995) also recommended revising the taxonomy to distinguish sensory tasks, in which signals are changes in perceptual features of stimuli (e.g., brightness, size), from cognitive tasks, which require numeric, linguistic, or semantic discriminations. Though Davies and Tune (Davies & Tune, 1969) had suggested that vigilance performance might differ between tasks using sensory (perceptual) and cognitive (alphanumeric) stimuli, research making that comparison was too sparse to warrant including the sensory-cognitive dimension in the taxonomy (Deaton & Parasuraman, 1993). Following the development of the vigilance taxonomy, however, Koelega et al. (1989) further examined the conditions required for sensitivity losses by comparing performance across four successive, high event rate tasks of varying demand, complexity, and stimulus type. The most notable difference in performance was between sensory and cognitive tasks; sensitivity declined in the two sensory tasks, but not for the two cognitive tasks. Similar results were observed in other studies (Loeb et al., 1987; Warm et al., 1984).

The reason for the performance differences between tasks using sensory and cognitive stimuli was unclear. Sensitivity may have remained stable in the cognitive task because the alphanumeric stimuli were familiar and well-learned, making them less demanding than sensory stimuli. Consistent with this idea, Fisk and Schneider (1981) found that variably-mapped signals, requiring effortful processing, produced sensitivity decrements while well-learned pairings, requiring automatic processing, did not.

Alternatively, sensitivity may have remained stable in cognitive tasks not because alphanumeric stimuli were more familiar, but because they required observers to make fundamentally different kinds of discriminations than sensory stimuli (Deaton & Parasuraman, 1993; See et al., 1995; Warm et al., 1984). Deaton and Parasuraman (1993) had participants make sensory and cognitive discriminations of an identical set of numerical stimuli. Correct detections declined over time in the sensory task, which required discriminations of digit size, and remained stable in the cognitive task, which required even/odd discriminations of digit value. This finding suggested that 1) sensory/cognitive differences are driven by factors beyond the physical features of the stimuli, and 2) tasks should be classified as sensory or cognitive on the basis of the discrimination required, rather than the nature of the stimuli.

Performance differences might have been driven by the additional processing step involved in cognitive discriminations (Deaton & Parasuraman, 1993; See et al., 1995). The sensory task involved extracting line length from the digits and comparing them, while the cognitive task involved extracting the name code for the digit, determining the number category to which it belongs (odd/even), and then comparing. However, the cognitive task was not considered more difficult and did not produce lower performance than the sensory task in Koelega et al.'s (1989) study. It is also possible that sensory tasks suffer more with time on task because participants are less able to predict how well they are performing (Deaton & Parasuraman, 1993). Deaton and Parasuraman suggested that the familiarity of cognitive stimuli offers built-in feedback that may motivate monitors to perform better.

Despite receiving relatively little attention in the literature, the sensory-cognitive dimension explained variance over and above the average level of sensitivity in the meta-analysis (See et al., 1995). Overall, sensitivity decrements were larger in sensory tasks than in cognitive tasks, except at very high event rates. However, a recent study (Claypoole et al., 2019) showed that event rate has a similar effect on cognitive and sensory stimuli. The current understanding, based on research summarized in See et al.'s (1995) meta-analysis, is that the presence and magnitude of a sensitivity decrement is a function of the type of discrimination, event rate, average sensitivity, and type of stimuli used in the task (sensory, cognitive). The mechanism by which sensitivity differs in sensory and cognitive tasks remains unspecified.

## Theories of Sensitivity Loss

In recent years, the focus of vigilance research has shifted from identifying the conditions in which sensitivity declines to understanding why it declines. The resource depletion theory attributes sensitivity decrements to information processing resources being overloaded and in turn, depleted. Mind-wandering theories, on the other hand, attribute sensitivity losses to resources being underloaded and in turn, reallocated elsewhere.

### *Resource Depletion*

The resource depletion hypothesis (Parasuraman, 1979) posits that vigilance tasks are effortful and therefore deplete information processing resources over time (Caggiano & Parasuraman, 2004; Parasuraman, 1979; Warm et al., 1996). As resources diminish, less attention can be directed to the task, reducing the monitor's ability to discriminate signal from noise. The resource depletion account of vigilance is consistent with findings that 1) increasing task demands increases the vigilance decrement, 2) vigilance tasks are subjectively effortful, and 3) cerebral metabolic activity declines with time on task.

**Increasing task demands increases vigilance decrement.** If sensitivity declines because task demands deplete resources (Parasuraman et al., 1987), then tasks that place greater demands on resources ought to deplete resources faster and produce a larger vigilance decrement than those that are less demanding (Caggiano &

Parasuraman, 2004). A substantial body of research finds that more demanding tasks produce not only lower overall detection rates but greater vigilance decrements than less demanding tasks (Gluckman et al., 1988; Helton & Russell, 2011, 2013; Helton & Warm, 2008; Smit et al., 2004; Joel S. Warm & Dember, 1998). For example, tasks with poor signal discriminability show steeper decrements over time that those with high signal discriminability, some of which are attributed directly to sensitivity losses (Helton et al., 2004, 2002; Helton & Warm, 2008; Matthews et al., 2000; Temple et al., 2000). MacLean et al. (2009) found that cueing visual attention to the location of a signal resulted in smaller sensitivity decrements than when stimuli required continuous monitoring. These findings demonstrate that more demanding tasks produce greater vigilance decrements, if not greater sensitivity decrements, consistent with the resource depletion hypothesis.

**Vigilance is subjectively effortful.** Vigilance tasks—though relatively straightforward—also appear to impose considerable mental workload on observers (R. A. Grier et al., 2003; Helton & Russell, 2013; Helton & Warm, 2008; Hitchcock et al., 1999; Temple et al., 2000; Warm et al., 1996). Workload describes the effort expended by an operator to achieve a particular level of performance (Hart & Staveland, 1988) and is the product of a task's demands, an operator's skill and response to a task, and the context in which the task is performed. Including a measure of workload is now commonplace in vigilance research, and for several decades, studies have consistently found that vigilance tasks produce moderate to high workload ratings (Dittmar et al., 1993; Finomore, 2006; Grier et al., 2003; Grubb et al., 1995; Helton et al., 2005; Helton & Russell, 2013, 2013; Helton & Warm, 2008; Hitchcock et al., 1999; Hollander et al., 2004; Matthews et al., 2000; Scerbo et al., 1993; Schoenfeld & Scerbo, 1997; Szalma et al., 2004; Temple et al., 2000) driven by high levels of mental demand and frustration (Hancock, 1984; Matthews et al., 2000).

Observers also find vigilance tasks highly stressful, making them feel less energetic, cheerful, motivated, able to concentrate, confident, and self-focused (Szalma et al., 2004). These effects do not seem to result simply from boredom, but from the information processing demands of the vigilance tasks themselves (Alikonis et al., 2002; Hitchcock et al., 1999; Warm et al., 1996).

The high workload and stress associated with vigilance tasks is interpreted as evidence that vigilance depletes resources (Parasuraman and Davies, 1977; Davies and Parasuraman, 1982; Parasuraman, 1984; Parasuraman, Warm and Dember, 1987; Warm and Dember, 1998). An alternative interpretation is that remaining vigilant for rare signals does not deplete resources, but that resource demanding tasks are susceptible to vigilance losses. If it is the vigilance component of these tasks that makes them stressful/taxing, and not other task, operator, or contextual factors, then workload should decrease when signals become more frequent.

Experiments manipulating signal rate have not found evidence that higher signal rates are associated with lower workload (R. A. Grier et al., 2003; Sawyer et al., 2014). At least one study trended in the opposite direction—Matthews (1996) compared low (.1) and high (.35) signal probabilities and found that mean overall workload was higher for the high-probability task. Although the resource depletion account predicts that increasing the signal rate reduces attentional demands, and therefore workload, Matthews interpreted the increased workload as evidence that resources were depleted by the increased need to respond when signals were more frequent. This presents an issue for testing the claim that vigilance itself causes high workload: if reducing attentional demands (by increasing the signal rate) proportionally increases response demands, then changes in workload as a function of signal rate are not diagnostic of resource depletion.

**Cerebral metabolic activity declines with time on task.** Although high workload is not proof that that the vigilance aspect of monitoring tasks depletes resources, the taxing nature of vigilance tasks is also demonstrated by brain imaging studies. Early PET and FMRI studies found that vigilance tasks were associated with increased activity in several brain areas (Deaton & Parasuraman, 1988) but unfortunately, did not examine changes in brain activity over time on task. Newer, less invasive neuroimaging techniques like transcranial doppler sonography (TCD; Aaslid, 1986) allow researchers to monitor participants for extended periods without hindering movement, making them more suitable for assessing changes over time.

During TCD, ultrasound waves are transmitted through the skull and are reflected off blood cells. Metabolic activity in the brain produces by-products such as carbon dioxide, which dilates blood vessels in the area, resulting in an increase in blood flow to the region. By measuring the difference in frequency between outgoing ultrasound signals and the reflected energy, TCD provides a real time measure of cerebral blood flow velocity (CBFV). CBFV reflects metabolic activity and is higher when a person is engaged in mental activity than when they are at rest (Duschek & Schandry, 2003; Helton et al., 2007; Stroobant & Vingerhoets, 2000; Vingerhoets & Stroobant, 1999a, 1999b). Individual differences in CBFV are correlated across tasks and are unrelated to changes in physiological arousal (e.g., pulse rate, heart rate, respiration, and blood pressure (Schnittger et al., 1997; Vingerhoets & Stoobant, 1999a, 1999b).

Several experiments have used TCD to analyze CBFV during vigilance tasks (Joel S. Warm et al., 2008; Joel S. Warm & Parasuraman, 2007). Participants performing vigilance tasks show an increase in CBFV in the right cerebral hemisphere relative to the left hemisphere. No hemispheric differences are observed in participants performing a control task (Helton et al., 2007; Hollander et al., 2003). The absolute blood flow velocity appears to be positively associated with the cognitive demands of the vigilance task (Hitchcock et al., 2003; Warm & Parasuraman, 2007). Additionally, right hemisphere CBFV declines over time for tasks that produce vigilance decrements (Hitchcock et al., 2003; Shaw et al., 2009) but remains stable when vigilance decrements are not observed (Funke et al., 2010). This covariation led proponents of resource depletion theory to suggest that declines in CBFV reflect declines in the availability of information processing resources (Joel S. Warm et al., 2012).

Other studies though, do not find evidence of a decline in CBFV over time. Helton et al. (Helton et al., 2007) and Hollander et al. (Hollander et al., 2003) both found that observers performing an abbreviated (12-minute) vigilance task had higher blood flow velocity in the right than the left cerebral hemisphere, compared to no hemispheric differences in in a control task. But CBFV remained stable over time in both cases, despite declines in correct detections. Both studies also included a measure of cerebral blood oxygen saturation, which has been previously shown to increase with the information processing demands of the task being performed (Punwani et al., 1998; Toronov et al., 2001). As with blood flow velocity, blood oxygenation was higher in the right hemisphere than the left for the vigilance group than the control, but did not decline over time. Helton and Hollander suggested that the abbreviated tasks they used may not have been long enough to allow declines in cerebral vascular dynamics to be observed, despite being long enough to produce a vigilance decrement.

If CBFV and oxygenation reflect the availability of resources, then under the depletion hypothesis, declines in CBFV should parallel declines in sensitivity. The aforementioned studies reveal associations between performing a vigilance task and increased blood flow velocity in the right hemisphere, and between declines in CBFV and correct detections. However, the false alarm rates were too low to allow for signal detection analysis of sensitivity and bias, meaning that observed declines in correct detections cannot be attributed confidently to declines in sensitivity. The exception is a more recent study by Matthews and colleagues (Matthews et al., 2010), who found that declines in CBFV were accompanied declines in *A'*—an alternative measure of sensitivity that is widely believed to be nonparametric (see p. 40 for discussion of the limitations of *A'*).

Collectively, the pattern of results observed in brain imaging studies leaves open the possibility that increased blood flow and oxygenation are not functionally important for maintaining vigilance but byproducts of engaging with the vigilance task. The observed decline in blood flow velocity may simply reflect a decline in resource utilization over time (e.g., if attention is diverted elsewhere), rather than a decline in resource availability. An explanation along these lines would account for the co-variation in brain activity and vigilance decrements without assuming a causal mechanism or loss of sensitivity.

### ***Mind-wandering***

In contrast to the view that vigilance tasks ‘overload’ observers, leaving insufficient resources for the task, mind-wandering accounts posit that the monotonous and understimulating nature of vigilance tasks ‘underloads’ observers, causing them to withdraw attention from the task. Theories of mind-wandering specify not only that attention is withdrawn from the task, but that it is redirected to internally generated task unrelated thought (TUT; Klinger, 1978; McVay & Kane, 2009; Risko et al., 2012; Seli et al., 2016; Thomson et al., 2015, 2016). As in a resource depletion account, a mind-wandering explanation predicts that fewer resources are allocated to the task over time, resulting in poorer quality information processing, and in turn, a decline in sensitivity.

Smallwood and Schooler (Smallwood & Schooler, 2006) argue that mind-wandering is automatically activated by task-unrelated goals that draw attention away from the primary task, but that resources are required to maintain the resulting TUTs. Therefore, in the context of vigilance, TUTs draw attentional resources away from the primary task, resulting in a decline in correct detections over time. This theory has been described as the executive-resource theory of mind-wandering (Feng et al., 2013; Thomson et al., 2015).

Conversely, McVay and Kane (2009, 2012) argue that mind wandering does not require attentional resources. Drawing on Watkins’ (2008) elaborated control theory, they propose a control-failure theory of mind-wandering, in which TUTs are initiated by failures of executive control to combat thoughts that interfere with the task and are maintained automatically. Both theories (Smallwood & Schooler, 2006; McVay and Kane, 2009; 2012) assume that mind-wandering occurs because the task fails to hold the observer's attention and, as such, attention becomes decoupled from the task (i.e., divided between external and internal information; Antrobus, 1968; Barron et al., 2011). Mind-wandering accounts of the vigilance decrement are consistent with findings that 1) increasing task engagement, arousal, and motivation reduces the vigilance decrement, 2) increasing task demands increases the vigilance decrement, and 3) TUTs increase with time on task.

**Increasing Task Engagement and Motivation Reduces the Vigilance Decrement.** Vigilance studies find that task engagement, as measured by the DSSQ (Matthews et al., 2002), typically declines over time (Matthews et al., 2002, 2017; Szalma et al., 2004) and is positively associated with detection performance (Finomore et al., 2009; Matthews et al., 2001; Matthews & Davies, 2001; Neigel et al., 2019). These findings are consistent with both resource depletion and mind-wandering accounts of the vigilance decrement. Under a resource depletion account, declines in task engagement have been interpreted as a byproduct of diminishing attentional resources (Matthews & Davies, 2001; Temple et al., 2000). Meanwhile, the mind-wandering account assumes that disengagement from the task causes attentional lapses, resulting in a decline in detections. This leads to the specific prediction that more engaging tasks will result in fewer attentional lapses and in turn, smaller vigilance decrements than less engaging tasks (Barron et al., 2011; Thomson et al., 2015).

Pop et al. (2012) manipulated task engagement by increasing the response demands of a simulated air traffic control task. Participants in both the "standard" and "engagement" conditions were tasked with monitoring for potential aircraft collisions. Those in the engagement had the additional task of clicking on aircraft as they entered the airspace. Participants performed the task on four consecutive days (1 hr/day). A vigilance decrement was observed for both conditions on Days 1, 2, and 3, but by Day 4, vigilance declined only in the standard condition and remained stable in the engagement condition. This suggests that engaging monitors in the task by having them respond to stimuli, rather than just passively monitor them, can mitigate the vigilance decrement—at least after extended practice.

In an earlier study, however, Pop et al. (2010) included two levels of cognitive engagement in which operators made decisions about aircraft based either on a single feature or the conjunction of two features. Vigilance declined in all conditions. Taken together, these findings (Pop et al., 2010, 2012) suggest that task engagement can mitigate the vigilance decrement, but perhaps only when the method of increasing engagement imposes minimal cognitive demands on the operator. This interpretation is consistent with Molloy and Parasuraman's (Molloy & Parasuraman, 1996) finding that detection performance was better in a complex single task, which involved monitoring an automated routine in a simulated flight system, than in a complex multitask condition, which involved monitoring additional gauges, and a simple single-task condition, which involved discriminating squares that differed in size.

Vigilance can also be maintained at relatively high levels via motivational influences. Esterman et al. (2016) examined the motivational effects of monetary losses in a vigilance task in which participants began with the potential to keep $18. Participants either lost small amounts of money for errors throughout the task (the continuous-small loss condition) or lost all $18 if they incorrectly responded to one particular stimulus toward the end of the task (the anticipated-large loss condition). When tasks were well-practiced, sensitivity (d') remained stable in the anticipated-large-loss condition, but declined over time in the continuous-small-loss condition, showing that vigilance can be affected by motivational differences. Esterman et al. attributed the effect of the anticipated-large loss to the fact that the opportunity cost was held constant in this condition. The opportunity cost model (Kurzban et al., 2013) posits that when the value of alternatives is greater than the value of the current task, information processing resources will be reallocated to an alternative that is more rewarding or lower effort (e.g., mind wandering). Thus, in the continuous-small-loss condition, the value of alternatives is more likely to exceed the value of attending to the task as the remaining reward decreases in monetary value. Previous work that has failed to find an effect of reward and punishment on vigilance have employed continuous-small rewards/punishments (Bergum & Lehr, 1964; Esterman et al., 2014).

Vigilance decrements can also be attenuated by providing monitors with knowledge of results (e.g., a message saying "hit" after a correct detection (Chadda, 1992; Shaw et al., 2009; Szalma et al., 2006; Warm et al., 2009). Baker (1963b) suggested that knowledge of results (KR) improves vigilance by allowing monitors to generate accurate expectations about the timing and frequency of signals. However, this interpretation is challenged by findings that false KR is just as effective at mitigating vigilance decrements as true KR (Antonelli & Karas, 1967; Loeb &

Schmidt, 1963; J. F. Mackworth, 1964; R. L. Smith, 1966; Warm et al., 2009; Weidenfeller et al., 1962). These findings suggest that the effectiveness of KR can be attributed to the process of receiving feedback rather than to its informational content.

To disentangle the motivating and arousing effects of receiving feedback, Loeb and Schmidt (1963) compared the effectiveness of messages that simply acknowledged monitors' responses to messages that provided true and false KR. Acknowledgement of responses did not maintain performance, suggesting that the effectiveness of KR is not driven by the arousing effect of added stimulation, but by receiving feedback in particular. These findings led researchers to conclude that the facilitative effects of KR are mostly motivational in nature (Loeb & Schmidt, 1963; Sipowicz et al., 1962; R. L. Smith, 1966), though it is possible that feedback messages are more arousing than task-irrelevant messages. Alternatively, since these studies did not isolate the effects of KR on sensitivity and bias, KR might have increased the detection rate simply by modifying bias (i.e., feedback of missed signals would promote a more liberal response bias).

In addition, monitors often demonstrate an "end spurt" whereby vigilance improves towards the end of the task (Beh, 1989; Catalano, 1973; Childs & Halcomb, 1972; De Joux et al., 2013; Johnston et al., 1966). Given that the increase in vigilance occurs only when participants have knowledge of the length of the vigil and time remaining, researchers attribute it to an increase in motivation due to anticipation of the end of the task (Bergum & Lehr, 1963; Catalano, 1973; Dannhaus et al., 1976). Collectively, the findings that increased engagement, rewards, knowledge of results, and knowledge of vigil length can attenuate vigilance decrements suggests that vigilance is not merely a function of the resources available, as predicted by the resource depletion account. Instead, vigilance appears to be at least partially driven by monitors' motivation and willingness to attend to the task, consistent with the mind-wandering account. These findings might be reconciled with a resource depletion account if we assume that motivation and willingness to attend to the task increase resource availability (Kahneman, 1973).

**Increasing Task Demands Increase the Vigilance Decrement.** As previously discussed, the resource depletion account of the vigilance decrement is supported by evidence that more demanding tasks produce larger vigilance decrements (Caggiano & Parasuraman, 2004; See et al., 1995). In contrast, the mind-wandering account assumes that attentional lapses are driven by the understimulating and monotonous nature of vigilance tasks, such that more demanding tasks should reduce attentional lapses and produce smaller vigilance decrements. Accordingly, rates of mind-wandering tend to be lower in more demanding tasks (Forster & Lavie, 2009; Giambra, 1995; Smallwood et al., 2004; Thomson et al., 2013) and increasing task demands should therefore decrease the size of the vigilance decrement.

However, Smallwood and Schooler's attentional resource theory of mind-wandering assumes that performing vigilance tasks and engaging in TUTs draw from the same pool of attentional resources. As such, attending to one leaves fewer resources for the other. It follows, then, that cost of allocating resources to mind-wandering would be greater in vigilance tasks that are more resource-intensive (e.g., Feng et al., 2013; Thomson et al., 2014). The resource depletion and mind-wandering accounts of the vigilance decrement therefore both predict that more demanding tasks produce greater vigilance decrements because fewer attentional resources are available to the task.

**Mind-wandering Covaries with Performance.** On their own, findings that more demanding tasks produce larger vigilance decrements are unable to distinguish between the resource depletion and mind-wandering hypotheses. However, if vigilance decrements are in fact a result of mind-wandering, then decreases in vigilance should be paralleled by increases in mind-wandering. Methods of measuring mind-wandering include the probe-caught method, in which participants are periodically asked to report whether they are currently engaging in task-related or task-unrelated thoughts (Smallwood et al., 2007), the self-caught method, in which participants indicate when they become aware of task-unrelated thoughts (Smallwood et al., 2004), and retrospective reporting, in which participants estimate the frequency of task-related and task-unrelated thoughts at the end of the task (e.g., using the "thinking content" subscale of the DSSQ;  Matthews et al., 2002).

The evidence that mind-wandering is negatively correlated with overall performance in is mixed, with some studies reporting that the frequency of self-reported TUTs is negatively correlated with detections (Helton & Warm, 2008; Robertson et al., 1997), and others reporting that TUTs did not predict performance (Head & Helton, 2014). However, very few tasks assess the frequency of mind-wandering over time. Those that do indicate that the frequency of TUTs increase throughout the duration of the vigil, while correct detections decrease (Cunningham et al, 2000) and response times increase (McVay & Kane, 2012). If we assume that TUTs demand attentional resources, then under the resource depletion account, TUTs should decline over time as the pool of available resources diminishes. Although research directly assessing the prediction that increases in mind-wandering parallel declines in vigilance is sparse, the existing findings that TUTs increase over time support the mind-wandering account and are inconsistent with the resource depletion account.

