Computational Modeling Reveals Minimal Vigilance Changes in a Cognitive Monitoring Task

Shannon Gyles[1], Yusuke Yamani[2], & Jason S. McCarley[1]

[1] School of Psychological Science, Oregon State University

[2] Department of Psychology, Old Dominion University

Authors' Accepted Manuscript

Author Note

This is the author's version of the manuscript accepted for publication in *Computational Brain & Behavior*. This version has not undergone final editing, copyediting, typesetting, or review of the resulting proof by the publisher. The final authenticated version is available at: https://doi.org/10.1007/s42113-024-00233-5.

Shannon Gyles https://orcid.org/0000-0001-5079-6033

Yusuke Yamani https://orcid.org/0000-0001-8990-0010

Jason S. McCarley https://orcid.org/0000-0002-8824-7491

This work was supported by a grant from the National Science Foundation (2240256) to Y.Y. and J.S.M. and a grant from the Human Factors and Ergonomics Society Perception and Performance Technical Group to S.G.

The authors have no relevant financial or non-financial interests to disclose.

The authors made the following contributions. Shannon Gyles: Conceptualization, Investigation, Software, Formal Analysis, Methodology, Writing - Original Draft Preparation, Writing - Review & Editing, Funding Acquisition; Yusuke Yamani: Conceptualization, Writing - Review & Editing, Funding Acquisition; Jason S. McCarley: Conceptualization, Investigation, Formal Analysis, Methodology, Writing - Original Draft Preparation, Writing - Review & Editing, Funding Acquisition.

Correspondence concerning this article should be addressed to Jason S. McCarley, 2950 SW Jefferson Way, Corvallis, OR, 97331. E-mail: jason.mccarley@oregonstate.edu

Abstract

The ability to monitor for rare critical events generally deteriorates over time on task, an effect termed the *vigilance decrement*. Although the decrement has been replicated many times, it has generally been studied with sensory discrimination tasks. Research using cognitive vigilance tasks, which require judgments of symbolic stimulus characteristics, has produced less consistent results. To test the robustness and nature of the cognitive vigilance decrement, the current study developed a computational performance model of a novel monitoring task. Participants performed a monitoring task that required them to estimate the central tendency of a set of three-digit readings each trial. For analysis, data from the first and last 4-min blocks of trials were fit with a model based on signal detection theory. The model assumed that participants could either perform the task in an attentive state in which decisions were stimulus-driven, or could lapse into an inattentive state in which decisions were guessed. Parameter estimats indicated an increase in mental lapse rate and decrease in positive guess rate over time, coupled with a decrease in internal processing noise. The effects of these latent changes on observable response rates, however, were modest and partially offsetting. Results suggest that mental lapses and a tendency to negative guesses are a common causes of vigilance loss across sensory and cognitive tasks, but may have small effects on observed responses.

Keywords: vigilance decrement, sustained attention, signal detection

Word count: 6206

Computational Modeling Reveals Minimal Vigilance Changes in a Cognitive Monitoring Task

Vigilance tasks like quality control and security surveillance require observers to monitor for infrequent signals over extended periods. A common finding is that detection falters over time on task (Mackworth, 1948), sometimes beginning within 5 minutes of task onset (Nuechterlein, Parasuraman, & Jiang, 1983). This pattern, the *vigilance decrement* (Proctor & Vu, 2023), has been replicated in hundreds of laboratory studies over multiple decades (See, Howe, Warm, & Dember, 1995; Warm, Finomore, Vidulich, & Funke, 2015), and has been observed in naturalistic tasks (Molloy & Parasuraman, 1996; Reinerman-Jones, Matthews, & Mercado, 2016).

Although the effect is familiar and well-studied, the psychological mechanisms underlying the vigilance decrement remain the subject of debate. Vigilance is commonly studied using yes-no detection tasks. In a task of this form, the observer is presented each trial with a stimulus from one of two categories, typically designated noise versus signal, and is asked to report whether a signal is present (yes) or absent (no). Data can be evaluated using signal detection theory (SDT: Green & Swets, 1966; Hautus, Macmillan, & Creelman, 2022). Under SDT, the observer encodes the evidence for the presence of a signal as a unidimensional decision variable, X. Variance in the decision variable is determined by the combination of external noise, variability that is inherent in the stimulus itself, and internal noise, variability that arises during the observer's sensory encoding and information processing. Confusability between signal and non-signal events exists when the distributions of X corresponding to the two categories overlap. Sensitivity, the observer's ability to distinguish signal from non-signal events, therefore increases as overlap between the distributions decreases. The observer transforms the decision variable to a yes or no judgment by comparing it to a response cutoff, rendering a positive judgment when X exceeds the cutoff value. Placement of the cutoff determines the observer's response bias; a low cutoff is *liberal*, favoring positive responses, and a high cutoff is *conservative*, favoring negative responses. A cutoff that favors neither positive or negative responses is *unbiased*.