So far, the mind-wandering wandering account of the vigilance decrement is supported by findings that vigilance decrements can be alleviated by increasing the extent to which monitors are engaged in and motivated to perform well in vigilance tasks and findings that mind-wandering increases over time. The relationship between increasing task demands and greater vigilance decrements can also be explained by the mind-wandering account. However, the idea that mind-wandering occurs because task it is under-stimulating is inconsistent with strong evidence that vigilance tasks are stressful (e.g., Caggiano & Parasuraman, 2004). Thus, neither the resource depletion theory nor the resource and control-failure theories of mind-wandering can fully account for all the findings observed.

### ***Resource Control Theory***

To better account for the full range of findings, Thomson, Besner and Smilek (2015) recently proposed a new theory of sustained attention that draws on aspects of the resource depletion account and the executive-resource and control-failure mind wandering accounts discussed above. Resource control theory posits that failures of executive control result in inappropriate allocation of attentional resources to mind-wandering. Helton and Warm (2008) had previously speculated from a resource depletion perspective that resources could be misallocated to TUTs.

Thomson et al. (2015) propose that executive control is responsible for distributing attentional resources among tasks, and that engaging executive control requires effort. The theory holds that the monotonous and unrewarding nature of vigilance tasks causes the motivation to engage executive control to wane over time, allowing information processing resources to be inappropriately allocated to mind-wandering. Specifically, executive control is said to fade because the cost—perceived as effort—of maintaining it outweighs the benefit of detecting occasional signals.

The resource control theory differs from the executive control theory of mind-wandering (McVay and Kane, 2009; 2012) in that it does not assume that executive control failures initiate instances of mind-wandering, only that control failures allow instances of mind-wandering to proceed unchecked and consume resources needed for the primary task. The resource control theory attributes the effortful and stressful nature of vigilance tasks to the increasing effort required to maintain the correct allocation of resources to the primary task over time.

**Re-examining evidence for sensitivity losses**

Overload and underload accounts of the vigilance decrement offer alternative explanations for the progressive decline in sensitivity over time. Under a resource depletion hypothesis, monitors' information processing resources deplete, limiting the amount of attention that can be directed to detecting signals. Under the mind-wandering hypotheses, the pool of resources does not shrink but is instead reallocated to TUT. Both the resource depletion account and the resource control theory of mind-wandering are consistent with the finding that vigilance tasks are subjectively effortful and stressful.

Although there is ongoing debate about the mechanism underlying the vigilance decrement, it seems clear that declines in sensitivity play an important role. A recent review, however, suggests that the evidence for sensitivity losses is much weaker than it appears (Thomson et al., 2016). Thomson and colleagues argue that shifts in response bias can masquerade as sensitivity losses when false alarms are extremely low. Although signal detection theory provides measures to isolate sensitivity from bias, measures of sensitivity based on binary responses are only bias-free when certain distributional assumptions are met. Recall that a decrease in correct detections may arise from a conservative shift in response criterion, in which case the observer makes fewer ‘signal’ judgments, resulting in fewer false alarms, or from a decline in sensitivity, in which case the observer more frequently confuses signals and noise stimuli, resulting in more false alarms.

The pattern of false alarms, then, is critical in distinguishing changes in sensitivity from changes in bias. As Thomson and collaborators (2016) explain, the issue is disentangling sensitivity and bias when false alarms are extremely low. This is rarely the case in regular signal detection tasks, in which the base rates of signal and noise events are typically equivalent, observers are required to respond on every trial, and signal and noise events are relatively difficult to discriminate. These factors encourage more moderate criterion placement than tasks with more frequent signals.

False alarms tend to be less frequent in vigilance tasks, though, as signals are rare and are usually quite distinct from noise events. Thus, task demands encourage more conservative criterion placement than tasks with equal base rates of signal and noise, and produce little overlap between the underlying signal and noise distributions. When false alarms are at or near floor, conservative shifts in an observer's response criterion can occur without detectably reducing false alarms. If a decrease in correct detections is not accompanied by a concurrent decrease in false alarms, then shifts in response bias could manifest as declines in the sensitivity metric, d'.

Thus, unless false alarms are high enough that decreases in hit rate could be accompanied by a commensurate decrease in false alarms, SDT metrics of sensitivity are not independent of response bias (Thomson et al., 2016). Several researchers have previously cautioned against the application of SDT to vigilance tasks for this reason (Caldeira, 1980; Craig, 1977; Craig & Colquhoun, 1975; Jerison, 1967; Williges, 1973). Yet, researchers continue to interpret apparent declines in hit rate as evidence that vigilance decrements are driven by sensitivity losses, even while acknowledging that false alarms are too low for meaningful analysis (e.g., (Ariga & Lleras, 2011; Helton et al., 2004, 2007, 2008; Temple et al., 2000).

By manipulating false alarm rates across conditions of a vigilance task, Thomson et al. (2016) showed that observed decreases in detections manifested as declines in sensitivity metrics when false alarms were low, but revealed changes in criterion when false alarms were sufficiently high. Both theoretically and empirically, Thomson et al. make a compelling case that within vigilance tasks, patterns of performance appearing to reflect declines in *d'* may simply reflect shifts in response bias.

To overcome the problem of low false alarm rates, researchers often opt to use the supposedly nonparametric measure of sensitivity, *A'*, that unlike d', can be calculated with hit rates of 1 and false alarm rates of 0.

$$A' = \frac{1}{2} + \frac{(HR - FAR) * (1 + HR - FAR)}{(4 * HR) * (1 - FAR)}$$

(J. B. Grier, 1971). *A'* corresponds to the area under the receiver operating characteristic curve, which is formed by plotting the cumulative HR against cumulative FAR at various criterion settings. The ability to calculate *A'* when false alarms are 0 makes it particularly appealing to vigilance researchers, since false alarms are often extremely low under conditions of low signal rates. It is now known, however, that *A'* increasingly underestimates sensitivity as bias becomes more extreme (McCarley et al., 2021; Pastore et al., 2003; Verde et al., 2006), meaning conservative shifts in response bias can manifest as declines in *A'* even when true sensitivity does not change.

Unfortunately, the use of *A'* is widespread within the vigilance literature (See et al., 1995). The preference for *A'* likely stems from the fact that the low signal rates in vigilance tasks often violate the assumption of equal variance required by d', and the misconception that *A'* is nonparametric (Craig, 1979). Many of the seminal findings linking memory demands, high event rates, and overall processing loads to sensitivity losses are based on reductions in *A'* over time.

Given the consensus that vigilance decrements are at least partially driven by conservative shifts in response bias over time, and the propensity for *A'* to underestimate sensitivity in these conditions, many of these apparent declines in *A'* might be driven purely by shifts towards more extreme response criteria.

Without clear evidence that performance decrements can be attributed to changes in sensitivity, the resource depletion account no longer stands out as the most plausible explanation of the vigilance decrement. Further, it suggests that vigilance researchers cannot simply assume that performance decrements reflect changes in sensitivity when tasks are designed according to Parasuraman and Davies' vigilance taxonomy. Thomson and colleagues urge researchers to "place higher value (and expend greater empirical effort) on theories of vigilant attention that do not hinge on declining observer sensitivity as the primary underlying cause of performance decrements" (p.24, 2016)

## Alternatives to Sensitivity Loss

The consensus among vigilance researchers is that the vigilance decrement is at least partially the result of conservative shifts in bias over time. Given that apparent sensitivity decrements may be spurious, it is possible that vigilance decrements are driven entirely by shifts in response criteria. Alternatively, vigilance decrements might be driven by an extreme form of mind-wandering—mindlessness—without implicating a loss of sensitivity.

### *Mindlessness*

The mindlessness hypothesis holds that the monotonous and understimulating nature of vigilance tasks encourages observers to withdraw their attention from the task, leading to increasingly automatized responding over time (Manly, 1999; Nachreiner & Hänecke, 1992; Robertson et al., 1997). Since signals are infrequent and observers need only respond when they detect a signal, the dominant ‘response’ in a vigilance task is to withhold responses from noise events. With time on task, withholding responses becomes more frequent, resulting in fewer correct detections. The key difference between mind-wandering and mindlessness, is that mindless responses are not stimulus driven. It is not merely that some attentional resources are redirected, but that attention is withdrawn from the task entirely, resulting in attentional lapses that are independent of signal strength. In contrast, mind-wandering reduces the quality of information processing, reducing sensitivity such that strong signals are detected more often than weak signals.

Assuming that false alarms are sufficiently high, attentional lapses would reduce hit and false alarm rates equally, making lapses indistinguishable from conservative criterion shifts in a binary signal detection task. Alternatively, if false alarms are near floor, the effect of lapses would only be observed on hit rates, appearing as a decline in sensitivity. However, the absence of a relationship between self-reported TUTs and time on task (Helton & Warm, 2008) suggests that attentional lapses alone cannot fully account for the vigilance decrement.

**Three mechanisms of vigilance decrement**

Together, current theories invoke three potential mechanisms by which vigilance may decline over time: criterion shifts, sensitivity losses, and attentional lapses. With only two degrees of freedom, binary signal detection data are unsuitable for discriminating between the proposed mechanisms. A novel method, employed by McCarley and Yamani (in press), is to analyze changes in the psychometric curve for a detection task. Psychometric curves plot signal response rates as a function of signal intensity and are characterized by three parameters. The shift parameter determines the curve's horizontal position and corresponds to the level of signal intensity at which a monitor reports a signal 50% of the time (i.e., response criterion). The scale parameter determines the slope of the curve and corresponds to how easily the monitor can discriminate signal from noise (i.e., sensitivity). A steeper slope implies better sensitivity, as the rate of responding changes more dramatically depending on the signal intensity. Finally, lapse rate determines the asymptote of the curve. A lower asymptote implies a higher rate of non-responses that are independent of stimulus intensity (i.e., attentional lapses).

As such, analyzing changes in psychometric curves over time can reveal the extent to which changes in response bias, sensitivity, and lapse rate contribute to decreased response rates in a vigilance task (See Figure 4).

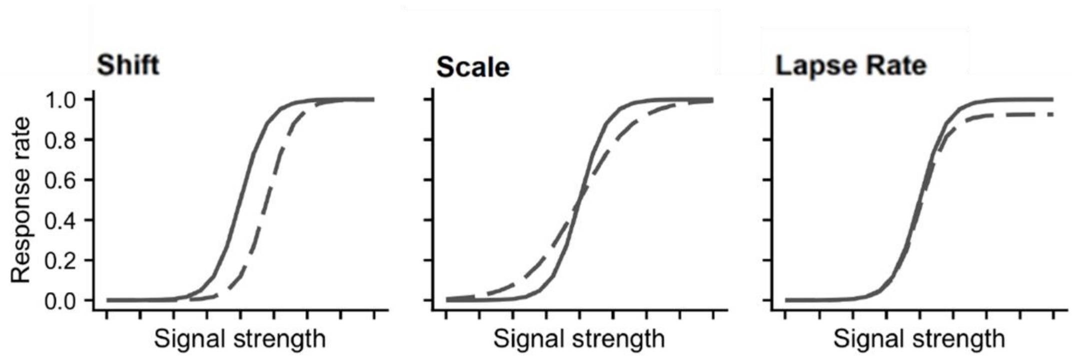


*Figure 4.* Sample psychometric curves. Curves are showing the effects of changes to shift (left), scale (middle), and lapse rate (right). Reproduced with permission from McCarley and Yamani (2021).

McCarley and Yamani (2021) employed Bayesian hierarchical modeling to fit psychometric curves to responses from a go/no go vigilance task. Vigilance declined over the course of the 20-minute task and parameter estimates revealed effects of shift, scale, and lapse rate. These findings provide preliminary evidence that all three mechanisms: bias shifts, sensitivity losses, and lapse rates, contribute to vigilance losses, though the effect of sensitivity losses might be small.

# Chapter 3: Assessing Online Vigilance Performance with Psychometric Curves

Chapter 3 presents two experiments that test the robustness of McCarley and Yamani's (2021) findings in an online sample. Experiment 1 employed a 12-minute sensory vigilance task adapted from McCarley and Yamani's 20-minute task. Using hierarchical Bayesian modeling, we fit detection data with psychometric curves to estimate changes in shift, scale, and lapse rate from the first to the last 4-minutes of the task. Experiment 2 is a large scale, preregistered replication of Experiment 1.

**Abstract**

When human monitors are required to detect infrequent signals among noise, they typically exhibit a decline in correct detections over time. Researchers have attributed this so-called vigilance decrement to three alternative mechanisms: shifts in bias, losses of sensitivity, and mind-wandering. The current experiments (n = 111, n = 194) examined the extent to which changes in these mechanisms contributed to the observed vigilance decrement in an online monitoring task. Participants completed an online visual signal detection task, judging whether the separation between two probes exceeded a criterion value. Separation was varied across trials and data were fit with logistic psychometric curves using Bayesian hierarchical parameter estimation. Parameters representing sensitivity, response bias, and attentional lapse rate were compared across the first and last four minutes of the vigil. Across two experiments, data gave evidence of an increased attentional lapse rate and conservative shifts in response bias, with inconclusive evidence for or against an effect of sensitivity. Sensitivity decrements appear less robust than criterion shifts or attentional lapses as causes of the vigilance loss.

**Assessing Online Vigilance Performance with Psychometric Curves**

Many tasks require people to watch for rare but critical events over a long time; consider baggage screeners and sonar operators watching for security threats, quality control inspectors watching for faults, and drivers watching for hazards on the road. Discriminating rare signals from background noise for long durations requires vigilance—a state of "psychological readiness to perceive and respond" to stimuli (Mackworth, 1948, p. 6). As time on watch progresses, vigilance declines and monitors miss signals they would otherwise be able to detect. This decline in detections over time, called the vigilance decrement, typically begins within the first half hour, sometimes in as little as 5 minutes (Nuechterlein et al., 1983).

Despite intensive study of the topic since the 1940s, researchers have yet to reach consensus on the mechanisms underlying the vigilance decrement. The most common framework for studying vigilance has been signal detection theory (SDT; Green & Swets, 1966; Hautus et al., 2022; Swets et al., 1961; Tanner & Swets, 1954), a model of the process by which observers transform probabilistic evidence into discrete judgments. In a conventional yes-no signal detection task, the observer is asked to discriminate between two possible states of the world, typically termed noise (N) and signal plus noise (S+N). Under the SDT model, the observer encodes evidence for or against the presence of a signal each trial as a scalar value (Pastore et al., 2003). Evidence values vary continuously and probabilistically, and by convention, the S+N distribution is assumed to have a mean greater than or equal to that of the N distribution.

Confusability between states of the world exists when the evidence distributions corresponding to N and S+N events overlap. To reach a discrete judgment, the observer compares the encoded evidence value to a cutoff (Green & Swets, 1966), responding with an S+N judgment if the evidence value is above the criterion and N judgment otherwise. Sensitivity, the observer's ability to distinguish signal from noise, increases as overlap between distributions decreases. Bias, the observer's tendency to favor either N or S+N judgments, is determined by placement of the response cutoff.

SDT thus allows two possible mechanisms by which signal detection rates might decrease over time--sensitivity losses and conservative bias shifts—and data have suggested that in fact, both can contribute to the vigilance decrement. Conservative bias shifts occur as observers adjust their behavior to the low signal rate, gradually moving their response cutoff upward. (Broadbent & Gregory, 1965; Craig, 1978; Williges, 1969). Bias shifts are common in vigilance tasks (Broadbent & Gregory, 1965; Colquhoun & Baddeley, 1964, 1967; Swets, 1977; Warm et al., 2015), and have been regarded as the primary cause of the vigilance decrement (Craig, 1978). Decreases in sensitivity are possible but less common, generally occurring only in tasks that impose high time stress and heavy demands on working memory and perception (Nuechterlein et al., 1983; Parasuraman, 1979; See et al.,1995; Swets, 1977).

The most popular account of these selective effects, the resource depletion theory of vigilance (Caggiano & Parasuraman, 2004; Grier et al., 2003; Helton & Warm, 2008; Parasuraman, 1979) proposes that sensitivity losses occur when sustained information-processing demands gradually consume the operator's attentional capacity. An alternative account, the resource control model (Thomson et al., 2015), suggests that the sensitivity decrement occurs when failures of executive control allow information processing resources to drift to off-task processing. Both models are consistent with findings that vigilance tasks are stressful and subjectively effortful (Dember et al., 1996; Warm et al., 2008).

Recent work, however, suggests that yes-no signal detection data may be inadequate for understanding the vigilance decrement. One concern has been the suggestion that apparent losses of sensitivity could be the result of a statistical artifact (Thomson et al., 2016). In general, conservative bias shifts and sensitivity losses—which both reduce hit (i.e., true-positive) rates—are distinguished in SDT by their effects on false alarm (i.e., false-positive) rates. All else being equal, false alarm rates decrease when bias becomes more conservative and increase when sensitivity declines. If evidence distributions for N and S+N events are assumed to be Gaussian and equal variance, the tradeoff between hits and false alarm rates in yes-no data is captured by the parametric sensitivity measure *d'* (Hautus et al, 2022),

$$d' = \text{Z(hit rate)} - \text{Z(false alarm rate)} \tag{1}$$

where Z denotes the inverse normal transformation.

Unless the false alarm rate is high enough to allow a statistically detectable decrease between task conditions, however, changes in sensitivity and bias are indistinguishable. Unfortunately, as Thomson et al. (2016) have noted, the low signal rate inherent to vigilance tasks encourages monitors to adopt an extremely conservative cutoff for S+N responses, often producing mean false alarm rates very near zero and introducing the risk of spurious sensitivity losses. Although alternative measures of sensitivity have been proposed for the analysis of yes-no data (e.g., Pollack & Norman, 1964; Craig, 1979; Szalma et al., 2006), none of them is bias free (Getty et al., 1995; McCarley et al., 2021; Macmillan & Creelman, 1996; Pastore et al. 2003). These considerations imply that the existing evidence for sensitivity losses in vigilance tasks may be less convincing than previously thought (Thomson et al., 2016).

Yes-no tasks also provide no direct method of testing for a third form of vigilance error, mindless responses. The mindlessness account of vigilance (Robertson et al., 1997; Manly et al., 1999; Thomson et al, 2014; Warm et al., 2015) argues that detection rates fall because of lapses that occur when attention is redirected from the monitoring task to self-generated, task-unrelated thoughts. The resource control hypothesis (Thomson et al., 2015) links these lapses to a traditional resource model, suggesting that lapses result when executive control failures allowing the mind to wander from the vigilance task to unrelated thoughts (Kane & McVay, 2012). Empirical data have confirmed that off-task thoughts increase over time in monitoring tasks, and that the frequency of off-tasks thoughts is correlated with vigilance performance (Thomson et al., 2014). But, because yes-no signal detection models provide at most two degrees of freedom (depending on whether the false alarm rate is above floor), they provide no way of accounting for the effects of mental lapses. Further, because lapses will tend to change hit and false alarm rate proportionately, they violate the equal-variance Gaussian assumption, and thereby distort calculations of sensitivity and bias from yes-no data.

The method commonly used to assess vigilance performance, the yes-no signal detection task, thus conflates three potential mechanisms by which vigilance may decline over time--criterion shifts, sensitivity losses, and attentional lapses. An alternative method of discriminating between the proposed mechanisms, is to analyze changes in the psychometric curve (McCarley & Yamani, 2021). The psychometric curves plot behavior in a psychophysical task as a function of a given stimulus property (Kingdom & Prins, 2016). The psychometric curve for a detection task, for example, might plot detection rates as a function of signal intensity, and are characterized by three parameters. The first, shift, determines the curve's horizontal position and provides a measure of response bias. The second parameter, scale, is inversely related to the slope of the curve and corresponds a measure of sensitivity. Finally, lapse rate determines the asymptote of the curve; an attentional lapse implies that a signal will go undetected even if it is well above the operator's response threshold, and a higher lapse rate therefore manifests as a decrease in asymptote. As such, analyzing changes in psychometric curves over time can reveal the extent to which changes in response bias, sensitivity, and lapse rate contribute to decreased response rates in a vigilance task. Figure 5 shows examples of curves, from a hypothetical vigilance task, differing in three parameters.

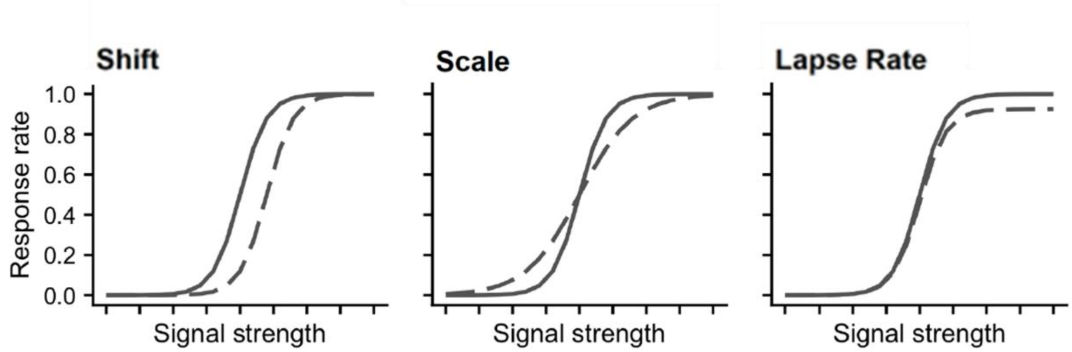


Figure 5. Sample psychometric curves. Curves are showing the effects of changes to shift (left), scale (middle), and lapse rate (right). Curves in the right panel are plotted under the assumption that a lapse leads to a failure to respond. Reproduced with permission from McCarley and Yamani (2021).

McCarley and Yamani (2021) employed Bayesian hierarchical modeling to fit psychometric curves to responses from a go/no go vigilance task. Vigilance declined over the course of the 20-minute task and parameter estimates revealed effects of shift, scale, and lapse rate, showing that all three mechanisms contribute to vigilance losses. We test the robustness of these findings in two experiments, in which we adapted McCarley and Yamani's (2021) procedure for use in an online task. Research examining vigilance performance online is sparse (Luna et al., 2021; Ralph et al., 2015; Thomson et al., 2016), perhaps, in part, because online studies offer less environmental and perceptual control than lab-based studies (Claypoole et al., 2018). But, despite variability in device type, screen size, testing location and environmental conditions, data collected online are generally consistent with, and of comparable quality to, data collected in the lab (Germine et al., 2012; McGraw et al., 2000). It is currently unknown whether this holds true for vigilance data.

At present, two published studies (Claypoole et al., 2018; Luna et al., 2021) have replicated an online vigilance task in a lab setting, producing mixed results. Collecting data in a laboratory, Claypoole and colleagues reported declines in correct detections like those that Thomson et al. (2016) found in an online version of the same task. However, they failed to replicate changes in metrics of sensitivity. In the online study, conventional *A'* scores indicated a sharp decline in sensitivity over time, whereas *A'* scores corrected for low false alarm rates did not significantly change. In the lab, *A'* and corrected *A'* scores both increased over time. Luna and colleagues observed similar declines in *A'* in their lab-based and online vigilance tasks, but interpretation of these results is limited, given that *A'* is known to vary with shifts in bias (McCarley et al., 2021; Pastore et al. 2003).

We expect our online vigilance task to produce a decline in detections over time, but it remains to be seen whether the effects of shift, scale, and lapse rate are comparable in the lab and online.

**Experiment 1**

In Experiment 1, we followed the procedure of McCarley and Yamani's (2021) study in an online sample. Participants viewed a series of probe circles and judged whether the gap between each pair of probes exceeded 2 cm. To maximize online completion rates, we shortened the task from 20-minutes to 12-minutes. There were also small differences in task and stimulus design as a result of reprogramming the experiment to run online. We fit gap discrimination data with psychometric curves, and analysis compared shift, scale and lapse rate parameters for the first and last 4-minutes of the task.

***Method***

**Participants.** One hundred and eleven undergraduate students were recruited from a large public university in the United States. Inclusion criteria were fluency in English, normal color vision, and normal or corrected-to-normal visual acuity. All participants gave informed consent to participate. Data were excluded from participants who failed to complete the full experimental session or to achieve d' scores of ≥ 0.25 in each 4-minute block of the task. Exclusions left 103 participants for analysis ($M_{age}$ = 22.58 years, gender = 80 females, 20 males, 2 non-binary, 1 not specified). All participants received course credit for a 30-minute experimental session.

**Apparatus and stimuli.** The experimental task was controlled by software written in PsychoPy 3 (Peirce et al., 2019) and hosted on Pavlovia (https://pavlovia.org/). Stimuli were scaled to participants' monitors to maintain consistent sizing across devices. The stimulus each trial was a pair of red probe circles embedded amongst five black distractor circles. Distractors which were intended to decrease signal discriminability, a characteristic associated with rapid sensitivity decrements (Nuechterlein et al., 1983). All circles were unfilled, drawn in 2-pixel stroke, and had a diameter of 0.3 cm. Stimuli were presented on a white background, within a circular search field 8 cm diameter.

The two probe circles were arranged horizontally, separated by a distance that varied across trials, as described below. The midpoint between probe circles was randomly assigned a position with an x-coordinate within an imaginary square of 2 × 2 cm centered within the display. Distractor circles were randomly assigned a position with an X-coordinate ± 2.8 cm from the center of the field and a Ycoordinate ± 1 cm from the center, allowing them to appear in any location along the X-axis that probe circles could appear.

**Procedure.** To match stimulus size across monitors, participants were first prompted to complete a screen-scaling procedure in which they resized an onscreen image of a credit card to match the size of a physical card. They then performed a signal detection task in which they judged whether the horizontal distance between probe circles each trial exceeded a criterion value of 2 cm. To account for any potential variability in stimulus sizing that remained after the scaling procedure, the task instructions did not describe the criterion distance in units of length, but simply showed an example of probe circles separated by the criterion distance. Participants were asked to press the space bar if the gap between probe circles exceeded the criterion distance on a given trial, and to withhold response otherwise.

We defined gaps greater than the criterion distance as signal events, and gaps equal to or less than the criterion distance as noise events. The distance d between probe circles varied across trials between 0.5 cm and 3.25 cm, in steps of 0.25 cm. The value of d for a given trial was determined probabilistically: on each trial, there was an 80% probability that the gap would be less than or equal to the criterion distance of 2 cm, and a 20% probability that gaps would exceed the criterion distance. After a trial was determined to be signal or noise, the gap size was selected randomly and with equal probability from amongst the range of possible values for that trial type. The range of non-signal values (0.5—2.0) was larger than the range of signal values (2.25—3.25) to deter participants from using the stimulus distribution midpoint as an implicit criterion.

Participants first completed a practice vigil of 90 trials, followed by a 12-minute experimental vigil. Vigil length was not disclosed to participants. Each trial comprised a 250 ms stimulus display followed by a blank interval of 1,250 ms, during which only the outline of the search field remained visible (see Figure 6). The subsequent trial began immediately thereafter, producing an event rate of 40 trials per minute. A response was attributed to trial i if it occurred before the onset of trial i + 1. Participants did not receive post-trial feedback.

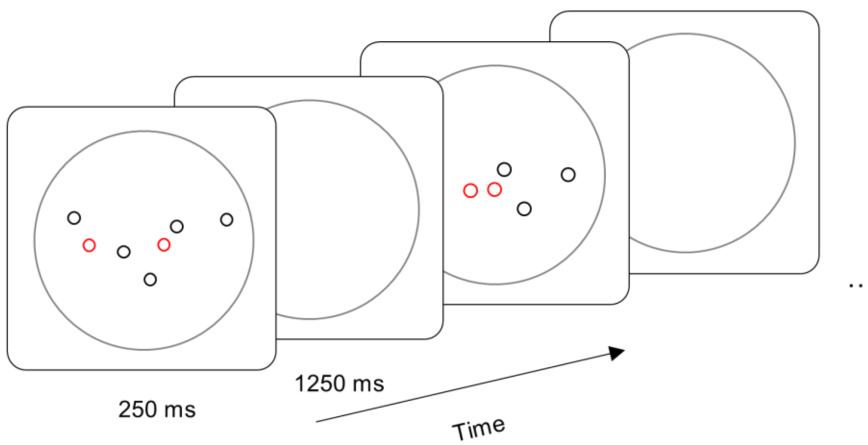


Figure 6. Sequence of events over a signal (left) and non-signal (right) trial. Not to scale.

The practice vigil was the same as the experimental vigil except that 1) signal and noise events were equally probable, creating a signal rate of 0.50, 2) for the first 25 trials of the vigil, the stimulus display remained visible for the full trial duration of 1,500 ms, and 3) response errors were followed by a 1-second feedback message reading either, “Oops! It was not a target.”, or “Oops! You missed a target.”, as appropriate. Error-free performance resulted in a practice vigil of 2 minutes 15 seconds and each error added 1 second.

At the end of the vigil, participants completed a computerized A-SWAT mental workload scale (Luximon & Goonetilleke, 2001). The A-SWAT consists of three subscales: time load, mental effort, and psychological stress, which were presented in order, one at a time. Participants made their rating of each subscale by clicking a horizontal line anchored with the text descriptions of subscale endpoints.

***Analysis***

To exclude participants who might have stopped attending to the task entirely, we converted participants' binary responses to the signal detection theory measure of sensitivity, d'. To correct for ceiling- or floor-level hit and false alarm rates, d′ scores were calculated using a log-linear correction (Hautus, 1995). Six participants who failed to achieve a d' of at least 0.25 in the first, middle, or last 4-minutes of the task were excluded from further analyses. With these exclusions, mean d' was 2.33, mean hit rate was 0.86, and mean false alarm rate was 0.16 for the whole vigil.

We used hierarchical Bayesian parameter estimation (Kruschke, 2015; Lee, 2018; Lee & Wagenmakers, 2013) to assess changes in vigilance between the first and last 4-minutes of the task. Signal detection responses were fit with logistic psychometric curves with three parameters: shift, $\alpha$, representing response bias; scale, $\beta$, representing sensitivity; and asymptote, $\lambda$, representing lapse rate (Kingdom & Prins, 2016). Standardized mean differences in shift, scale, and lapse rate between first and last blocks were modeled with normalized effects and unit normal priors. For consistency and ease of comparison across parameters, the model placed priors on the probit-transformed lapse rate rather than on lapse rate directly, such that prior values corresponded to a uniform distribution over the interval [0, 1].

Follow-up analyses used Savage-Dickey density ratios to assess evidence for the effects of block on the standardized mean differences in scale, shift, and lapse rate from the first to last block. The Savage-Dickey ratio is the height of the posterior distribution divided by height of the prior distribution at the parameter value of interest, in this case, $\delta\alpha = 0$, $\delta\beta = 0$, and $\delta\lambda = 0$. The resulting Bayes factor, denoted*B10*, is the ratio of the likelihood of the data under the alternative hypothesis versus the null, and therefore summarizes the strength of the evidence for or against the alternative. A ratio of 1 indicates no evidence in either direction, values greater than 1 support the alternative hypothesis, and values between 0 and 1 support the null. For ease of interpretation, we describe the strength of evidence using the qualitative guidelines proposed by Jeffreys (1961), whereby a Bayes factor of 1-3 is considered anecdotal evidence, 3-10 is considered substantial evidence, 10-30 is considered strong evidence, 30-100 is considered very strong evidence, and > 100 is considered decisive evidence for the alternative hypothesis.

Mean ratings for each of the A-SWAT subscales were estimated separately within a hierarchical model that placed a normal likelihood function on observed ratings, and uniform priors, U(1, 100), on the group means and standard deviations of the ratings.

All analyses were conducted in R (R Core Team, 2019). Estimation procedures ran four MCMC chains for 10,000 warmup trials, followed by 250,000 sample steps each, using the JAGS package (Plummer, 2015). Chains were thinned to every fifth step, leaving 50,000 samples for analysis. All parameter estimates showed $\hat{R}$ convergence values of < 1.1, indicating satisfactory convergence of MCMC chains.

***Results***

The psychometric curves in the left panel of Figure 7 show the proportion of trials on which participants responded for the first and last 4-minute blocks of the task. Symbols represent the empirical data, with responses for gaps ≤2 cm corresponding to false alarms and responses for gaps > 2 cm corresponding to hits. The right panel shows mean differences in response rate between blocks, with negative values indicating lower response rates in the last block. The error bars in both panels represent 95% posterior predictive credible intervals, based on data simulated from the posterior distribution. Here, the intervals demonstrate that the model captures the trend of the empirical data. Response rates were lower in the last block than in the first, showing a vigilance decrement.

Figure 8 shows the posterior distributions of the standardized mean differences in shift, scale, and the probit-transformed lapse rate between the first and last blocks. Shift,$B10$ = 94.53, and lapse rate,$B10$ = 2891.56, both increased between blocks, indicating a conservative change in response bias and an increased rate of attentional lapses in the last block. The scale parameter trended in the direction of a sensitivity loss, with the mass of the posterior distribution falling above 0. However, the Bayes factor gave no conclusive evidence for or against a change in scale between blocks,$B10$ = 0.85. The standardized mean difference for lapse rate, $M_{diff}$ = 1.24, 95% BCI[0.61, 2.03], was larger than for shift, $M_{diff}$ = 0.42, 95% BCI[0.2, 0.65] or scale, $M_{diff}$ = 0.30, 95% BCI[-0.02, 0.64].

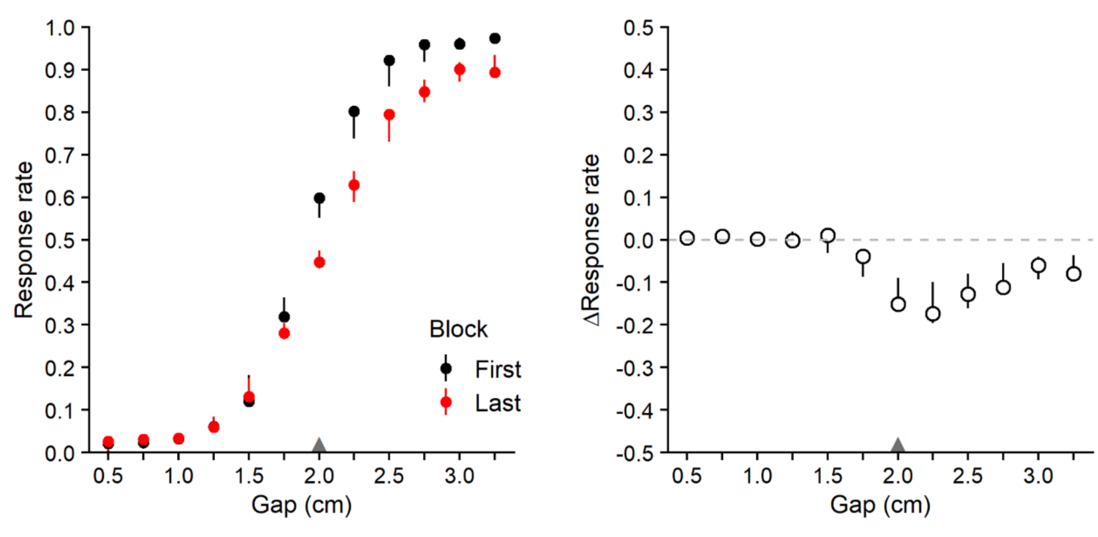


Figure 7. Response rates for the first (black) and last (red) 4-minute blocks of the vigilance task. Symbols represent empirical means, error bars represent 95% posterior predictive credible intervals, and the gray triangle denotes the boundary between noise and signal events.

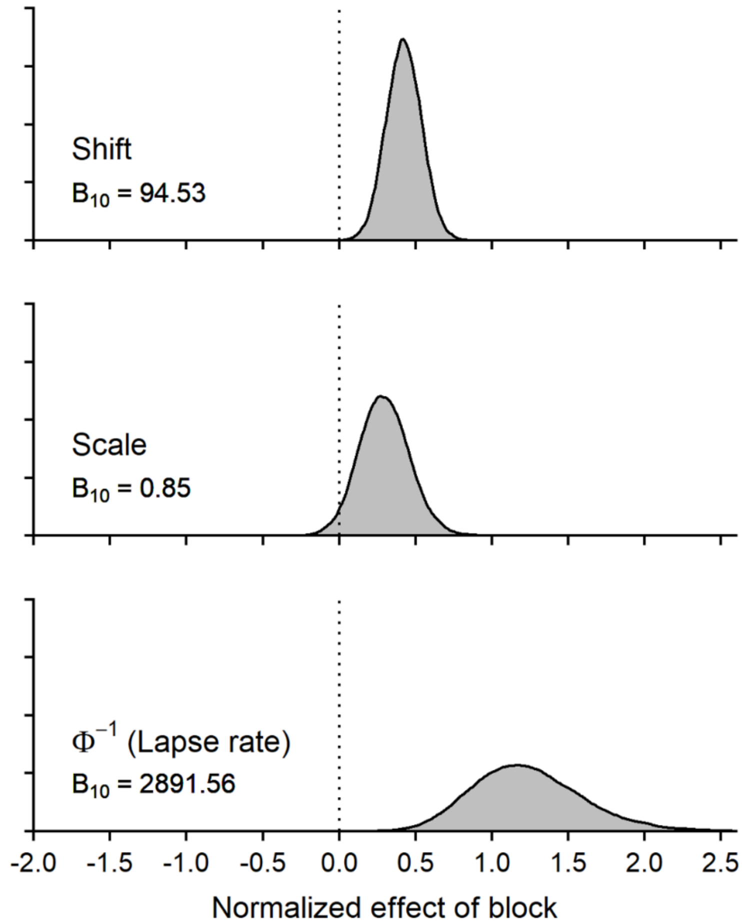


Figure 8. Posterior density plots of the standardized mean differences in shift, scale, and probit-transformed lapse rate between first and last blocks.

While standardized effect sizes are useful for comparing effects on parameters that differ in scale, they do not directly reveal the impact of the observed effects on raw data (Pek & Flora, 2018). To compare the selective effects of shift, scale, and lapse rate changes on raw response rates between blocks, McCarley and Yamani (2021) calculated posterior predictive data for each parameter individually, while holding other parameters fixed at their mean values. They found that conservative shifts in response bias explained most of the observed differences in response rate between blocks. Their data indicated that changes in lapse rate produced a decrease in response rates for signal events, while changes in scale explained small increases in response rates for noise events just below threshold, and small decreases in response rates for signal events just above threshold. We applied the same analysis to the current data. Figure 9 presents the posterior predictive distributions for the selective effects of block on shift (holding scale and lapse rate constant), scale (holding shift and lapse rate constant), and lapse rate (holding shift and scale constant). The mean change in response rate between the first and last blocks is plotted for comparison.

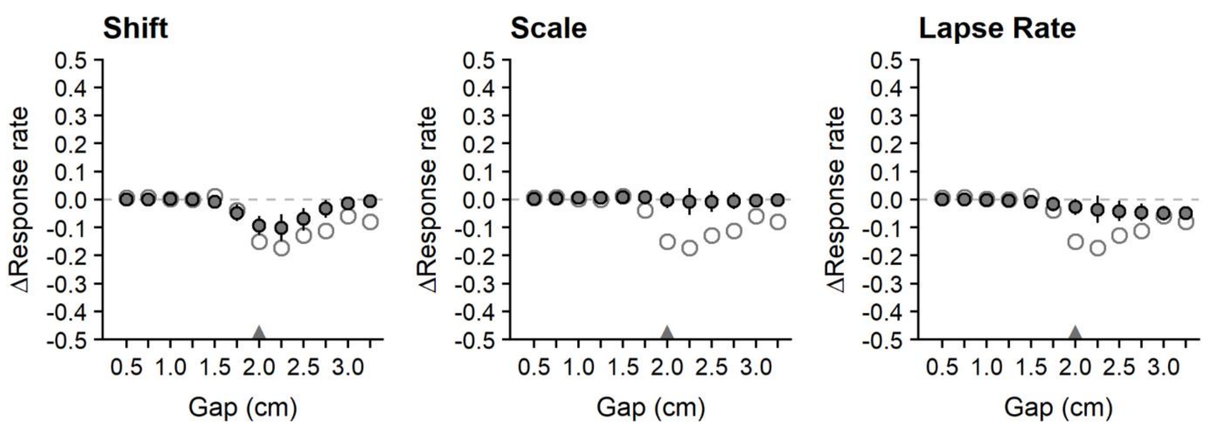


Figure 9. Selective effects of shift, scale, and lapse rate on posterior predictive differences in response rate. Filled symbols represent mean response rate differences and error bars represent 95% credible intervals. Open symbols are mean changes in response rate, representing the combined effects of shift, scale, and lapse rate.

Of the selective effects, changes in shift and lapse rate accounted for most of the decrease in response rates between blocks. Changes in shift produced large decreases in response rate around gaps of 2 cm, showing that participants required more evidence (i.e., larger gap sizes) to produce a signal response over time. Response rates for gap sizes that were clearly above or below the criterion for signal were unaffected by the change in shift, as expected. Changes in lapse rate produced decreases in response rate for gaps that were unambiguously larger than 2 cm. Gaps less than 2 cm showed little effect of attentional lapses since the correct ‘response’ in these cases was to refrain from responding. The selective effect of scale had a negligible effect on raw response rates from the first to last block.

Estimated mean A-SWAT ratings were M = 33.61, 95% BCI[28.33, 38.86] for time load, M = 81.02, 95% BCI[77.43, 84.65] for mental effort, and M = 59.96, 95% BCI[54.83, 65.08] for stress. High ratings for mental effort and psychological stress are consistent with findings that vigilance tasks impose a high workload on monitors (e.g., Caggiano & Parasuraman, 2004; Helton & Warm, 2008).

### ***Discussion***

Psychometric curves gave decisive evidence for changes in lapse rate, strong evidence for changes in shift, and inconclusive evidence for changes in scale. Changes in each parameter were associated with decreased responses in the last block, with lapse rate accounting for the largest decrease in response rate. These results, though preliminary, demonstrate that the vigilance decrement was driven by attentional lapses and conservative shifts in bias in an online task. We conducted a second experiment to validate these findings in a pre-registered replication.

## **Experiment 2**

Experiment 2 was a close, pre-registered replication of Experiment 1. We recruited paid participants from Prolific (https://prolific.co) rather than from the undergraduate subject pool. Research suggests that participants compensated with pay outperform those who are compensated with course credit (Brase et al., 2006), and

Prolific encourages highly engaged participants by enforcing fair rewards and rigorous data quality rules.

Details of the sample, method, and analyses were pre-registered on Open Science Framework (https://osf.io/nt6u3) prior to data collection. Analysis scripts and data are also available on OSF.

***Method***

**Participants.** The sample size was determined using the following pre-registered plan: We recruited an initial 125 participants from Prolific and continued in increments of 25 until (i) the Bayes factors for effects of shift, scale, and lapse rate indicated an evidence ratio of at least 1:10 in either direction (i.e., until the data were 10 times as likely under the alternative or null hypothesis than the other), or (ii) reaching a maximum of 200. Data collection ceased at 200 participants, at which point the Bayes factor for the effect of scale had not yet reached the 1:10 evidence criterion.

Participants gave informed consent to participate and were screened for English fluency, normal color vision, and normal or corrected-to-normal visual acuity. Data were excluded from three participants who failed to complete the full experimental session or to achieve *d'* scores of ≥ 0.25 in each 4-minute block of the task. Data files were blank for an additional three participants, leaving 194 participants for analysis (Mage = 27.77 years, gender = 77 females, 118 males, 2 not specified). All participants received 10.00 USD/hour for a 30-minute experimental session.

All other methodological details were identical to Experiment 1.

***Results***

The left panel of Figure 10 shows response rates for the first and last 4-minute blocks of the task. The right panel shows the mean change in response rates between the first and last blocks. Visual inspection shows that the response rate in Experiment 2 was higher in the first 4-minutes of the task than the last 4-minutes, indicating a vigilance decrement. The error bars represent 95% posterior predictive credible intervals. Although they capture the overall trend of the data, the posterior predictive intervals slightly overestimate the response rate for values just below the signal threshold, and slightly underestimate the response rate for values above the signal threshold.