Assuming that task demands and stimulus characteristics are held constant, detection rates in a vigilance task might drop over time for either of two reasons under SDT. One possibility is an increase in internal noise, reducing the observer's sensitivity. The second is a conservative shift in response cutoff. Data suggest that, in fact, both mechanisms can contribute to the vigilance decrement, but tend to do so unequally. Cutoff shifts are common (e.g., Broadbent & Gregory, 1963, 1965; A. Craig, 1987; Parasuraman, 1979; Swets, 1977), and presumably occur as the observer adapts to the low frequency of signal events (Colquhoun & Baddeley, 1967; A. Craig, 1978). Sensitivity losses are more selective, and in particular, are believed to obtain when the event rate of the task (i.e., number of trials per minute) is high and information processing demands are heavy (Nuechterlein et al., 1983; Parasuraman, 1979; Parasuraman & Mouloua, 1987; See et al., 1995). Theoretical accounts attribute sensitivity losses to gradual reductions in the attention allocated to the vigilance task. *Resource depletion theory* proposes that maintaining vigilance is mentally taxing (Grier et al., 2003; Warm, Parasuraman, & Matthews, 2008) and exhausts "reservoirs of energy" (Warm et al., 2015, p. 261) that determine the observers' information processing capacity (Caggiano & Parasuraman, 2004; Neigel et al., 2020; Schumann et al., 2022; Warm et al., 2015). *Resource control theory* argues that processing capacity remains constant, but that executive control failures or strategic choices let resources drift to task-unrelated thoughts (Thomson, Besner, & Smilek, 2015). Under either model, resources dedicated to the vigilance task dwindle over time, resulting in poorer sensitivity. However, generative modeling suggests that even when they occur, sensitivity decrements might produce relatively small declines in raw detection rates (Gyles, McCarley, & Yamani, 2023; McCarley & Yamani, 2021).

Mechanisms outside standard signal detection theory can also contribute to vigilance failures. At times, attention may lapse, disengaging entirely from the vigilance task (Esterman & Rothlein, 2019; Gyles et al., 2023; McCarley & Yamani, 2021). Lapses might result from any of multiple causes (Unsworth & Robison, 2016), including external distractions (Drody, Pereira, & Smilek, 2023; Robison & Unsworth, 2015; Unsworth & McMillan, 2014), microsleeps (Buckley, Helton, Innes, Dalrymple-Alford, & Jones, 2016), an intentional or unintentional (Seli, Risko, & Smilek, 2016; Thomson et al., 2015) drift of processing to off-task thoughts or stimuli (McVay & Kane, 2009, 2012), or a breakdown of goal representation (Ariga & Lleras, 2011; Manly, Robertson, Galloway, & Hawkins, 1999). Lapses become more common with longer time on task and are associated with performance losses (Cunningham, Scerbo, & Freeman, 2000; Krimsky, Forster, Llabre, & Jha, 2017; McVay & Kane, 2009, 2012; Unsworth & Robison, 2016; Zanesco, Denkova, & Jha, 2024), and thus contribute to the vigilance decrement (Esterman & Rothlein, 2019; Gyles et al., 2023; McCarley & Yamani, 2021) .

Lapses also introduce a further mechanism of declining response rates over the course of a vigil. Whatever its cause, the effect of an attention lapse within a signal detection task is that the participants' response is selected independent of the stimulus (Kingdom & Prins, 2016). Responses during lapses can therefore be modeled as guesses (Kuss, Jäkel, & Wichmann, 2005; Lee, 2018). A decrease in the probability of guessing a positive response is then an additional source of vigilance loss; as the participant adapts to the low signal rate, the positive guess rate declines (Gyles et al., 2023).

Unfortunately, attention lapses and guesses complicate efforts to distinguish sensitivity from bias with yes-no response data. The most common measure of sensitivity, *d'*, assumes that the values of X associated with signal and noise events are normal with equal variance. Guessed responses violate this parametric assumption. Although it is often described as non-parametric, an alternative measure of yes-no sensitivity, *A'*, also incorporates parametric assumptions that are violated by response guessing (Macmillan & Creelman, 1996; Pastore, Crawley, Berens, & Skelly, 2003). More generally, yes-no data provide only two degrees of freedom, hit and false alarm rate, too few to identify a model with sensitivity, response bias, lapse rate, and cutoff rate as parameters.