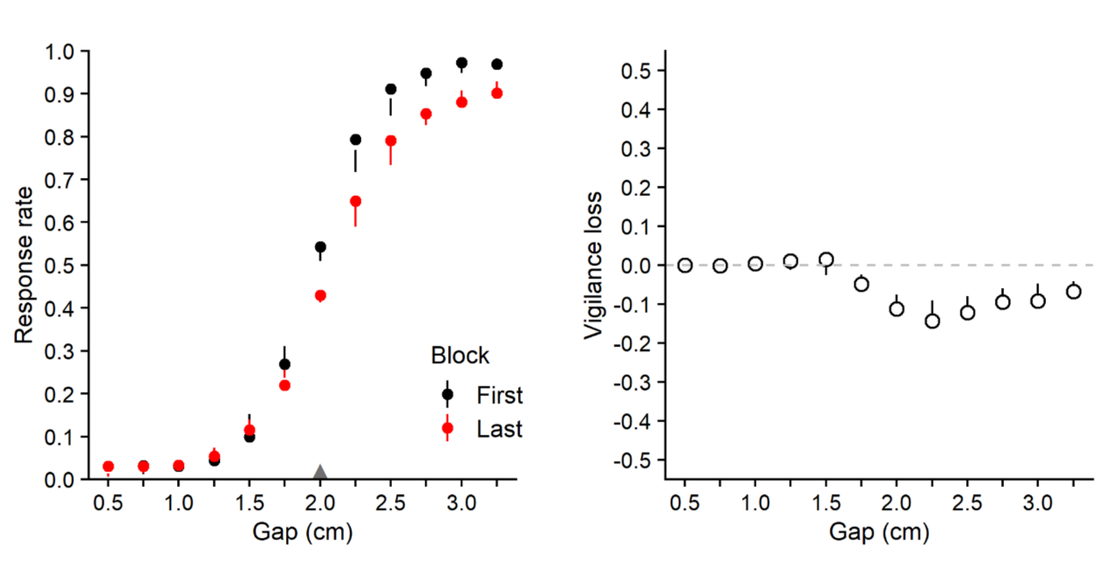


Figure 10. Response rates for the first (black) and last (red) 4-minute blocks of the vigilance task. Symbols represent empirical means, error bars represent 95% posterior predictive credible intervals, and the gray triangle denotes the boundary between noise and signal events.

Figure 11 shows the posterior distributions of the standardized mean differences in shift, scale, and the probit-transformed lapse rate between the first and last blocks. Shift,*B10* = 841.97, and lapse rate,*B10* = 3012.02, increased between blocks, indicating a conservative shift in response bias and an increased rate of attentional lapses in the last block, respectively. Evidence for a change in scale favored the null hypothesis (i.e., that sensitivity did not decline) but not strongly,*B10* = 0.27.

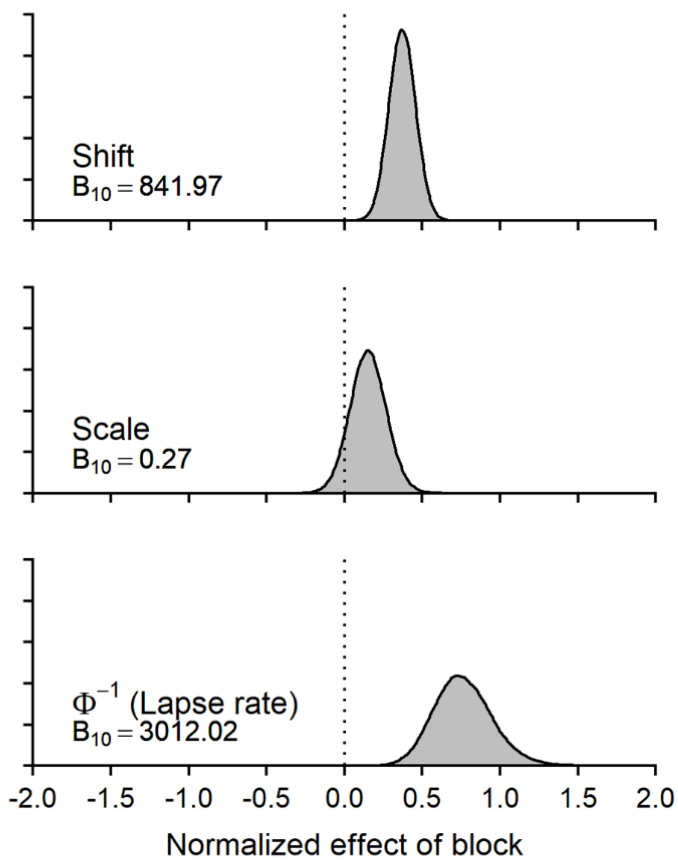


Figure 11. Posterior density plots of the standardized mean differences in shift, scale, and probit-transformed lapse rate between first and last blocks.

Figure 12 presents the posterior distributions for the selective effects of block on shift, scale, and lapse rate, with the mean change in response rate between the first and last blocks plotted for comparison.

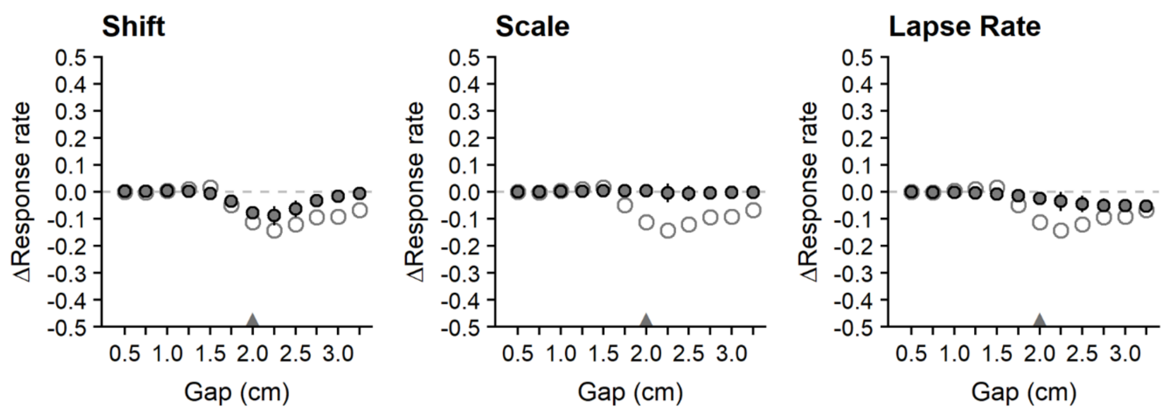


Figure 12. Selective effects of shift, scale, and lapse rate on posterior predictive differences in response rate. Filled symbols represent mean response rate differences and error bars represent 95% credible intervals. Open symbols are mean changes in response rate, representing the combined effects of shift, scale, and lapse rate.

Again, changes in shift produced decreases in response rates for gaps around 2 cm, indicating more conservative responding over time, and changes in lapse rate produced decreases in response rate for gaps clearly above criterion. Scale did not account for any meaningful changes in raw response rates across blocks.

Estimated mean A-SWAT ratings were M = 35.84, 95% BCI[31.64, 40.05] for time load, M = 80.05, 95% BCI[77.62, 83.37] for mental effort, and M = 51.34, 95% BCI[47.15, 55.54] for stress.

***Discussion***

Data gave decisive evidence for changes in lapse rate and shift, but not scale, over time, replicating the pattern of results observed in the first experiment. However, compared to Experiment 1, the current experiment found a larger effect of shift and a smaller—though still decisive—effect of lapse rate.

## General Discussion

In two experiments (N = 103 and N = 194), we analyzed psychometric curves for the first and last 4-minutes of an online vigilance task, testing the extent to which changes in response bias, sensitivity, and lapse rate contributed to the vigilance decrement. In both experiments, the vigilance decrement was largely driven by (i) conservative shifts in response bias and (ii) more frequent attentional lapses. In other words, over the course of the task, people (i) became less willing to call to a given gap size a signal, independent of their ability to discriminate between signal and noise, and (ii) experienced a greater number of failures to respond to even the most intense signals.

Neither experiment found strong evidence for or against an effect of scale, which corresponds to sensitivity in the SDT framework. In both experiments, the mass of the posterior distribution fell above 0, which may be interpreted as a trend toward an effect, yet the Bayes factors were inconclusive in Experiment 1 and favored a null effect by roughly a 4-to-1 ratio in Experiment 2. Although we placed diffuse priors on parameter estimates, Bayes factors can be highly sensitive to prior distributions (Liao et al., 2021; Liu & Aitkin, 2008). Given the relatively weak evidence, a different prior specification might have produced different results.

Altogether, the findings of null sensitivity changes should be treated as uncertain, though the data do suggest sensitivity losses were small, at best.

The present experiments partially replicate McCarley and Yamani's findings (2021). Although both studies found effects of bias shifts and lapses, there were differences in the relative sizes of these effects. While McCarley and Yamani reported vigilance losses primarily driven by bias shifts, the decrement was driven strongly by both bias shifts and attentional lapses in the present study. McCarley and Yamani also reported decisive changes in scale over the course of their task, indicating a sensitivity loss between the first and last 4-minutes of their 20-minute task. Cumulatively, we found weak evidence against an effect of block on scale.

Discrepancies in the magnitude of lapse rate and scale effects between the current study and Yamani and McCarley's (2021) may have been driven by differences in face-to-face versus online data collection. Participants in the present study completed the online vigilance tasks in uncontrolled, unsupervised locations, likely exposing them to interruptions and distractions that could have affected attention allocation. For example, dividing attention between the online vigilance task and an alternative task, such as checking one's phone or watching television, may have 1) inflated the lapse rate, and 2) reduced total information processing demands, potentially attenuating sensitivity losses. This possibility accords with data reported by Casner and Schooler (2015), who found that pilots in a simulated flight task used external activities strategically to maintain their alertness, but that engagement in external activities, ironically, led to monitoring failures.

It is also possible that participants completing the online tasks simply expended less effort than those in-person, leaving less room for sensitivity to decline.

However, perceived mental effort was higher in the current experiments ($M_{effort}$ = 81.02 and $M_{effort}$ = 80.05 for Experiments 1 and 2, respectively) than in the lab experiment ($M_{effort}$ = 72.3; McCarley & Yamani, 2021), suggesting that the absence of a sensitivity loss online was not driven by a lack of effort.

Alternatively, differences between the current findings and McCarley and Yamani's could be explained by differences in vigil length or stimulus characteristics. To increase the likelihood that participants would complete the online task, we reduced the duration of the vigil from 20 to 12 minutes. Previous use of abbreviated vigilance tasks has found that 12-minutes is sufficient to produce a vigilance decrement (Craig & Klein, 2019; Temple et al., 2000). It is possible, though, that sensitivity does not decline as rapidly as changes in bias and lapse rate occur, such that the abbreviated task may have produced smaller sensitivity losses than a longer task would have. Additionally, stimuli in the lab-based task were embedded in dynamic Gaussian visual noise to decrease discriminability. Because were unable to implement this noise online, probe stimuli were instead embedded among distractor circles of a different color. Visual noise was therefore less dense in the online task, which also could have limited sensitivity losses. Future work should test these hypotheses directly.

Results of this study indicate that the vigilance decrement was primarily driven by attentional lapses and conservative criterion shifts in this online vigilance task. While sensitivity losses may contribute to the decrement under alternative conditions, the effect of sensitivity appears to be less robust and less consequential than the effects of bias and attentional lapses. These findings are at odds with theories that assign resource depletion a significant role in the vigilance decrement, and suggest that vigilance decrements may reflect different mechanisms when tasks are conducted online rather than in laboratory settings. Further, these results imply that interventions targeting response bias and attentional lapses might be more effective at mitigating vigilance decrements than interventions focused on reducing sensitivity losses.

# Chapter 4: Mechanisms of Vigilance Loss: A Cognitive Modeling Approach

As an alternative to fitting detection data with psychometric curves, Chapter 4 presents a cognitive process model with parameters directly representing observers' response criteria, internal decision noise, and attentional lapse rate. This approach provides quantitative predictions and links performance changes directly to putative cognitive mechanisms. The generative cognitive model is used to reanalyze data from two previous experiments, demonstrating that it captures trends in the data that were revealed by psychometric curve analyses.

## Abstract

The ability to detect rare signals among background noise declines with time on task—a phenomenon called the *vigilance decrement*. Current theories attribute the vigilance decrement to three alternative mechanisms: shifts in bias, losses of sensitivity, and attentional lapses. Two experiments examined the extent to which the vigilance decrement reflects each mechanism. In Experiment 1, 194 participants completed an online visual signal detection task, judging whether the gap between two probe circles exceeded a criterion value. In Experiment 2, 132 participants completed a similar lab-based task. Gap size was varied across trials and data were fit with a generative cognitive model with parameters representing response bias, sensitivity, and attentional lapse rate. Parameter estimates were compared across the first and last four minutes of the vigil in a hierarchical Bayesian analysis. Data gave evidence that the decrement was driven by a conservative shift in response bias and an increased frequency of attentional lapses from the first to the last block. Although data trended in the direction of a sensitivity loss, the effects were not statistically credible.

## Mechanisms of Vigilance Loss: A Process Modeling Approach

Vigilance is the state of readiness to perceive and respond to stimuli (Mackworth, 1948) and is difficult to maintain for long durations. As such, human operators watching for rare signals among frequent noise events (e.g., baggage screeners watching for security threats or sonar operators watching for enemy activity) usually experience a decline in detections over time. This decline, known as the *vigilance decrement,* typically appears within the first 30 minutes of beginning a task, but has been observed within as little as 5 minutes (Nuechterlein et al., 1983).

Although the study of vigilance began during World War II (Mackworth, 1948) and has remained a topic of interest since, researchers have yet to agree upon the mechanisms underlying the vigilance decrement. The most common framework for analyzing vigilance data has been *signal detection theory* (SDT; Green & Sweats, 1966). SDT models an observer's ability to discriminate between signal and noise events using the relationship between correct detections and false alarms to compute separate measures of *sensitivity* and *response bias*. Both sensitivity—the observer's ability to distinguish signal from noise, and response bias—the observer's tendency to make signal or noise judgments, can explain a decline in detections over time.

SDT data suggest that the vigilance decrement is primarily driven by conservative bias shifts, whereby observers adopt progressively stricter cutoffs for making signal judgments in response to the low signal rate (Broadbent & Gregory, 1965; Craig, 1978). Analyses sometimes reveal a concurrent sensitivity decrement, whereby monitors get worse at discriminating signal from noise over time (See et al., 1995). Sensitivity losses generally only occur when task demands are high, for example, when discriminations load memory, event rates impose time pressure, or stimuli are not very discriminable (Nuechterlein et al., 1983; Parasuraman, 1979).

The selective nature of the sensitivity decrement has been explained by two theoretical accounts. The *resource depletion* theory (Parasuraman, 1979) proposes that sensitivity decrements occur when the demands on information processing resources deplete the observer’s attentional capacity, while the *resource control model* (Thomson et al., 2015) proposes that executive control failures allow attention to drift off-task. In both cases, fewer attentional resources are available to the task. Both accounts are supported by findings that vigilance tasks are subjectively effortful, and that increasing task demands reduces sensitivity (Caggiano & Parasuraman, 2004; Grier et al., 2003; Helton & Warm, 2008).

Recent work, however, suggests that yes-no signal detection data may not be suitable for understanding the vigilance decrement. First, when false alarms are near-zero, bias shifts and sensitivity losses—which usually have opposing effects on false alarms—become indistinguishable. Because of the low signal rates inherent in vigilance tasks, observers tend to adopt very conservative criteria for responding ‘signal’. This, in turn, produces low false alarm rates, allowing shifts in bias to mimic declines in the sensitivity measure, *d’* (Thomson et al., 2016).

Second, although an alternative measure, *A’*, is widely used to analyze sensitivity when false alarms are near-zero, it is not actually bias free (Getty et al., 1995; Macmillan & Creelman, 1996; Pastore et al., 2003). In fact, a recent simulation study found that shifts in bias produce spurious changes in *A’* (McCarley et al., 2021). This work suggests that much of the existing evidence for sensitivity losses in vigilance tasks may be misleading, which prompts reconsideration of the mechanisms contributing to the vigilance decrement.

Finally, SDT is unable to capture declines in detections driven by a third potential mechanism—attentional lapses (Jerison et al., 1965; Robertson et al., 1997). Lapses occur when attention is temporarily withdrawn from the task, resulting in failures to respond to stimuli entirely, regardless of signal intensity. With only two degrees of freedom, binary signal detection data are unsuitable for discriminating criterion shifts, sensitivity losses, *and* attentional lapses.

Employing a novel method of analysis to account for all three alternatives, McCarley and Yamani (2021) fit psychometric curves to detection responses in a go/noGo vigilance task. Data gave decisive evidence that the vigilance decrement was driven by losses of sensitivity, conservative shifts in response bias, and attentional lapses, with bias shifts accounting for most of the decrement.

To generalize these findings, we adapted McCarley and Yamani's (2021) vigilance task for use in an online sample. To maximize completion rates, we shortened the task from 20 minutes to 12 minutes and increased the signal rate from 0.15 to 0.20. The punctate visual noise used to decrease signal discriminability in the original experiment could not be replicated online and was substituted with distractor stimuli that were intended to serve the same purpose.

Across two online experiments (N = 103, N = 194), we replicated the conservative bias shifts and increased frequency of attentional lapses observed in the lab. However, neither experiment found conclusive evidence for or against a sensitivity loss, suggesting that sensitivity losses—though present in the lab—do not always play a significant role, even in tasks to designed to elicit sensitivity losses (i.e., those with high event rates and successive discriminations; Parasuraman & Davies, 1977).

An alternative way of analyzing these data is to fit them with a generative process model. This approach should provide more statistical sensitivity than fitting psychometric curves to the data and, since it simulates cognitive processes directly, might give more insight into the psychological processes involved in generating the observed data.

## Experiment 1

Experiment 1 re-examined data collected from a preregistered (https://osf.io/nt6u3) online adaptation of McCarley and Yamani's (2021) vigilance task. The data were originally analyzed using psychometric curves and are reported in Experiment 2 of Chapter 3.

### *Method*

**Participants.** Data were analyzed from 194 participants recruited from the online research platform Prolific (www.prolific.co). Inclusion and exclusion criteria are detailed on page 61.

**Apparatus and stimuli.** The stimulus each trial was a pair of red probe circles embedded in five, black distractor circles. Stimuli were presented on a white background, within a circular search field 8 cm diameter. The two probe circles were arranged horizontally, separated by a distance (*d*) that varied each trial between 0.5 cm and 3.25 cm, in steps of 0.25 cm. Full details are reported on pages 51-52.

**Procedure.** Participants viewed a series of probe circles and judged whether the gap between each pair of probes exceeded 2 cm. The distance between probe circles varied across trials between 0.5 cm and 3.25 cm, in 0.25 cm increments. Participants were asked to press the space bar if the gap between probe circles exceeded the criterion distance on a given trial, and to withhold response otherwise. The mean signal rate was 0.20. Each trial comprised a 250 ms stimulus display followed by a blank interval of 1,250 ms, during which only the outline of the search field remained visible. The subsequent trial began immediately thereafter, producing an event rate of 40 trials per minute. Participants first completed a practice vigil of 90 trials, followed by a 12-minute experimental vigil. The full procedure is reported on pages 52-54.

### *Analysis*

Participants' binary responses from the 12-minute vigil were converted to the signal detection theory measure of sensitivity, *d'.* To correct for ceiling- or floor-level hit and false alarm rates, *d′* scores were calculated using a log-linear correction (Hautus, 1995). Of the 197 participants whose data were saved, three were excluded for failing to achieve a *d'* of at least 0.25 in first, middle, or last 4-minutes of the task. With these exclusions, mean *d'* = 2.33, mean hit rate = 0.86, and mean false alarm rate = 0.16.

We used hierarchical Bayesian parameter estimation (Kruschke, 2015; Lee, 2018; Lee & Wagenmakers, 2013) to assess changes in vigilance between the first and last 4-minutes of the task. Signal detection responses were fit with a hierarchical process model that estimated parameters representing response bias, sensitivity, and attentional lapses at the individual- and group-level. Mean differences in group-level parameters between the first and last block were modeled with normalized effects and unit normal priors. The model placed priors on the log of internal noise to ensure positive variance values, and on probit-transformed lapse rate rather than lapse rate directly for ease of comparison across parameters. Priors for group level estimates were assigned the following values:

mu_lognoise2 ~ normal (0, .001)

sigma_lognoise2 ~ uniform (0, 100)

mu_c ~ normal (0, .001)

sigma_c ~ uniform (0, 100)

mu_probitlapse ~ normal (0, 1)

sigma_probitlapse ~ uniform (0, 100)

Priors on effects of block were assigned the following values:

mu_delta_lognoise2 ~ normal (0, .001)

sigma_delta_lognoise2 ~ uniform (0, 100)

mu_delta_c ~ normal (0, .001)

sigma_delta_c ~ uniform (0, 100)

mu_delta_probitlapse ~ normal (0, .001)

sigma_delta_probitlapse ~ uniform (0, 100)

We report mean parameter estimates with 95% Bayesian credible intervals (BCIs) representing the middle (i.e., most credible) 95% values from the posterior distribution. Effects are considered statistically credible if the BCI excludes zero. All analyses were conducted in R (R Core Team, 2019). All analyses were conducted in R (R Core Team, 2019). Estimation procedures ran four MCMC chains for 1000 warmup trials, followed by 12,500 sample steps each, using the JAGS package (Plummer, 2015). Chains were thinned to every fourth step. All parameter estimates showed $\hat{\mathrm{R}}$ convergence values of < 1.1, indicating satisfactory convergence of MCMC chains.

***Results***

Figure 13 shows the proportion of trials on which participants responded for the first and last 4-minute blocks of the vigilance task. Visual inspection of Figure 1 shows that correct detections (i.e., signal trials on which participants made a “yes” response) decreased over time, indicating a vigilance decrement.

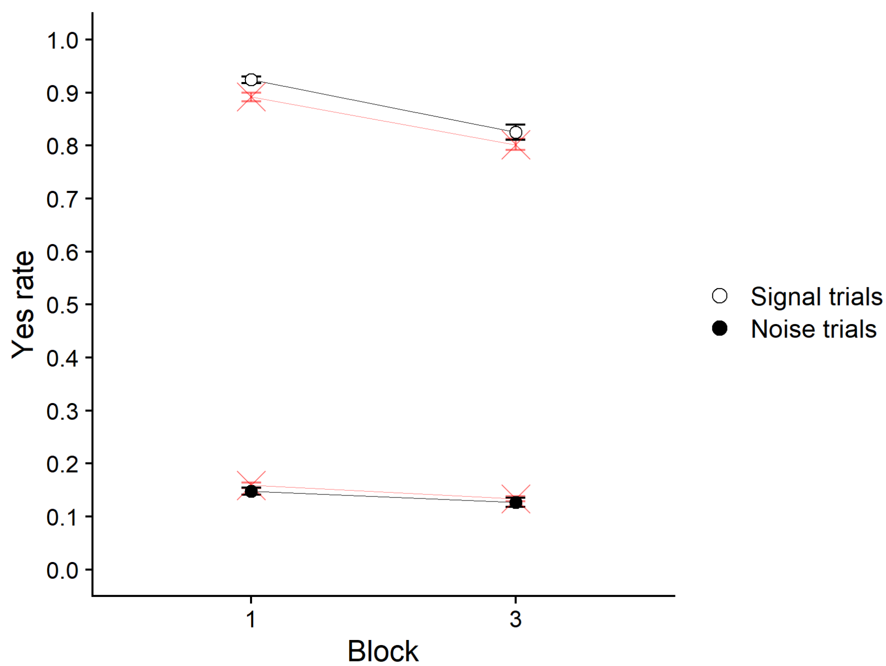


*Figure 13.* Response rates for the first and last 4-minute blocks of the vigilance task. Black and white symbols are empirical means with error bars representing standard errors. Red symbols are posterior predictive values with error bars

representing 95% BCIs.

Figure 14 shows the posterior distributions of the mean differences in the response criterion, the log of internal noise, and the probit-transformed lapse rate between the first and last blocks.

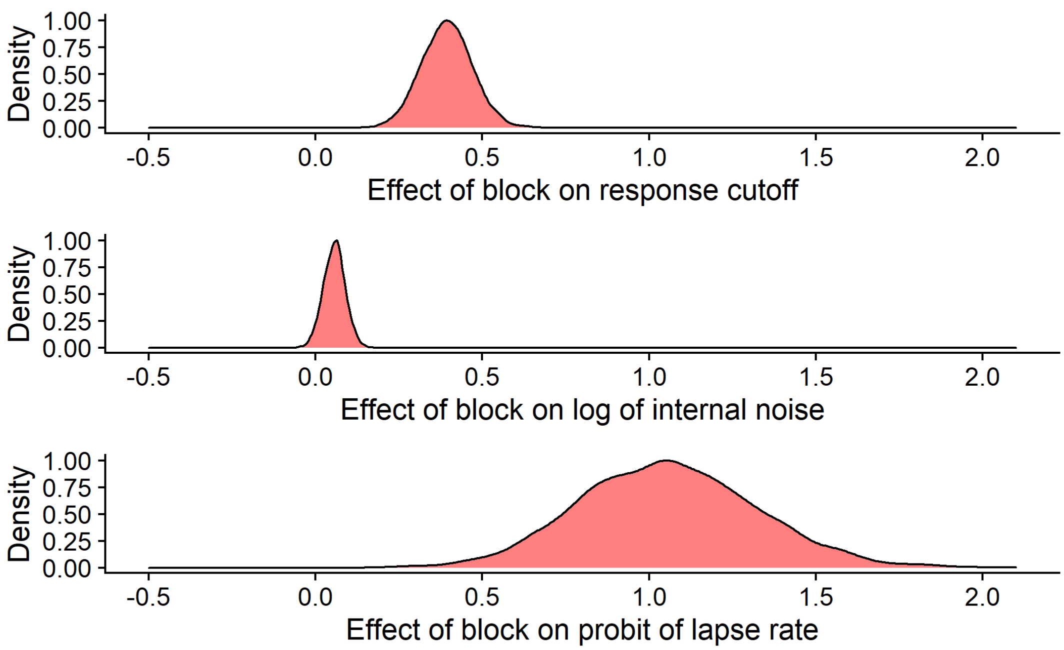


*Figure 14.* Posterior distributions of effects of block on three parameters: response cutoff (top), log of internal noise (middle), and the probit of lapse rate (bottom).

The top panel shows that the cutoff for responding to stimuli (i.e., a "signal present" judgment) was higher in the last block of the task than the first, *M* = 0.65, BCI = [0.60, 0.71]. The middle panel shows a trend toward an increase in internal noise between blocks, *M* = 0.06, BCI = [-0.01, 0.12], but the effect is not credible. The bottom panel shows an increased rate of attentional lapses in the last block, *M* =

1.06, BCI = [0.56, 1.6]. The difference between blocks was larger for lapse rate than for response cutoff or internal noise.

### *Discussion*

The current analysis compared the extent to which three proposed mechanisms of vigilance loss contributed to the vigilance decrement in an online signal detection task. Parameter estimates revealed credible effects of block on response cutoff and attentional lapse rate. Both effects were associated with a decrease in response rate, indicating that monitors adopted a more conservative response bias and experienced a greater number of attentional lapses with time on task. There was also a trend toward an increase in internal noise over time. More internal noise corresponds to a greater degree of overlap between internal representations of signal and noise events, thereby reducing observers' sensitivity. However, the 95% BCI includes negative values, suggesting that the effect is not credibly different from zero. This reanalysis provides converging evidence that the vigilance decrement was primarily driven by bias shifts and attentional lapses in this online task.

## Experiment 2

To further test the adequacy of the process model, Experiment 2 reanalyzed data from a preregistered lab-based replication of McCarley and Yamani's (2021) study. As in McCarley and Yamani's original lab-based study and our online adaptation, participants were asked to respond when the gap between pairs of probes exceeded 2 cm. The preregistered replication data were originally analyzed using psychometric curves. Figure 15 shows the posterior density plots for normalized effects of block on the shift, scale, and lapse rate parameters. Data gave decisive evidence for an effect of block on lapse rate,*B10* = 3700.4, but only anecdotal evidence for effects on response bias,*B10* = 2.21, and sensitivity,*B10* = 0.37.

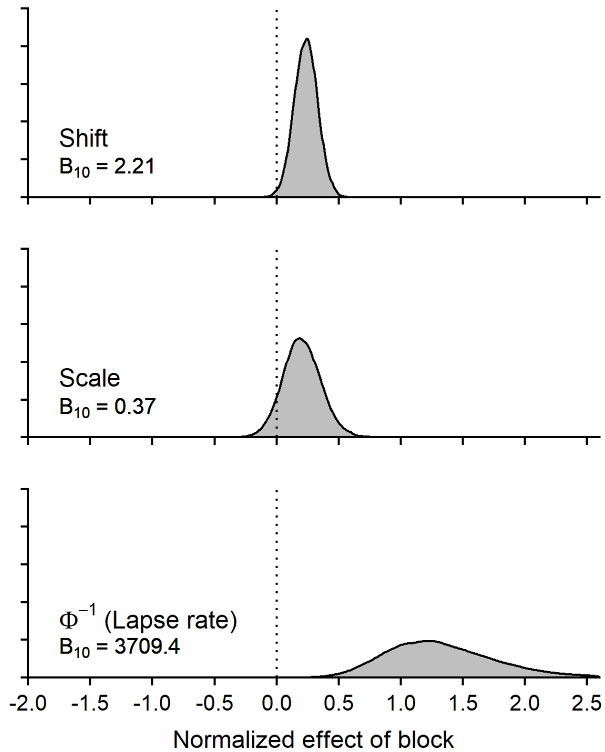


*Figure 15.* Posterior distributions of effects of block on three parameters: shift, corresponding to response bias (top), scale, corresponding to sensitivity (middle), and the probit of lapse rate, corresponding to attentional lapses (bottom).

Here, we reanalyze these data using the cognitive process model from Experiment 1 to estimate changes in response bias, sensitivity, and attentional lapse rate between the first and last 4-minutes of the task.

***Method***

**Participants.** One hundred thirty-two undergraduate students were recruited from two large public universities in the United States. Inclusion criteria were fluency in English, normal color vision, and normal or corrected-to-normal visual acuity. One participant was excluded for failing to achieve a *d'* score of ≥ 0.25 in each 4-minute block of the task. All methodological details, including exclusion criteria, were preregistered at osf.io/3np2v/. Participants received 10 USD/hour for an experimental session designed to be completed in < 30 minutes.

**Apparatus and stimuli.** Experimental software was written in PsychoPy 3 (Peirce et al., 2019) and was identical to that used by McCarley and Yamani (2021). The stimulus each trial was a pair of red probe circles, but unlike in the previously reported online tasks, probes were embedded in dynamic Gaussian visual noise rather than distractor circles. Probes were unfilled, drawn in 3-pixel stroke, with a diameter of 2 mm. The two probe circles were arranged horizontally, separated by a distance (*d*) that varied each trial between 0.75 cm and 3.5 cm, in steps of 0.25 cm. Stimuli were presented on a 24-inch LED monitor with a resolution of 1024 x 768 pixels and a refresh rate of 75 Hz, and on a 24-inch LED monitor with a resolution of 1920 x 1080 pixels and a refresh rate of 60 Hz, depending on the site of data collection.

**Procedure.** Each trial, participants judged whether the horizontal separation between two probe circles exceeded a target criterion of 2 cm. They were asked to press the space bar if the gap between probe circles exceeded the criterion distance on a given trial, and to withhold response otherwise. The value of *d* for a given trial was determined probabilistically with each trial randomly designated to be a signal or noise event. The probability of signal was 0.15. Gap size was then selected randomly and with uniform probability from the range of values corresponding to the trial type. Non-signal values were 0.75, 1.25, and 1.75. Signal values were 2.25, 2.75, 3.25, and 3.75.

On each trial, stimuli were displayed for 500 ms, followed by a blank screen for 1000 ms. Figure 16 depicts the sequence of events. The next trial began immediately thereafter, producing an event rate of 40 trials per minute. A response was attributed to trial *i* if it occurred before the onset of trial *i* + 1. Participants did not receive post-trial feedback on experimental trials.

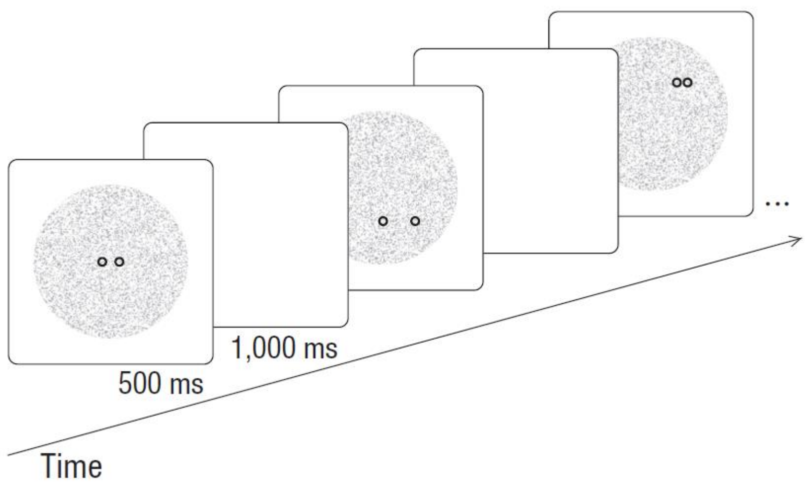


*Figure 16.* Sequence of events over a series of trials. Not to scale. Reproduced with permission from McCarley and Yamani (2021).

Each participant began by completing a 3-minute block of practice trials. The practice vigil was the same as the experimental vigil except that 1) signal and noise events were equally probable, 2) for the first 25 practice trials, the stimulus display remained visible for the full 1,500 ms duration, and 3) response errors were followed by a 1-second feedback message reading either, “Oops! It was not a target.”, or “Oops! You missed a target.”, as appropriate. Following practice, participants completed the 20-minute vigil. Immediately after the experimental trials, participants were asked to complete the abbreviated Subject Workload Assessment Technique (ASWAT)  in which they provided ratings of mental effort, time load, and psychological stress.

***Results***

Figure 17 shows that correct detections (yes responses on signal trials) decreased over time, indicating a vigilance decrement.

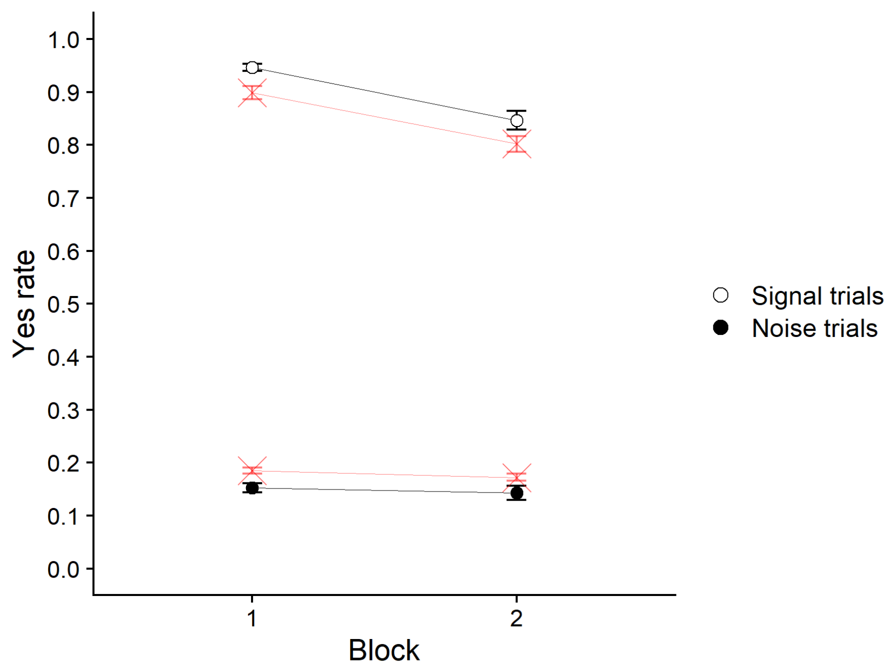


*Figure 17.* Response rates for the first and last 4-minute blocks of the vigilance task. Black and white symbols are empirical means with error bars representing standard

errors. Red symbols are posterior predictive values with error bars representing 95% BCIs.

Figure 17 shows the posterior distributions of the mean differences in the response criterion, the log of internal noise, and the probit-transformed lapse rate between the first and last blocks.

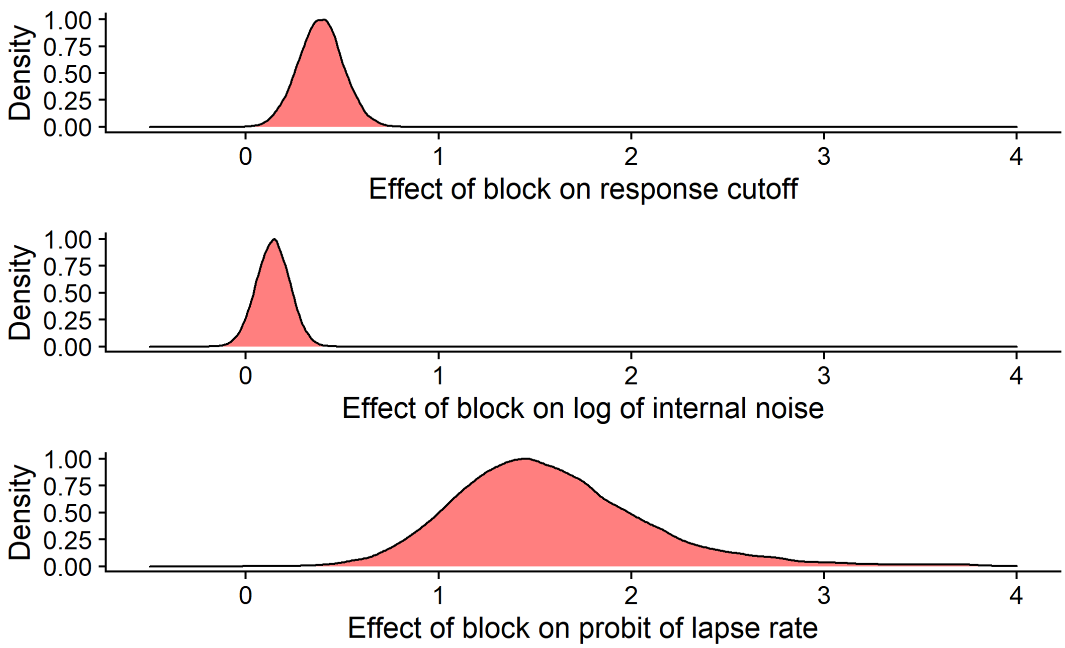


*Figure 17.* Posterior distributions of effects of block on three parameters: response cutoff (top), log of internal noise (middle), and the probit of lapse rate (bottom).

The top panel shows that observer's cutoff for making a signal present response was higher in the last block of the task than in the first, *M* = 0.39, BCI = [0.16, 0.62]. The middle panel shows a noncredible trend toward an increase in internal noise between blocks, *M* = 0.14, BCI = [-0.03, 0.31]. The bottom panel shows an increased rate of attentional lapses in the last block compared to the first, *M* = 1.59, BCI = [0.76, 2.79]. Again, the effect of block was largest for attentional lapse rate, followed by response cutoff.

### *Discussion*

The current analysis examined changes in response cutoff, internal noise, and lapse rate in a lab-based sensory vigilance task. Parameter estimates revealed credible effects of block on response cutoff and attentional lapse rate, indicating that monitors adopted a more conservative response bias and experienced more attentional lapses as the task progressed. Although there was a trend toward an increase in internal noise between blocks, the 95% BCI included negative values indicating that the effect was not credibly different from zero. These results are consistent with those revealed by the psychometric curve analysis and demonstrate that the vigilance decrement was largely driven by an increase in the rate of attentional lapses.

## General Discussion

In the present study, we fit data from two previous experiments (N = 194 and N = 132) with a generative cognitive model to test the model's ability to capture changes in response bias, sensitivity, and attentional lapses. Experiment 1 re-examined data from the preregistered online sensory vigilance task reported in Chapter 3. Consistent with results from the original psychometric curve analysis, the generative model revealed credible effects on response cutoff and lapse rate, indicating that conservative shifts in response bias and an increase in the frequency of attentional lapses both contributed to the vigilance decrement.

Experiment 2 re-examined data from a preregistered lab-based sensory task directly replicating McCarley and Yamani (2021). The replication is not reported in this dissertation. However, consistent with its findings, the generative cognitive model revealed credible effects on response cutoff and lapse rate. These results indicate again that the observed vigilance decrement was driven by monitors adopting a more conservative response bias and experiencing more attentional lapses with time on task. More broadly, Experiments 1 and 2 demonstrate that the generative model discriminates between the three potential mechanisms of vigilance loss and captures the same trends in parameter estimates as the psychometric curves while providing greater statistical sensitivity.

**Chapter 5: Mechanisms of Vigilance loss in an Online Cognitive Task**

Using the generative cognitive model described and tested in Chapter 4, Chapter 5 assesses the extent to which bias shifts, losses of sensitivity, and attentional lapses contribute to the vigilance decrement in an online cognitive vigilance task. While sensory vigilance tasks require discriminations of perceptual features (e.g., the separation two circles), cognitive tasks require discriminations of the semantic meaning or numeric value of stimuli. Cognitive vigilance tasks typically show smaller vigilance decrements over time than sensory tasks, but interaction effects are inconsistent and the reasons for more stable performance are unclear. To date, no prior research has compared all three potential mechanisms of vigilance loss within a cognitive task.

**Abstract**

In vigilance tasks, a common finding is that monitors experience a progressive decline in detections over time called the vigilance decrement. However, most vigilance research has examined performance within sensory tasks, which require judgments of perceptual features of stimuli. The vigilance decrement appears to be less consistent within cognitive vigilance tasks, which require judgments of the meaning or value of alphanumeric stimuli. The current experiment (N = 180) examined the extent to which changes in response bias, sensitivity, and attentional lapses contributed to the vigilance decrement in a cognitive vigilance task. Each trial participants viewed a set of four three-digit numbers and judged whether they were drawn from a distribution with a mean of 490 (noise) or 510 (signal). Data gave substantial evidence against a shift in response bias and strong evidence against a sensitivity loss. Although the vigilance decrement was trivial compared to decrements observed in prior sensory tasks, data gave substantial evidence for an increase in attentional lapse rate over time.

## Chapter 5: Mechanisms of Vigilance Loss in an Online Cognitive Task

*Vigilance tasks* like quality control and security surveillance require observers to monitor for infrequent signals over extended periods of time. A common finding is that vigilance deteriorates, producing a progressive decline in detections termed the *vigilance decrement* (Mackworth, 1948; Nuechterlein et al., 1983; Parasuraman, 1979; See et al., 1995; Warm et al., 1996). The vigilance decrement generally occurs within 30 minutes (Mackworth, 1948), but under some circumstances, can begin within 5 minutes of starting a task (Nuechterlein et al., 1983). Despite considerable research dedicated to mitigating the vigilance decrement, the mechanisms underlying the effect remain the subject of debate.

Typically, vigilance studies employ yes-no detection tasks in which operators judge each trial whether a signal is present (yes) or absent (no). As such, responses can be evaluated using signal detection theory (SDT; Green & Swets, 1966) to isolate the effects of sensitivity and response bias on performance. SDT analyses have suggested that the vigilance decrement largely reflects a tendency for operators to adopt a progressively more conservative response criterion throughout the task, such that they become less willing to classify events as signals over time (Broadbent & Gregory, 1965; Broadbent, 1971; Swets, 1973, 1977). This shift toward more cautious responding is assumed to represent a change in the observer’s signal expectancy (Parasuraman, 1979; Parasuraman & Mouloua, 1987), and reduces both hits (true positive detections) and false alarms (false positive detections).

When tasks impose a high demand on attentional resources, SDT analyses may also reveal a decline in sensitivity, whereby operators become worse at discriminating signals from background noise. Sensitivity losses tend to occur in tasks that place heavy demands on attentional resources, such as those combining a high memory load with time pressure or poor signal discriminability (Nuechterlein et al., 1983; Parasuraman, 1979; See et al., 1995; Smit et al., 2004). Theoretical accounts attribute the selective occurrence of sensitivity losses to gradual reductions in the attentional resources allocated to the vigilance task. *Resource depletion theory* proposes that maintaining vigilance is mentally taxing (Grier et al., 2003; Warm et al., 2008) and when demand on attentional resources exceeds availability, the pool of resources shrinks (Caggiano & Parasuraman, 2004; Warm et al., 1996). Alternatively, *resource control theory* argues that the pool of resources remains constant in size, but that executive control wanes over time, allowing resources drift to task-unrelated thoughts (Thomson et al., 2015). Under either model, resources dedicated to the vigilance task dwindle over time, resulting in poorer sensitivity.

Although signal detection theory distinguishes between two potential mechanisms of vigilance decrement—bias shifts and sensitivity losses—an alternative theory proposes that the decrement is driven by attentional lapses (Jerison et al., 1965; Manly et al., 1999; Robertson et al., 1997). The mindlessness theory holds that that the monotonous nature of vigilance tasks (Scerbo, 1998) encourages observers to respond in an automatic, routinized manner, resulting in brief periods in which attention is fully decoupled from the vigilance task. Unlike periods of mind-wandering, which reduce response rates for weaker, more ambiguous signals, periods of mindlessness reduce response rates regardless of signal intensity.

Support for the mindlessness theory comes from studies employing the Sustained Attention to Response Task (SART; Robertson et al., 1997): a modified vigilance task in which observers respond to frequent noise events and withhold responses to infrequent signals to promote rapid response automatization. Research with the SART has shown that detection failures are preceded by periods of faster, presumably more automatic, responding (Dockree et al., 2004; Manly et al., 1999), and that trait absentmindedness is positively associated with number of detection failures (Manly et al., 1999; Robertson et al., 1997). However, the vigilance decrement does not appear to be solely the result of mindless responding; subjective workload data indicate that both traditional vigilance tasks and the SART are effortful (Dillard et al., 2015; Grier et al., 2003). It remains a possibility, though, that the decrement is at least partially driven by attentional lapses.

A recent study accounting for all three proposed mechanisms found that conservative shifts in response bias, attentional lapses, and—to a lesser extent—losses of sensitivity, all contributed to the vigilance decrement (McCarley & Yamani, 2021). An online adaptation (Gyles et al., 2022) replicated the effects of conservative bias shifts and attentional lapses but did not find strong evidence for or against a sensitivity loss. Together, these findings suggest that sensitivity losses are possible, both potentially less robust than bias shifts and lapses.

Within the vigilance literature, researchers often make a distinction between sensory and cognitive tasks (Davies & Tune, 1969, See et al., 1995). Most previous research has examined vigilance performance within sensory tasks, which commonly require observers to detect perceptual features of visual or auditory stimuli (Deaton &

Parasuraman, 1993). Participants in a sensory vigilance task, for example, might be asked to detect larger than usual jumps in a clock hand (Mackworth, 1948), increases in the brightness of a light (Broadbent & Gregory, 1965) or amplitude of a tone (Cahoon, 1973), or differences in the color of a disk (Chan & Chan, 2022). Sensory tasks are also common in applied settings. Consider the earlier examples of baggage screening, quality control, and security surveillance; each of these tasks requires the operator to monitor perceptual features like size, shape, and movement to detect a signal.

Less research has examined vigilance performance in cognitive tasks, which require observers to judge the meaning or value of alphanumeric stimuli (Deaton & Parasuraman, 1993). Of the relatively small number of studies that have examined cognitive vigilance, many have asked participants to perform simple arithmetic operations on pairs of digits (Claypoole & Szalma, 2018; Claypoole et al., 2019; Warm et al., 1984) (e.g., whether paired digits differ by more or less than 1) or to monitor a stream of digits for successive odd numbers (Matthews et al., 1990). These discriminations are abstractions of real-world tasks in which operators must monitor displays of symbolic information. For example, combat aircraft pilots must process strings of letters and numbers to identify aircraft as friendly or hostile (Deaton & Parasuraman, 1993), and cyber security officers must monitor network activity to detect keywords, IP addresses, and changes in network traffic (McIntire et al., 2013).

A meta-analysis examining conditions that produce sensitivity losses (See et al., 1995) revealed that the sensory-cognitive dimension explains variation in whether, and to what extent, sensitivity declines in the vigilance decrement. However, performance differences between sensory and cognitive tasks are inconsistent. The meta-analysis revealed that sensitivity losses are generally smaller in cognitive tasks than in sensory tasks, except at very high event rates, in which case the opposite pattern is observed. This effect was also dependent on whether the tasks required simultaneous or successive discriminations. Within cognitive tasks, event rate was positively associated with the magnitude of sensitivity loss for simultaneous tasks but had little effect on successive tasks. In contrast, a recent study examining the effects of event rate on cognitive vigilance found that event rate was positively associated with the magnitude of sensitivity losses even in a successive cognitive task (Claypoole et al., 2019).

Other studies have reported a vigilance *increment* in cognitive tasks, whereby target detection rates increased over the course of the vigil (Deaton & Parasuraman, 1993; Dember et al., 1984; Lysaght et al., 1984; Noonan et al., 1985; Warm et al., 1984). Although many of these studies did not report signal detection measures of sensitivity and bias, See and colleagues (1995) found that effect sizes for sensitivity losses were negative at low event rates, suggesting that previously reported vigilance increments could reflect increases in sensitivity. However, many studies attempting to replicate the cognitive vigilance increment have failed to find an increase in detections over time (Ash et al., 1983; Loeb et al., 1982; Noonan et al., 1984; Noonan et al., 1985), while others have seen temporal improvements only in response times (Ash et al., 1985; Dember et al., 1984).

Although vigilance decrements and sensitivity losses are generally less consistent in cognitive tasks than in sensory tasks, the reason for this difference is not well understood (Deaton & Parasuraman, 1993; See et al., 1995). Researchers have suggested that sensitivity might be more stable in cognitive tasks because alphanumeric stimuli are highly familiar, and therefore impose a lower load on information processing resources than unfamiliar sensory stimuli (Koelega et al., 1989; Warm et al., 1984). However, cognitive tasks are sometimes rated as more demanding than sensory tasks (Claypoole et al., 2019; Deaton & Parasuraman, 1993), even when they do not demonstrate a vigilance decrement, indicating that cognitive tasks do not universally impose a lower workload than sensory tasks.

Warm and colleagues (1984) have also speculated that observers might be more motivated to attend to cognitive tasks because the well-learned stimuli make it easy for observers to recognize when they have made an error. An increase in motivation and task engagement might prevent attentional resources from drifting to task unrelated thoughts. But this hypothesis was partially ruled out by a study in which participants made sensory and cognitive discriminations of an identical set of numerical stimuli (Deaton & Parasuraman, 1993). The sensory task required discrimination of digit size, with signals defined as pairs of digits in which one digit was physically smaller than the other. The cognitive task, on the other hand, required discrimination of digit values, with signals defined as pairs containing one even and one odd digit. Detections declined over time in the sensory task but remained stable in the cognitive task, despite initial performance levels being similar. Although this effect cannot be attributed to differences in stimuli, these results do not rule out the possibility that noticing errors is easier for cognitive judgments than for sensory judgments.

We propose another potential explanation: that performance differences observed in sensory and cognitive vigilance tasks might be driven by differences in whether the tasks involve continuous- or discrete-state judgments. Signal detection theory assumes that observers make judgments by comparing a continuous decision variable to a cutoff to render a binary response (Green & Swets, 1966), and therefore models a continuous-state decision process. In sensory tasks, the decision variable (e.g., size or brightness of a stimulus) is subject to both external noise (i.e., uncertainty inherent in the data) and internal noise (i.e., variation in the observer's neural response to stimuli). As such, it is often ambiguous whether the extracted evidence represents a signal or noise event (MacMillan & Creelman, 2005).

Cognitive vigilance tasks, however, tend to map stimuli to discrete rather than continuous states. For instance, when observers judge whether the difference between two digits is equal to 1, or whether a digit is odd or even, a stimulus is unambiguously signal or noise. So long as the stimuli are well above sensory thresholds for detection and recognition, performance will be unaffected by graded internal noise, and judgments are unlikely to be subject to drifts of response bias.

To further explore mechanisms of the vigilance decrement in cognitive tasks, the current study analyzed vigilance performance in a signal detection task requiring continuous-state judgments of numerical value (Healy & Kubovy, 1981). The task was framed as a cyber monitoring task in which participants monitored a display of four briefly presented numbers and provided a keypress response when their average value exceeded a criterion. The numbers presented on signal and noise trials were drawn from distributions with means of 510 and 490, respectively, each with a standard deviation of 20. On average, signal trials tended to produce values greater than 500, but the overlap between signal and noise distributions introduced variability, such that numbers drawn from the signal distribution could sometimes exceed numbers drawn from the noise distribution. Further, the brief presentation of the numbers did not allow observers to calculate the precise arithmetic mean, but rather, required rapid estimation, introducing uncertainty in the mapping of the stimuli to the observer's decision variable. Therefore, unlike cognitive tasks requiring discrete-state judgments (e.g., whether a number is odd or even), performance in this cognitive task was subject to both external and internal decision noise, allowing examination with a continuous-state decision model and comparison with existing sensory vigilance data.

The conventional way of analyzing these data is to estimate sensitivity and response bias within the framework of signal detection theory. However, recent work suggests that binary detection data are unsuitable for distinguishing mechanisms of vigilance loss. Though many studies report decreases in the SDT measure of sensitivity *d'*, this decline may merely be the result of low false alarm rates. When false alarm rates are at or near floor, bias shifts and sensitivity losses—which usually have opposing effects on false alarms—become indistinguishable, allowing bias shifts to mimic declines in *d'* (Thomson et al., 2016). Bias shifts also produce spurious changes in *A'* (Pollack & Norman, 1964), an alternative measure of sensitivity often preferred by vigilance researchers when false alarms are low (McCarley et al., 2021). Even when false alarm rates are high enough to allow researchers to distinguish changes in bias from changes in sensitivity, yes-no detection data are limited by two degrees of freedom. Binary signal detection tasks are therefore unable to account for the third potential mechanism of vigilance loss, attentional lapses.

An alternative approach is to fit a generative cognitive process model (Haines et al., 2020) to responses to estimate bias, sensitivity, and lapse rate parameters directly. Here, participants were asked to estimate the mean of a set of numbers each trial, and to judge whether the numbers were sampled from signal or noise distribution (Healy & Kubovy, 1981). To model the task, we assumed that operator's decision variable was equal to the arithmetic mean plus noise (Brusovansky et al., 2018). To analyze changes in bias, sensitivity, and lapse rate over time, the current study employs Bayesian hierarchical modeling to compare parameter estimates for the first and last 4 minutes of a 12-minute vigil, using the generative model tested in Chapter 4.

## Method

### *Participants*

Two hundred participants were recruited from the online research platform Prolific (https://prolific.co/). All participants gave informed consent and self-reported fluency in English, normal color vision, and normal or corrected-to-normal visual acuity. Data were excluded from participants who failed to complete the full experimental session or to achieve $d'$ scores of $\geq 0.25$ in each 4-minute block of the task. Exclusions left 180 participants for analysis (Mage = 22.58 years, gender = 73 females, 100 males, 4 non-binary, 3 not specified). Participants were reimbursed at US$10.00/hour for an experimental session lasting approximately 25 minutes.

***Apparatus and stimuli***

The experimental task was controlled by software written in PsychoPy 3 (Peirce et al., 2019) and hosted on Pavlovia (https://pavlovia.org/).The stimulus each trial was a column of four three-digit numbers, presented in black text on a white background.

***Procedure***

Participants performed a signal detection task in which they judged whether the average value of a set of numbers each trial was sampled from a distribution with a mean of 490 or 510. Participants were asked to press the space bar if they thought the average value was drawn from the signal distribution (i.e., with a mean of 510) and withhold response otherwise. Participants first completed a practice vigil of 90 trials, followed by a 12-minute experimental vigil. To avoid potential end-spurt effects (Bergum & Lehr, 1963) participants were not informed of vigil length.

Whether a trial represented a signal or noise event was determined probabilistically, with experimental trials assigned as signals with a probability of 20%. If a trial was determined to be a signal, the four numbers were drawn randomly from a normal distribution with $M = 510$ and $SD = 20$. If a trial was determined to be noise, the four numbers were drawn randomly from a normal distribution with $M = 490$ and $SD = 20$.

Experimental trials comprised a 1500 ms stimulus display (see Figure 18), during which time participants could provide a keypress response. The subsequent trial began immediately thereafter, producing an event rate of 40 trials per minute. Participants did not receive post-trial feedback on experimental trials.

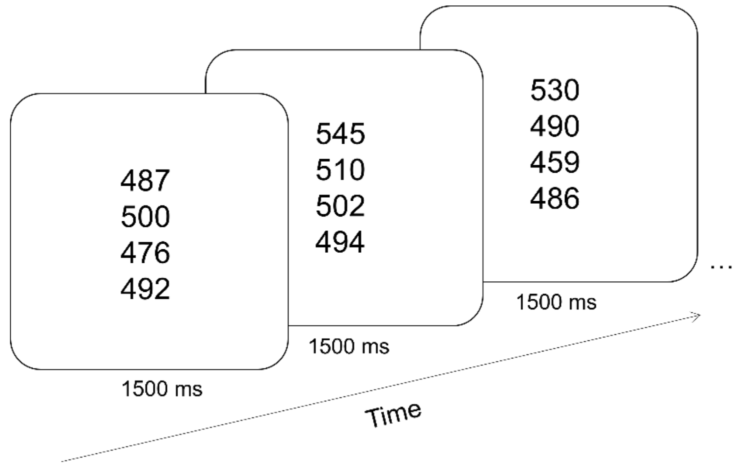


*Figure 18.* Sequence of events in the cognitive vigilance task. Not to scale.

The practice vigil was the same as the experimental vigil except that 1) the signal rate was 0.50, 2) for the first 25 trials of the vigil, the stimulus display remained visible for 3000 ms, and 3) response errors were followed by a 1-second feedback message reading either, “Oops! It was not a target.”, or “Oops! You missed a target.”, as appropriate. Error-free performance resulted in a practice vigil of 2 minutes 15 seconds and each error added 1 second.

Upon task completion, participants completed a computerized A-SWAT mental workload scale (Luximon & Goonetilleke, 2001) consisting of three subscales: time load, mental effort, and psychological stress. Participants made their rating of each subscale, presented one at a time, by clicking a horizontal line anchored with the text descriptions of subscale endpoints.

**Analysis**

To exclude participants who might have stopped attending to the task entirely, we converted participants' binary responses to the sensitivity measure *d'* using a log-linear correction (Hautus, 1995) to correct for floor and ceiling level hit and false alarm rates. Twenty participants who failed to achieve a *d'* of at least 0.25 in the first, middle, or last 4-minutes of the task were excluded from further analyses. With these exclusions, mean *d'* was 1.35, mean hit rate was 0.75, and mean false alarm rate was 0.30, collapsed across experimental blocks.

We used hierarchical Bayesian parameter estimation (Kruschke, 2015; Lee, 2018; Lee & Wagenmakers, 2013) to assess changes in vigilance between the first and last 4-minutes of the task. Signal detection responses were fit with a hierarchical generative model that estimated parameters representing response bias, sensitivity, and attentional lapses at the individual- and group-level. Mean differences in group-level parameters between the first and last block were modeled with normalized effects and unit normal priors. The model placed priors on the log of internal noise to ensure positive variance values, and on probit-transformed lapse rate rather than lapse rate directly for ease of comparison across parameters. Prior values are reported in Chapter 4 (p. 79).

We report mean parameter estimates with 95% Bayesian credible intervals (BCIs) to describe the posterior distributions, and Bayes factors to summarize the strength of evidence for (or against) each parameter. We calculated Bayes factors from the Savage-Dickey density ratio (Wagenmakers et al., 2010): the height of the posterior distribution divided by the height of the prior distribution at the parameter value of interest (i.e., 0). The resulting Bayes factor, denoted *B10*, is the ratio of the likelihood of the data under the alternative hypothesis versus the null. A ratio of 1 indicates no evidence in either direction, values greater than 1 support the alternative hypothesis, and values between 0 and 1 support the null. We interpret the strength of evidence in line with guidelines proposed by Jeffreys (1961) (i.e., 1-3 = anecdotal, 3-10 = substantial, 10-30 = strong, and >100 = decisive evidence for the alternative hypothesis, compared to the null hypothesis).

Mean ratings for each of the A-SWAT subscales were estimated separately within a hierarchical model that placed a normal likelihood function on observed ratings, and uniform priors, U(1, 100), on the group means and standard deviations of the ratings.

All analyses were conducted in R (R Core Team, 2019). Estimation procedures ran four MCMC chains for 1000 warmup trials, followed by 12,500 sample steps each, using the JAGS package (Plummer, 2015). Chains were thinned to every fourth step. All parameter estimates showed $\widehat{R}$ convergence values of < 1.1, indicating satisfactory convergence of MCMC chains.

**Results**

Figure 19 shows the proportion of responses for signal and noise trials in the first and last blocks of the cognitive vigilance task. Visual inspection shows suggests that the response rate remained relatively stable between blocks.

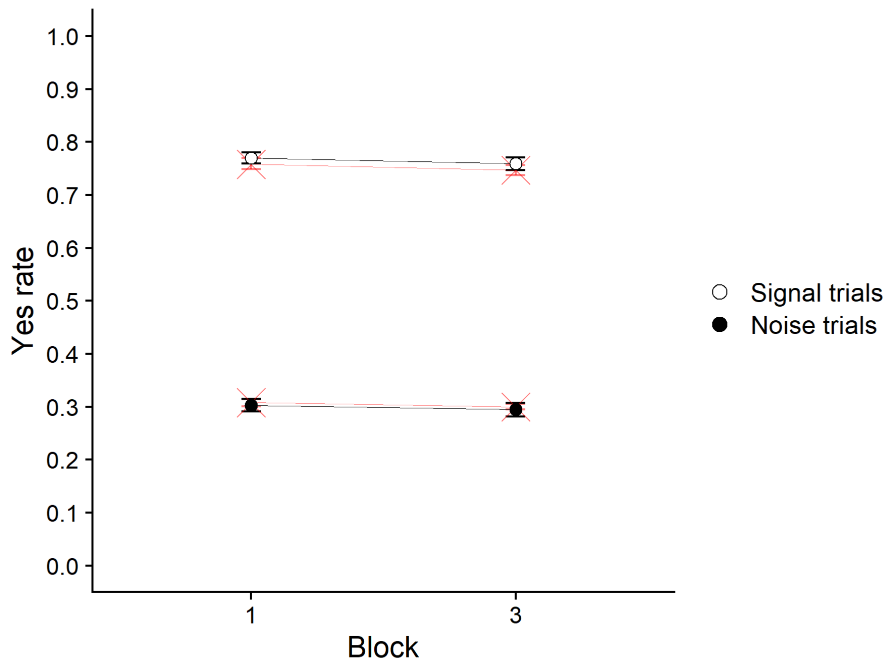


*Figure 19.* Response rates for the first and last 4-minute blocks of the vigilance task for signal and noise trials. Black and white symbols are empirical means with error bars representing standard errors. Red symbols are posterior predictive values with error bars representing 95% BCIs.

Figure 20 shows the posterior distributions of the mean differences in the response criterion, the log of internal noise, and the probit-transformed lapse rate between the first and last blocks. The top panel shows that the posterior distribution for the effect of block on response cutoff was highly variable but not credibly different from zero, *Mdiff* = 0.03, BCI = [-0.54, 0.60]. Data gave substantial evidence against a shift in response cuttoff from the first to last block,*B10* = 0.29. Data in the middle panel give strong evidence,*B10* = 0.06, against an effect of block on internal noise, *Mdiff* = -0.03, BCI = [-0.13. 0.07]. Data in the bottom panel give substantial evidence,*B10* = 3.90, in favor of an effect of block on lapse rate. The change in lapse rate was associated with a decrease in response rate from the first to last block, indicating an increase in attentional lapses over time, *Mdiff* = 0.49, BCI = [0.09, 0.95].

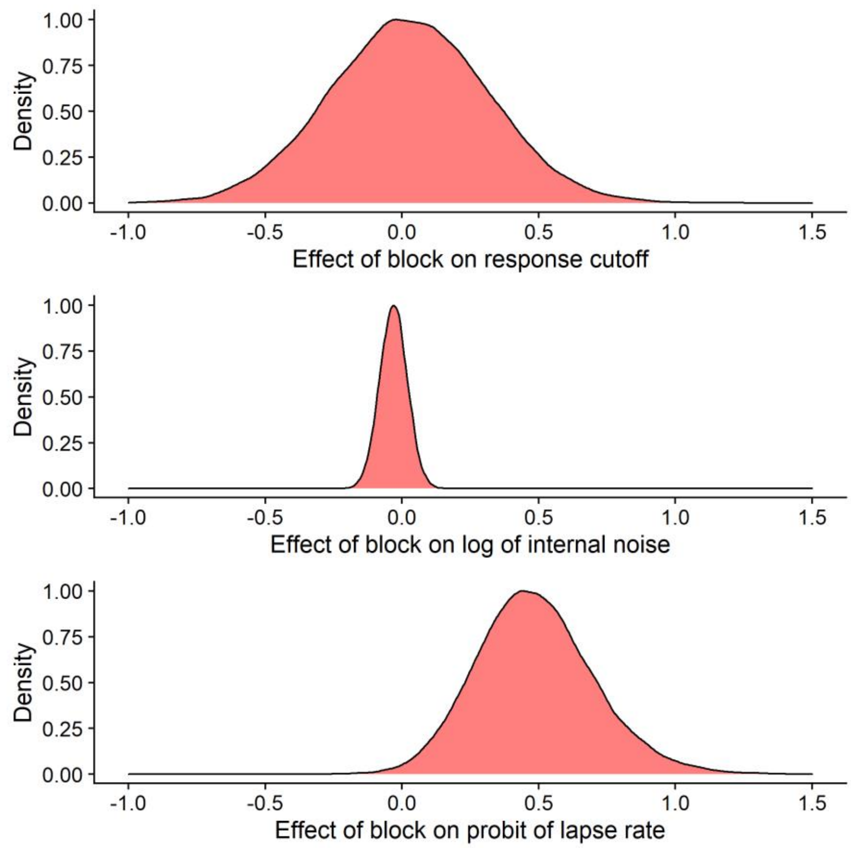


*Figure 20.* Posterior distributions of effects of block on three parameters: response cutoff (top), log of internal noise (middle), and the probit of lapse rate (bottom).

## Discussion

The present study assessed changes in parameters representing response bias, sensitivity, and lapse rate across the first and last 4-minutes of a cognitive vigilance task.  Data gave substantial evidence against an effect of response cutoff and strong evidence against an effect of sensitivity, suggesting that neither bias shifts nor sensitivity losses contributed to changes in response rates in this task. Data did give substantial evidence for an effect of lapse rate, whereby lapses increased from the first to the last block, but the decrease in response rates was minimal.

Like previous comparisons of sensory and cognitive vigilance, the studies reported within this dissertation reveal marked differences in performance between tasks requiring sensory discriminations versus those requiring cognitive discriminations. Chapters 2-4 report data exclusively from sensory tasks, for which analyses revealed substantial vigilance decrements. In contrast, detections in the present study remained relatively stable over time, despite the cognitive task meeting conditions known to encourage sensitivity losses in sensory tasks (i.e., high event rates and successive discriminations). This pattern accords with previous findings that vigilance losses are generally smaller in cognitive tasks than in sensory tasks (Ash et al., 1983; Loeb et al., 1982; Noonan et al., 1984; Noonan et al., 1985; See et al., 1995).

The present results are not, however, consistent with the vigilance increment—an increase in detections over time—that has been observed in some cognitive tasks (Deaton & Parasuraman, 1993; Dember et al., 1984; Koelega et al., 1989; Loeb et al., 1987; Lysaght et al., 1984; Noonan et al., 1985; Warm et al., 1984). While See and colleagues (1995) suggest that the vigilance increment might be driven by an increase in sensitivity, the present study found stronger evidence *against* a change in sensitivity than in any of the sensory tasks reported in earlier chapters. It remains a possibility, though, that sensitivity increments occur only under alternative task conditions, such as in low event rate or simultaneous discrimination tasks that are less demanding (Davies & Parasuraman, 1982).

We proposed that performance differences between sensory and cognitive tasks might be driven by differences in whether the tasks required continuous- or discrete-state judgments. Previous cognitive tasks have asked participants to make judgments that are discrete in nature, such as judging whether number is odd or even. In an odd/even discrimination, the number unambiguously belongs to one category or the other, limiting the extent to which responses can be affected by graded internal noise and shifts in response bias. The present task was designed to allow these effects, yet still did not produce changes in sensitivity or bias.

One possibility is that participants avoided making continuous state judgments by adopting a simpler heuristic for responding. Rather than judging whether the average of the numbers was drawn from a distribution with a mean of 490 or 510, as intended, participants could have simplified the task by responding “yes” whenever the average was greater than a criterion value (e.g., 500). Alternatively, performance differences between sensory and cognitive tasks might not be driven by differences in discrete- versus continuous-state decision models. Although previous work (Deaton & Parasuraman, 1993) ruled out the possibility that people perform better in cognitive tasks because alphanumeric stimuli make it easier for them to recognize when they have made an error, it remains possible that people are better able to notice errors when making cognitive discriminations, irrespective of stimulus type. Future work should directly test the effects of knowledge of results for tasks requiring sensory and cognitive discriminations.

# Chapter 7: General discussion

The systematic study of vigilance began during World War II as a means of understanding why radar and sonar operators missed targets at the end of their watch that they were usually able to detect. This decline in the ability to detect infrequent signals over time, termed the vigilance decrement, is now one of the best-established findings in the human performance literature (Hancock, 2013). Early theories of vigilance assumed that the decrement was caused by a decline in observers' perceptual ability to distinguish signals from noise (i.e., sensitivity; Frankman & Adams, 1962; J. F. Macworth, 1968b). However, the development of signal detection theory (Green & Swets, 1966), allowed researchers to distinguish changes in observers' sensitivity from changes in their response bias by analyzing the pattern of tradeoffs between hits and false alarms.

Following the adoption of SDT by vigilance researchers in the 1960's, many studies reported that the vigilance decrement primarily resulted from monitors adopting progressively more conservative response criteria over time, rather than declines in sensitivity (Broadbent, 1971; Mackworth, 1970; Swets & Kristofferson, 1970). But SDT analyses did sometimes reveal concurrent declines in sensitivity (Loeb & Binford, 1968; Smith & Barany, 1970). To identify the types of tasks that produced sensitivity losses, Parasuraman and Davies (1977) evaluated vigilance studies that reported SDT analyses, concluding that sensitivity only declined in tasks that combined high event rates (>24/min) with successive discriminations.

Although many studies found sensitivity losses in successive, high event rate tasks (e.g., Claypoole et al., 2019; Galinksy et al., 1993; Smith et al., 2002; Warm &

Jerison, 1984), others observed sensitivity losses in tasks with low event rates (e.g., Mackworth, 1970; Tomporowski & Simpson, 1990; Williges, 1971) and simultaneous discriminations (Dittmar et al., 1985.; Parasuraman & Mouloua, 1987), suggesting that the original taxonomy was incomplete. Proposed revisions to the taxonomy suggest that sensitivity losses depend on overall task load (Parasuraman & Mouloua, 1987), rather than the specific combination of high event rates and successive discriminations, as well as whether tasks require sensory or cognitive discriminations (See et al., 1995).

Theorists generally attribute the selective pattern of sensitivity losses to limitations in effortful attention. Specifically, the *resource depletion account* (Caggiano & Parasuraman, 2004; Grier et al., 2003; Parasuraman, 1979) proposes that tasks imposing a high load on observers gradually exhaust the limited supply of attentional resources. Meanwhile, the *resource control hypothesis* (Thomson et al., 2015) proposes that executive control wanes over time, allowing resources to drift to task-unrelated thoughts. In either case, fewer resources are available to the vigilance task, reducing information-processing quality. Both models are consistent with findings that vigilance tasks are stressful and subjectively effortful (Dember et al., 1996; Warm et al., 2008).

Recent work, however, suggests that traditional SDT analyses of binary detection data are inadequate for understanding the vigilance decrement. One concern is that apparent losses of sensitivity might be a statistical artifact of very low false alarm rates, which are common in vigilance tasks (Thomson et al., 2016). Unless the false alarm rate is high enough to allow a statistically detectable decrease between task conditions (which is often not the case in vigilance tasks; Thomson et al., 2016) changes in sensitivity and bias are indistinguishable. To circumvent this issue, researchers often prefer to use *A*' in place of *d'*, as it is widely believed to be a nonparametric measure of sensitivity. However, *A'* is not actually bias-free (Getty et al., 1995; Macmillan & Creelman, 1996; Pastore et al., 2003) and shifts in bias have been shown to produce spurious changes in *A'* (McCarley et al., 2021). Together, these issues prompt reconsideration of the mechanisms underlying vigilance decrements.

An alternative theory, *mindlessness*, proposes that the vigilance decrement is driven by attentional lapses: brief periods in which attention is fully decoupled from the task (Jerison et al., 1965; Manly et al., 1999; Robertson et al., 1997). But, with only two degrees of freedom, binary detection data are unable to account for changes in bias shifts, sensitivity losses *and* attentional lapses.

To address these limitations, the goal of the work presented within this dissertation was to re-examine the mechanisms underlying the vigilance decrement using methods suitable for distinguishing between bias shifts, sensitivity losses and attentional lapses. To this end, Chapter 3 presented two studies that examined changes in the psychometric curve between the first and last 4-minutes of an online sensory vigilance task. Psychometric curves revealed that the decrement was largely driven by an increase in the frequency of attentional lapses over time, followed by observers adopting a progressively stricter cutoff for making signal present responses. Across both studies, data failed to give convincing evidence for or against an effect of block on sensitivity. Although McCarley and Yamani (2021) found decisive evidence of a sensitivity loss in their lab-based version of the task, its absence in the online studies suggests that 1) the vigilance decrement may reflect different mechanisms when tasks are conducted online versus in the lab, and 2) sensitivity losses are less robust than bias shifts and attentional lapses.

Chapter 4 introduced an alternative modeling approach for estimating changes in response bias, sensitivity, and attentional lapses over the course of a vigilance task. The generative cognitive model was written to directly simulate observer's response cutoffs, internal noise (corresponding to sensitivity), and attentional lapse rates, linking changes in response rates directly to the underlying cognitive mechanisms. The chapter presents two demonstrations that the generative model captures trends in the data previously revealed by analyzing changes in the psychometric curve. Results provide converging evidence that the vigilance decrement primarily reflects increases in attentional lapses and conservative bias shifts, in both an online and a lab-based sensory task.

Chapter 5 shifted from examining mechanisms of vigilance loss in sensory tasks to examining vigilance within a cognitive task. Given the breadth of research examining sensory vigilance, relatively few studies have employed cognitive tasks. Those that have, tend to find that vigilance decrements and sensitivity losses are smaller than they are in sensory tasks. We proposed that these effects might have been driven by cognitive vigilance tasks mapping stimuli to discrete rather than continuous states, thereby limiting the extent to which responses are affected by internal noise and shifts in response criterion. However, despite designing a cognitive task that required continuous-state judgments, response rates remained relatively stable across blocks. Data gave evidence against a bias shift and a sensitivity loss. The negligible decline in response rates appears to have been driven entirely by an increase in attentional lapses.

Overall, these results support the distinction between sensory and cognitive vigilance tasks made by other researchers (Deaton & Parasuraman, 1993; Koelega et al., 1989; See et al., 1995) and suggest that performance differences between sensory and cognitive tasks might not be driven entirely by differences in discrete versus continuous-state decision processes.

Collectively, the large-N, preregistered studies presented in this dissertation provide convincing evidence that the vigilance decrement largely reflects attentional lapses and conservative shifts in bias over time. Although analyses do not rule out sensitivity losses, the results reported here are not consistent with theories that assign resource depletion a significant role in the vigilance decrement. Finally, the results suggest that interventions aimed at reducing attentional lapses and bias shifts might be more effective at mitigating the vigilance decrement than those focused on minimizing sensitivity losses.

**References**


Aaslid, R. (1986). Transcranial Doppler examination techniques. In *Transcranial doppler sonography* (pp. 39–59). Springer.

Alikonis, C. R., Warm, J. S., Matthews, G., Dember, W. N., Hitchcock, E. M., & Kellaris, J. J. (2002). Vigilance, Workload, and Boredom: Two Competing Models. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(17), 1531–1535. https://doi.org/10.1177/154193120204601701

Antonelli, D. C., & Karas, G. G. (1967). Performance on a Vigilance Task under Conditions of True and False Knowledge of Results. *Perceptual and Motor Skills*, *25*(1), 129–138. https://doi.org/10.2466/pms.1967.25.1.129

Antrobus, J. S. (1968). Information Theory and Stimulus-Independent Thought. *British Journal of Psychology*, *59*(4), 423–430. https://doi.org/10.1111/j.2044-8295.1968.tb01157.x

Ariga, A., & Lleras, A. (2011). Brief and rare mental “breaks” keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*, *118*(3), 439–443. https://doi.org/10.1016/j.cognition.2010.12.007

Baddeley, A., & Colquhoun, W. (1969). Signal probability and vigilance: A reappraisal of the ‘signal-rate’effect. *British Journal of Psychology*, *60*(2), 169–178.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Elsevier. https://doi.org/10.1016/S0079-7421(08)60452-1

Bakan, P. (1955). Discrimination decrement as a function of time in a prolonged vigil. *Journal of Experimental Psychology*, *50*(6), 387.

Baker, C. (1963). Consistency of performance in two visual vigilance tasks. *Vigilance: A Svrnoosium*, 43–52.

Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in Thought: The Effect of Mind Wandering on the Processing of Relevant and Irrelevant Events. *Psychological Science*, *22*(5), 596–601. https://doi.org/10.1177/0956797611404083

Beh, H. C. (1989). The effect of passive smoking on vigilance performance. *Ergonomics*, *32*(10), 1227–1236. https://doi.org/10.1080/00140138908966892

Bergum, B. O., & Lehr, D. J. (1964). Monetary incentives and vigilance. *Journal of Experimental Psychology*, *67*(2), 197–198. https://doi.org/10.1037/h0040966

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica*, *134*(3), 330–343. https://doi.org/10.1016/j.actpsy.2010.03.006

Broadbent, D., & Gregory, M. (1963). Vigilance considered as a statistical decision. *British Journal of Psychology*, *54*(4), 309–323.

Broadbent, Donald E. (1954). Some effects of noise on visual performance. *Quarterly Journal of Experimental Psychology*, *6*(1), 1–5.

Broadbent, Donald E., & Gregory, M. (1965). Effects of Noise and of Signal Rate upon Vigilance Analysed by Means of Decision Theory. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *7*(2), 155–162. https://doi.org/10.1177/001872086500700207

Broadbent, Donald Eric. (1971). *Decision and stress.*

Brusovansky, M., Glickman, M., & Usher, M. (2018). Fast and effective: Intuitive processes in complex decisions. *Psychonomic Bulletin & Review*, *25*(4), 1542-1548.

Buckner, D. N., Harabedian, A., & McGrath, J. J. (1960). *A study of individual differences in vigilance performance*. Human Factors Research Inc. Los Angeles, CA.

Buckner, D. N., & McGrath, J. J. (1963). *Vigilance: A symposium.*

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, *11*(5), 932–937. https://doi.org/10.3758/BF03196724

Cahoon, R. L. (1973). Auditory vigilance under hypoxia. *Journal of Applied Psychology*, *57*(3), 350.

Caldeira, J. D. (1980). Parametric Assumptions of Some “Nonparametric” Measures of Sensory Efficiency. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *22*(1), 119–120. https://doi.org/10.1177/001872088002200113

Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, *35*, 33-41.

Catalano, J. F. (1973). Effect of Perceived Proximity to End of Task upon End-Spurt. *Perceptual and Motor Skills*, *36*(2), 363–372. https://doi.org/10.2466/pms.1973.36.2.363

Chadda, K. G. (1992). *The effects of knowledge of results and aging on vigilance.*

Chan, L. K., & Chan, W. W. (2022). Target-rate effect in continuous visual search. *Cognitive research: principles and implications*, *7*(1), 1-22.

Childs, J. M., & Halcomb, C. G. (1972). Effects of Noise and Response Complexity upon Vigilance Performance. *Perceptual and Motor Skills*, *35*(3), 735–741. https://doi.org/10.2466/pms.1972.35.3.735

Claypoole, V. L., Dever, D. A., Denues, K. L., & Szalma, J. L. (2019). The Effects of Event Rate on a Cognitive Vigilance Task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(3), 440–450. https://doi.org/10.1177/0018720818790840

Colquhoun, W. P., & Baddeley, A. D. (1967). Influence of signal probability during pretraining on vigilance decrement. *Journal of Experimental Psychology*, *73*(1), 153–155. https://doi.org/10.1037/h0024087

Colquhoun, W. Peter, & Edwards, R. S. (1970). Practice Effects on a Visual Vigilance Task with and without Search. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *12*(6), 537–545. https://doi.org/10.1177/001872087001200603

Craig, A. (1977). Broadbent and Gregory Revisited: Vigilance and Statistical Decision. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *19*(1), 25–36. https://doi.org/10.1177/001872087701900103

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

Davies, D. R., & Tune, G. S. (1969). *Human vigilance performance.*

De Joux, N., Russell, P. N., & Helton, W. S. (2013). A functional near-infrared spectroscopy study of sustained attention to local and global target features. *Brain and Cognition*, *81*(3), 370–375. https://doi.org/10.1016/j.bandc.2012.12.003

De Waard, D. (2002). *Mental workload*.

Deaton, J. E., & Parasuraman, R. (1988). Effects of Task Demands and Age on Vigilance and Subjective Workload. *Proceedings of the Human Factors Society Annual Meeting*, *32*(19), 1458–1462. https://doi.org/10.1177/154193128803201932

Deaton, J. E., & Parasuraman, R. (1993). Sensory and Cognitive Vigilance: Effects of Age on Performance and Subjective Workload. *Human Performance*, *6*(1), 71–97. https://doi.org/10.1207/s15327043hup0601_4

Deaton, M., Wilkinson, R. T., & Tobias, J. S. (1971). The effect of sleep deprivation on signal detection parameters. *Quarterly Journal of Experimental Psychology*, *23*(4), 449–452. https://doi.org/10.1080/14640747108400257

Dember, W. N., Joshi, A., Warm, J. S., & Scerbo, M. W. (1985). Effects of event rate and spatial uncertainty on performance in simultaneous and successive vigilance tasks. *Meeting of the Psychonomic Society, Boston, MA.*

Ditchburn, R. (1943). Some factors affecting efficiency of work of lookouts. *Admiralty Res. Lab. Rep., Great Britain*.

Dittmar, M. L., Warm, J. S., & Dember, W. (1985). Effects of knowledge of results on performance in successive and simultaneous vigilance tasks: A signal detection analysis. *Trends in Ergonomics/Human Factors II*, 195–202.

Dittmar, Mary Lynne, Warm, J. S., Dember, W. N., & Ricks, D. F. (1993). Sex differences in vigilance performance and perceived workload. *The Journal of General Psychology*, *120*(3), 309–322.

Dockree, P. M., Kelly, S. P., Roche, R. A., Hogan, M. J., Reilly, R. B., & Robertson, I. H. (2004). Behavioural and physiological impairments of sustained attention after traumatic brain injury. *Cognitive Brain Research*, *20*(3), 403-414.

Duschek, S., & Schandry, R. (2003). Functional transcranial Doppler sonography as a tool in psychophysiological research. *Psychophysiology*, *40*(3), 436–454.

Eason, R. G., Beardshall, A., & Jaffee, S. (1965). Performance and physiological indicants of activation in a vigilance situation. *Perceptual and Motor Skills*, *20*(1), 3–13.

Egan, J. P., Greenberg, G. Z., & Schulman, A. I. (1961). Operating characteristics, signal detectability, and the method of free response. *The Journal of the Acoustical Society of America*, *33*(8), 993–1007.

Eilers, K., Hänecke, K., Peper, J., Nachreiner, F., & Leonard, J. P. (1988). Time of day effects in vigilance performance. *Vigilance Methods, Models and Regulation. Frankfurt: Peter Lang*, 181–190.

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W. (2018). Working Memory and Executive Attention: A Revisit. *Perspectives on Psychological Science*, *13*(2), 190–193. https://doi.org/10.1177/1745691617720478

Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of Monetary Reward Can Attenuate the Vigilance Decrement. *PLOS ONE*, *11*(7), e0159741. https://doi.org/10.1371/journal.pone.0159741

Esterman, M., Reagan, A., Liu, G., Turner, C., & DeGutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *Journal of Experimental Psychology: General*, *143*(6), 2287–2295. https://doi.org/10.1037/xge0000019

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, *20*(3), 586–592. https://doi.org/10.3758/s13423-012-0367-y

Finomore Jr., V. S. (2006). *Effects of feature presence/absence and event asynchrony on vigilance performance and perceived mental workload* [PhD Thesis]. University of Cincinnati.

Finomore, V., Matthews, G., Shaw, T., & Warm, J. (2009). Predicting vigilance: A fresh look at an old problem. *Ergonomics*, *52*(7), 791–808. https://doi.org/10.1080/00140130802641627

Fisk, A. D., & Schneider, W. (1981). Control and Automatic Processing During Tasks Requiring Sustained Attention: A New Approach to Vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *23*(6), 737–750. https://doi.org/10.1177/001872088102300610

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, *111*(3), 345–355. https://doi.org/10.1016/j.cognition.2009.02.006

Frankman, J. P., & Adams, J. A. (1962). *Theories of vigilance. 59*, 243–254.

Funke, M. E., Warm, J. S., Matthewi, G., & Finomore Jr, V. (2010). Static and Dynamic in Vigilance: Effects on Cerebral Hemodynamics and Workload. *Advances in Understanding Human Performance: Neuroergonomics, Human Factors Design, and Special Populations*, 80.

Galinsky, T. L., Rosa, R. R., Warm, J. S., & Dember, W. N. (1993). Psychophysical Determinants of Stress in Sustained Attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(4), 603–614. https://doi.org/10.1177/001872089303500402

Giambra, L. M. (1995). A Laboratory Method for Investigating Influences on Switching Attention to Task-Unrelated Imagery and Thought. *Consciousness and Cognition*, *4*(1), 1–21. https://doi.org/10.1006/ccog.1995.1001

Gluckman, J. P., Dember, W. N., & Warm, J. S. (1988). Capacity Demand in Dual-Task Monitoring of Simultaneous and Successive Vigilance Tasks. *Proceedings of the Human Factors Society Annual Meeting*, *32*(19), 1463–1465. https://doi.org/10.1177/154193128803201933

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *75*(6), 424.

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The Vigilance Decrement Reflects Limitations in Effortful Attention, Not Mindlessness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 349–359. https://doi.org/10.1518/hfes.45.3.349.27253

Grodsky, A., & Giambra, L. M. (1990). The Consistency across Vigilance and Reading Tasks of Individual Differences in the Occurrence of Task-Unrelated and Task-Related Images and Thoughts. *Imagination, Cognition and Personality*, *10*(1), 39–52. https://doi.org/10.2190/6QG5-CXVV-4XUR-7P3K

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 Annual Meeting*, *39*(21), 1360–1364.

Guralnick, M. J. (1972). Observing responses and decision processes in vigilance. *Journal of Experimental Psychology*, *93*(2), 239. https://doi.org/10.1037/h0032458

Haines, N., Kvam, P. D., Irving, L. H., Smith, C., Beauchaine, T. P., Pitt, M. A., ... & Turner, B. (2020). Learning from the reliability paradox: How theoretically informed generative models can advance the social, behavioral, and brain sciences. *PsyArXiv, psyarxiv. com/xr7y3*.

Hancock, P.A. (1984) “Environmental Stressors” in Sustained Attention in Human Performance, Warm, J.S. (ed.) (Chichester, UK: Wiley), 103−142.

Hancock, P. A., & Weaver, J. L. (2005). On time distortion under stress. *Theoretical Issues in Ergonomics Science*, *6*(2), 193–211. https://doi.org/10.1080/14639220512331325747

Hancock, P. A. (1989). A Dynamic Model of Stress and Sustained Attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *31*(5), 519–537. https://doi.org/10.1177/001872088903100503

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology* (Vol. 52, pp. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Hatfield, J. L., & Loeb, M. (1968). Sense mode and coupling in a vigilance task. *Perception & Psychophysics*, *4*(1), 29–36. https://doi.org/10.3758/BF03210443

Hatfield, J. L., & Soderquist, D. R. (1970). Coupling Effects and Performance in Viligance Tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *12*(4), 351–359. https://doi.org/10.1177/001872087001200401

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values ofd′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46-51.

Head, J., & Helton, W. S. (2014). Sustained attention failures are primarily due to sustained cognitive load not task monotony. *Acta Psychologica*, *153*, 87–94. https://doi.org/10.1016/j.actpsy.2014.09.007

Healy, A. F., & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(5), 344.

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., Beauchamp, G., Parasuraman, R., & Hancock, P. A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology*, *96*(2), 249–261.

Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K., Matthews, G., Dember, W. N., Parasuraman, R., & Hancock, P. A. (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of Clinical and Experimental Neuropsychology*, *29*(5), 545–552.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437.

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*(3), 429–436. https://doi.org/10.1007/s00221-012-3322-2

Helton, W. S., Shaw, T. H., Warm, J. S., Matthews, G., Dember, W. N., & Hancock, P. A. (2004). Demand transitions in vigilance: Effects on performance efficiency and stress. *Human Performance, Situation Awareness and Automation: Current Research and Trends, HPSAAII*, *1*, 258–262.

Helton, W. S., Shaw, T., Warm, J. S., Matthews, G., & Hancock, P. (2008). Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress & Coping*, *21*(2), 173–184. https://doi.org/10.1080/10615800801911305

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*(1), 18–25. https://doi.org/10.1016/j.actpsy.2008.04.002

Helton, W. S., Warm, J. S., Matthews, G., Corcoran, K. J., & Dember, W. N. (2002). Further tests of an abbreviated vigilance task: Effects of signal salience and jet aircraft noise on performance and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(17), 1546–1550.

Hertzum, M., & Holmegaard, K. D. (2013). Perceived Time as a Measure of Mental Workload: Effects of Time Constraints and Task Success. *International Journal of Human-Computer Interaction*, *29*(1), 26–39. https://doi.org/10.1080/10447318.2012.676538

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*(3), 365–372.

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., Mayleben, D. W., & Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 89–112.

Hollander, T. D., Helton, W. S., Tripp, L. D., Parsons, K., Warm, J. S., Matthews, G., Dember, W. N., & Parasuraman, R. (2003). Cerebral vascularity and performance on an abbreviated vigilance task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(13), 1663–1667.

Hollander, T. D., Warm, J. S., Matthews, G., Shockley, K., Dember, W. N., Weiler, E., Tripp, L. D., & Scerbo, M. W. (2004). Feature Presence/Absence Modifies the Event Rate Effect and Cerebral Hemovelocity in Vigilance Performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(16), 1943–1947. https://doi.org/10.1177/154193120404801632

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Jenkins, H. M. (1958). The effect of signal-rate on performance in visual monitoring. *The American Journal of Psychology*, *71*(4), 647–661.

Jerison, H. J., & Pickett, R. M. (1964). Vigilance: The Importance of the Elicited Observing Rate. *Science*, *143*(3609), 970–971. https://doi.org/10.1126/science.143.3609.970

Jerison, H. J., & Wing, J. F. (1961). Human Vigilance and Operant Behavior. *Science*, *133*(3456), 880–881. https://doi.org/10.1126/science.133.3456.880

Jerison, H. J. (1967). Signal Detection Theory in the Analysis of Human Vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *9*(3), 285–288. https://doi.org/10.1177/001872086700900310

Jerison, H. J., Pickett, R. M., & Stenson, H. H. (1965). The Elicited Observing Rate and Decision Processes in Vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *7*(2), 107–128. https://doi.org/10.1177/001872086500700204

Johnston, W. A., Howell, W. C., & Williges, R. C. (1969). The components of complex monitoring. *Organizational Behavior and Human Performance*, *4*(2), 112–124. https://doi.org/10.1016/0030-5073(69)90002-6

Joshi, A. (1985). *Effects of event rate and spatial uncertainty on performance in simultaneous and successive vigilance tasks* [PhD Thesis]. University of Cincinnati.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Kantowitz, B., & Knight, J. (1974). Testing tapping time-sharing. *Journal of Experimental Psychology*, *103*(2), 331–336.

Kingdom, F. A. A., Prins, N. (2016). Psychophysics: A practical introduction (2nd ed.). Academic Press.

Klinger, E. (1978). Modes of normal conscious flow. In *The stream of consciousness* (pp. 225–258). Springer.

Koelega, H. S., Brinkman, J.-A., Hendriks, L., & Verbaten, M. N. (1989). Processing demands, effort, and individual differences in four different vigilance tasks. *Human Factors*, *31*(1), 45–62.

Lanzetta, T. M., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of Task Type and Stimulus Heterogeneity on the Event Rate Function in Sustained Attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(6), 625–633. https://doi.org/10.1177/001872088702900602

Lanzetta, T. M., Warm, J. S., Dember, W. N., & Berch, D. B. (1985). Information Processing Load and the Event Rate Function in Sustained Attention.

*Proceedings of the Human Factors Society Annual Meeting*, *29*(12), 1084–1088. https://doi.org/10.1177/154193128502901201

Levine, J. M. (1966). The Effects of Values and Costs on the Detection and Identification of Signals in Auditory Vigilance. *Human Factors*, *8*(6), 525–537. https://doi.org/10.1177/001872086600800607

Levine, J. M., Romashko, T., & Fleishman, E. A. (1973). Evaluation of an abilities classification system for integrating and generalizing human performance research findings: An application to vigilance tasks. *Journal of Applied Psychology*, *58*(2), 149.

Liao, J. G., Midya, V., & Berg, A. (2021). Connecting and contrasting the Bayes factor and a modified ROPE procedure for testing interval null hypotheses. *The American Statistician*, *75*(3), 256-264

Liu, C. C., & Aitkin, M. (2008). Bayes factors: Prior sensitivity and model generalizability. *Journal of Mathematical Psychology*, *52*(6), 362-375.

Lindsley, D., & Anderson, I. (1944). Radar operator fatigue: The effects of length and repetition of operating periods on efficiency of performance. *Office of Scientific Research Research and Development*.

Loeb, M., & Schmidt, E. A. (1963). A comparison of the effects of different kinds of information in maintaining efficiency on an auditory monitoring task.. *Ergonomics*, *6*(1), 75–81. https://doi.org/10.1080/00140136308930677

Loeb, Michel, & Binford, J. R. (1964). Vigilance for Auditory Intensity Changes as a Function of Preliminary Feedback and Confidence Level. *Human Factors*, *6*(5), 445–458. https://doi.org/10.1177/001872086400600502

Loeb, Michel, & Binford, J. R. (1968). Variation in performance on auditory and visual monitoring tasks as a function of signal and stimulus frequencies. *Perception & Psychophysics*, *4*(6), 361–367.

Loeb, Michel, Noonan, T. K., Ash, D. W., & Holding, D. H. (1987). Limitations of the cognitive vigilance increment. *Human Factors*, *29*(6), 661–674.

Luximon, A., & Goonetilleke, R. S. (2001). Simplified subjective workload assessment technique. *Ergonomics*, *44*(3), 229–243. https://doi.org/10.1080/00140130010000901

Mackworth, J. F. (1968a). The effect of signal rate on performance in two kinds of vigilance task. *Human Factors*, *10*(1), 11–17.

Mackworth, J. F. (1968b). Vigilance, arousal, and habituation. *Psychological Review*, *75*(4), 308.

Mackworth, J. F. (1970). *Vigilance and attention: A signal detection approach.*

Mackworth, N. (1948). The breakdown of vigilance during prolonged visual search: Quarterly Journal of Experimental Psychology: Vol 1, No 1. *Quarterly Journal of Experimental Psychology*, *1*(1), 6–21.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Researches on the Measurement of Human Performance.*, *268*.

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*(5), 1042–1058. https://doi.org/10.3758/APP.71.5.1042

Manly, T. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661–670. https://doi.org/10.1016/S0028-3932(98)00127-4

Matthews, G. (1996). Signal probability effects on high-workload vigilance tasks. *Psychonomic Bulletin & Review*, *3*(3), 339–343. https://doi.org/10.3758/BF03210757

Matthews, G., Campbell, S. E., & Falconer, S. (2001). Assessment of Motivational States in Performance Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(13), 906–910. https://doi.org/10.1177/154193120104501302

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., ... & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. Emotion, 2(4), 315.

Matthews, G., & Davies, D. R. (2001). Individual differences in energetic arousal and sustained attention: A dual-task study. *Personality and Individual Differences*, *31*(4), 575–589. https://doi.org/10.1016/S0191-8869(00)00162-8

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

Matthews, G., Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied*, *16*(2), 187.

Matthews, G., Warm, J. S., & Smith, A. P. (2017). Task Engagement and Attentional Resources: Multivariate Models for Individual Differences and Stress Factors in Vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(1), 44–61. https://doi.org/10.1177/0018720816673782

McCann, P. H. (1969). The Effects of Ambient Noise on Vigilance Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *11*(3), 251–256. https://doi.org/10.1177/001872086901100307

McCarley, J. S. & Yamani, Y. (2021). Psychometric curves reveal three mechanisms of vigilance decrement. *Psychological Science.*

McCarley, J. S., Yamani, Y., & Gyles, S.P. (2021, September). Comparing measures of sensitivity for vigilance tasks: Bayesian-estimated and log-linear corrected measures outperform A!. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 65, No. 1, pp. 1459-1463). Sage CA: Los Angeles, CA: SAGE Publications.

McCormack, P. (1959). Performance in a vigilance task with and without knowledge of results. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *13*(2), 68.

McGrath, J. J. (1961). *Cross-validation of some correlates of vigilance performance*. HUMAN FACTORS RESEARCH INC LOS ANGELES CA.

McGrath, J. J., Harabedian, A., & Buckner, D. N. (1960). *An exploratory study of the correlates of vigilance performance*. HUMAN FACTORS RESEARCH INC LOS ANGELES CA.

McIntire, L., McKinley, R. A., McIntire, J., Goodyear, C., & Nelson, J. (2013). Eye metrics: An alternative vigilance detector for military operators. *Military Psychology*, *25*(5), 502-513.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196–204. https://doi.org/10.1037/a0014104

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "d'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 525–549. https://doi.org/10.1037/a0025896

Meuter, R. F. I., & Lacherez, P. F. (2016). When and Why Threats Go Undetected: Impacts of Event Rate and Shift Length on Threat Detection Accuracy During Airport Baggage Screening. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(2), 218–228. https://doi.org/10.1177/0018720815616306

Milošević, S. (1974). Effect of time and space uncertainty on a vigilance task. *Perception & Psychophysics*, *15*(2), 331–334. https://doi.org/10.3758/BF03213953

Molloy, R., & Parasuraman, R. (1996). Monitoring an Automated System for a Single Failure: Vigilance and Task Complexity Effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*(2), 311–322. https://doi.org/10.1177/001872089606380211

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, *27*, 84–92.

Mouloua, M., & Parasuraman, R. (1995). Aging and Cognitive Vigilance: Effects of Spatial Uncertainty and Event Rate. *Experimental Aging Research*, *21*(1), 17–32. https://doi.org/10.1080/03610739508254265

Nachreiner, F., & Hänecke, K. (1992). Vigilance. In *Handbook of human performance, Vol. 1: The physical environment; Vol. 2: Health and performance; Vol. 3: State and trait* (pp. 261–288). Academic Press.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*(3), 214. https://doi.org/10.1037/0033-295X.86.3.214

Neigel, A. R., Claypoole, V. L., Fraulini, N. W., Waldfogle, G. E., & Szalma, J. L. (2019). Where is my mind? Examining mind-wandering and vigilance performance. *Experimental Brain Research*, *237*(2), 557–571. https://doi.org/10.1007/s00221-018-5438-5

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science*, *220*(4594), 327–329. https://doi.org/10.1126/science.6836276

Parasuraman, R. (1976). Consistency of individual differences in human vigilance performance: An abilities classification analysis. *Journal of Applied Psychology*, *61*(4), 486.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*(4409), 924–927. https://doi.org/10.1126/science.472714

Parasuraman, R. (1985). Sustained attention: A multifactorial approach. *Attention and Performance XI*, *1482*, 493–511.

Parasuraman, R., & Davies, D. R. (1977). A Taxonomic Analysis of Vigilance Performance. In R. R. Mackie (Ed.), *Vigilance: Theory, Operational Performance, and Physiological Correlates* (pp. 559–574). Springer US. https://doi.org/10.1007/978-1-4684-2529-1_26

Parasuraman, R., Davies, D. R., & Toh, K. Y. (1984). *Time of day, memory load, and vigilance performance*. 9–14.

Parasuraman, R., & Giambra, L. (1991). Skill development in vigilance: Effects of event rate and age. *Psychology and Aging*, *6*(2), 155. https://doi.org/10.1037/0882-7974.6.2.155

Parasuraman, R., & Mouloua, M. (1987). Interaction of signal discriminability and task type in vigilance decrement. *Perception & Psychophysics*, *41*(1), 17–22. https://doi.org/10.3758/BF03208208

Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). “Nonparametric”A’ and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, *10*(3), 556–569. https://doi.org/10.3758/BF03196517

Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological methods*, *23*(2), 208.

Plummer, M. (2015). JAGS Version 4.0. 0 user manual.

Pop, V. L., Stearman, E. J., & Durso, F. T. (2010). Vigilance decrements in the NextGen environment. Proceedings of HCI–Aero 2010 Conference.

Pop, V. L., Stearman, E. J., Kazi, S., & Durso, F. T. (2012). Using Engagement to Negate Vigilance Decrements in the NextGen Environment. *International Journal of Human-Computer Interaction*, *28*(2), 99–106. https://doi.org/10.1080/10447318.2012.634759

Punwani, S., Ordidge, R. J., Cooper, C. E., Amess, P., & Clemence, M. (1998). MRI measurements of cerebral deoxyhaemoglobin concentration [dHb]—Correlation with near infrared spectroscopy (NIRS). *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, *11*(6), 281–289.

R Core Team (2019). R Foundation for Statistical Computing. R: A language and environment for statistical computing.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday Attention: Variation in Mind Wandering and Memory in a Lecture: Mind wandering. *Applied Cognitive Psychology*, *26*(2), 234–242. https://doi.org/10.1002/acp.1814

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*(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8

Rose, C. L., Murphy, L. B., Schickedantz, B., & Tucci, J. (2001). The Effects of Event Rate and Signal Probability on Childrens Vigilance. *Journal of Clinical and Experimental Neuropsychology*, *23*(2), 215–224. https://doi.org/10.1076/jcen.23.2.215.1205

Sawyer, B. D., Finomore, V. S., Funke, G. J., Mancuso, V. F., Funke, M. E., Matthews, G., & Warm, J. S. (2014). Cyber vigilance: Effects of signal probability and event rate. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 1771–1775.

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. *The Journal of General Psychology*, *120*(3), 293–307.

Scerbo, M. W., Warm, J. S., Doettling, V. S., Parasuraman, R., & Fisk, A. D. (1987). Event asynchrony and task demands in sustained attention. In *Ergonomics and Human Factors* (pp. 33–39). Springer.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schnittger, C., Johannes, S., Arnavaz, A., & Münte, T. F. (1997). Relation of cerebral blood flow velocity and level of vigilance in humans. *NeuroReport*, *8*(7), 1637–1639.

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 Annual Meeting*, *41*(2), 1288–1292.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, *117*(2), 230–249. https://doi.org/10.1037/0033-2909.117.2.230

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-Wandering With and Without Intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. https://doi.org/10.1016/j.tics.2016.05.010

Shaw, T. H., Matthews, G., Warm, J. S., Finomore, V. S., Silverman, L., & Costa Jr, P. T. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research in Personality*, *44*(3), 297–308.

Shaw, T. H., Parasuraman, R., Sikdar, S., & Warm, J. (2009). Knowledge of Results and Signal Salience Modify Vigilance Performance and Cerebral Hemovelocity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(17), 1062–1065. https://doi.org/10.1177/154193120905301704

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*(4), 657–690. https://doi.org/10.1016/j.concog.2004.06.003

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, *14*(3), 527–533. https://doi.org/10.3758/BF03194102

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Smit, A. S., Eling, P. A. T. M., & Coenen, A. M. L. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*(1), 35–42. https://doi.org/10.1016/j.actpsy.2003.11.001

Smith, C., Mikulka, P., Freeman, F., & Scerbo, M. (2002). The Effects of Event Rate and Array Size on a Cognitive Vigilance Task with Associated EEG Rhythms and a Derived Engagement Index. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(17), 1674–1678. https://doi.org/10.1177/154193120204601731

Smith, L. A., & Barany, J. W. (1970). An Elementary Model of Human Performance on Paced Visual Inspection Tasks. *A I I E Transactions*, *2*(4), 298–308. https://doi.org/10.1080/05695557008974768

Stroobant, N., & Vingerhoets, G. (2000). Transcranial Doppler ultrasonography monitoring of cerebral hemodynamics during performance of cognitive tasks: A review. *Neuropsychology Review*, *10*(4), 213–231.

Swets, J. A. (1973). The Relative Operating Characteristic in Psychology: A technique for isolating effects of response bias finds wide use in the study of perception and cognition. *Science*, *182*(4116), 990–1000. https://doi.org/10.1126/science.182.4116.990

Swets, J A, & Kristofferson, A. B. (1970). Attention. *Annual Review of Psychology*, *21*(1), 339–366. https://doi.org/10.1146/annurev.ps.21.020170.002011

Swets, John A. (1977). Signal Detection Theory Applied to Vigilance. In R. R. Mackie (Ed.), *Vigilance* (pp. 705–718). Springer US. https://doi.org/10.1007/978-1-4684-2529-1_34

Swets, John A, Tanner Jr, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*(5), 301.

Szalma, J. L., Hancock, P. A., Dember, W. N., & Warm, J. S. (2006). Training for vigilance: The effect of knowledge of results format and dispositional optimism and pessimism on performance and stress. *British Journal of Psychology*, *97*(1), 115–135. https://doi.org/10.1348/000712605X62768

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors*, *46*(2), 219–233.

Tanner Jr., W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*(6), 401–409. https://doi.org/10.1037/h0058700

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: The Journal of the Human Factors and Ergonomics Society*, *42*(2), 183–194. https://doi.org/10.1518/001872000779656480

Thomson, D. R., Besner, D., & Smilek, D. (2013). In pursuit of off-task thought: Mind wandering-performance trade-offs while reading aloud and color naming. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00360

Thomson, D. R., Besner, D., & Smilek, D. (2015). A Resource-Control Account of Sustained Attention: Evidence From Mind-Wandering and Vigilance Paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96. https://doi.org/10.1177/1745691614556681

Thomson, D. R., Besner, D., & Smilek, D. (2016). A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychological Review*, *123*(1), 70–83. https://doi.org/10.1037/rev0000021

Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, *27*, 14–26. https://doi.org/10.1016/j.concog.2014.04.001

Thurmond, 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*(2), 100–102. https://doi.org/10.3758/BF03210142

Tomporowski, P. D., & Simpson, R. G. (1990). Sustained attention and intelligence. *Intelligence*, *14*(1), 31–42. https://doi.org/10.1016/0160-2896(90)90012-I

Toronov, V., Webb, A., Choi, J. H., Wolf, M., Michalos, A., Gratton, E., & Hueber, D. (2001). Investigation of human brain hemodynamics by simultaneous near-infrared spectroscopy and functional magnetic resonance imaging. *Medical Physics*, *28*(4), 521–527.

Tripp, L. D., & Warm, J. S. (2007). Transcranial Doppler sonography. In R. Parasuraman & M. Rizzo (Eds.), Neuroergonomics: The brain at work (pp. 82-94). New York, NY: Oxford University Press.

Tyler, D. M., Waag, W. L., & Halcomb, C. G. (1972). Monitoring Performance Across Sense Modes: An Individual Differences Approach. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *14*(6), 539–547. https://doi.org/10.1177/001872087201400605

Verde, M. F., Macmillan, N. A., & Rotello, C. M. (2006). Measures of sensitivity based on a single hit rate and false alarm rate: The accuracy, precision, and robustness of’,A z, andA’. *Perception & Psychophysics*, *68*(4), 643–654. https://doi.org/10.3758/BF03208765

Vingerhoets, G., & Stroobant, N. (1999a). Between-task habituation in functional transcranial Doppler ultrasonography. *Neuroreport*, *10*(15), 3185–3189.

Vingerhoets, G., & Stroobant, N. (1999b). Lateralization of cerebral blood flow velocity changes during cognitive tasks: A simultaneous bilateral transcranial Doppler study. *Stroke*, *30*(10), 2152–2158.

Ware, J. R., Baker, R. A., & Sipowicz, R. R. (1962). Performance of mental deficients on a simple vigilance task. *American Journal of Mental Deficiency*, *66*, 647–650.

Ware, J. R., Sipowicz, R. R., & Baker, R. (1961). Auditory vigilance in repeated sessions. *Perceptual and Motor Skills*, *13*(2), 127–129.

Warm, J. S., Chin, K., Dittmar, M. L., & Dember, W. N. (1987). Effects of head restraint on signal detectability in simultaneous and successive vigilance tasks. *The Journal of General Psychology*, *114*(4), 423–431.

Warm, J. S., & Dember, W. N. (1998). Tests of vigilance taxonomy. In *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 87–112). American Psychological Association. https://doi.org/10.1037/10290-004

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In *Automation and human performance: Theory and applications* (pp. 183–200). Lawrence Erlbaum Associates, Inc.

Warm, J. S., Finomore, V., Shaw, T. H., Funke, M. E., Hausen, M. J., Matthews, G., Taylor, P., Vidulich, M. A., Repperger, D. W., Szalma, J. L., & Hancock, P. A. (2009). Effects of Training with Knowledge of Results on Diagnosticity in Vigilance Performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(17), 1066–1070. https://doi.org/10.1177/154193120905301705

Warm, J. S., Howe, S. R., Fishbein, H. D., Dember, W. N., & Sprague, R. L. (1984). Cognitive demand and the vigilance decrement. *Trends in Ergonomics/Human Factors I*, 15–20.

Warm, J. S., & Jerison, H. J., (1984). The psychophysics of vigilance. In: Warm, J.S. (Ed.), Warm, J. S., & Parasuraman, R. (2007). Cerebral hemodynamics and vigilance. *Neuroergonomics: The Brain at Work*, 146–158.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Warm, J. S., Tripp, L. D., Matthews, G., & Helton, W. S. (2012). Cerebral hemodynamic indices of operator fatigue in vigilance. *Handbook of Operator Fatigue*, 197–207.

Watkins, E. R. (2008). *Constructive and unconstructive repetitive thought. 134*(2), 163.

Weidenfeller, E. W., Baker, R. A., & Ware, J. R. (1962). Effects of Knowledge of Results (True and False) on Vigilance Performance. *Perceptual and Motor Skills*, *14*(2), 211–215. https://doi.org/10.2466/pms.1962.14.2.211

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickens, T. D. (2002). *Elementary Signal Detection Theory*. Oxford University Press.

Williams, P. S. (1986). Processing demands, training, and the vigilance decrement. *Human Factors*, *28*(5), 567–579.

Williges, R. C. (1969). Within-session criterion changes compared to an ideal observer criterion in a visual monitoring task. *Journal of Experimental Psychology*, *81*(1), 61–66. https://doi.org/10.1037/h0027432

Williges, R. C. (1971). The Role of Payoffs and Signal Ratios in Criterion Changes during a Monitoring Task. *Human Factors: The Journal of the Human*

*Factors and Ergonomics Society*, *13*(3), 261–267. https://doi.org/10.1177/001872087101300307

Williges, R. C. (1973). Manipulating the Response Criterion in Visual Monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *15*(2), 179–185. https://doi.org/10.1177/001872087301500209

Wyatt, S., & Langdon, J. N. (1932). Inspection Processes in Industry. (A Preliminary Report). *Inspection Processes in Industry. (A Preliminary Report).*, *63*. https://www.cabdirect.org/cabdirect/abstract/19322701072

Yadav, A. K., Singh, I. L., & Tiwari, T. (2015). Impact of cognitive demand on vigilance decrement: An overview. *Indian Journal of Positive Psychology*, *6*(2), 215.