As an alternative method to isolate sensitivity losses, bias changes, lapses, and guesses, McCarley and Yamani (2021) proposed the analysis of psychometric functions for vigilance tasks. The psychometric function for a detection task presents the positive response rate as a function of stimulus intensity, and is typically an S-shaped curve. The form and position of the function reflect the four mechanisms of vigilance decrement discussed above. Placement of the response cutoff, specifically, determines the horizontal position of the function, sensitivity determines its steepness, and lapse and guess rates determine its upper and lower asymptotes. Changes in the psychometric function over time on task can therefore reveal mechanisms of vigilance loss. Studies using psychometric curves to analyze vigilance data have suggested that sensitivity losses are possible (McCarley & Yamani, 2021), but that changes of response cutoff, lapse rate, and guess rate are more common (Gyles et al., 2023; McCarley & Yamani, 2021; Román-Caballero, Martín-Arévalo, & Lupiáñez, 2022). Conservative cutoff shifts appear to account for the majority of the change in raw detection rates over time on task.

Sensory vs. Cognitive Vigilance

Much of what's known about the vigilance decrement has come from studies of sensory detection and discrimination. In tasks of this type, noise and signal conditions are distinguished by differences in physical stimulus properties, for example, shape (e.g., Helton & Warm, 2008; Nuechterlein et al., 1983), brightness (Broadbent & Gregory, 1965), spatial alignment (Dillard et al., 2019; Hitchcock, Dember, Warm, Moroney, & See, 1999) or size (Colquhoun & Baddeley, 1964; e.g., Deaton & Parasuraman, 1993; McCarley & Yamani, 2021). These can be distinguished from cognitive discrimination tasks, in which signal and noise stimuli are defined by symbolic properties (See et al., 1995). Given pairs of digits as stimuli, for instance, the signal in a cognitive discrimination task might be defined as a trial in which one digit is even and the other is odd (Deaton & Parasuraman, 1993) or a trial in which the difference between values is between -1 and +1 (Claypoole, Dever, Denues, & Szalma, 2019).

Comparisons between sensory and cognitive discrimination tasks are potentially valuable for theories of the vigilance decrement. A finding that the vigilance decrement was similar for sensory and cognitive tasks, for instance, would imply that vigilance failures occurred at a post-sensory, supramodal stage of information processing (cf., Greenlee, DeLucia, & Lui, 2022; Shaw et al., 2009). A finding that the effect was small or absent for cognitive tasks would implicate sensory or perceptual processing limitations as a major source of the conventional vigilance decrement.

Studies of cognitive vigilance, though, have produced wildly inconsistent effects. Experiments by Warm and colleagues (1984), using pairs of digits as stimuli, found a drop in detection rates over time when signal events were defined by a simple rule (digits differ by no more than ± 1), but a gradual increase in detection rates–that is, a vigilance increment–when signals were defined by a more complex rule (digits differ by no more than ± 1 and have a sum between 4 and 14). Other experiments failed to recreate the vigilance increment under conditions of high task complexity (Loeb, Noonan, Ash, & Holding, 1987), however, and a meta-analysis suggested that the effect might obtain only under very select task conditions (See et al., 1995). Several studies have reported null or modest effects of time on task in cognitive monitoring tasks (Deaton & Parasuraman, 1993; Koelega, Brinkman, Hendriks, & Verbaten, 1989; Loeb et al., 1987), and others have reported conventional vigilance decrements, including sensitivity losses (Claypoole et al., 2019; Claypoole & Szalma, 2018b, 2018a; Matthews, Davies, & Holley, 1993; Mouloua & Parasuraman, 1995) and apparent cutoff shifts (Matthews, Warm, Reinerman-Jones, Washburn, & Tripp, 2010).

Methodological considerations seem likely to explain some of these inconsistencies. In at least some studies, the tasks used to test cognitive vigilance might not have been well-suited for analysis using SDT. SDT presumes that states of knowledge are inherently continuous, and that errors arise from external and internal noise in the evidence representing alternative states of the world (Hautus et al., 2022; Wixted, 2020). Although this model will probably hold for sensory vigilance tasks, it might be less appropriate to some of the tasks used in studies of cognitive vigilance. For instance, consider the task of judging whether the difference between two one-digit numbers is less than or equal to one (Claypoole et al., 2019; Claypoole & Szalma, 2018b, 2018a; Warm et al., 1984), or whether one digit in a pair is even and the other is odd (Deaton & Parasuraman, 1993). In these cases, signal and noise categories are discrete, and assuming that stimuli are not perceptually degraded and processing is not terminated prematurely, encoding or decisional noise seems unlikely to cause confusability. Errors more probably reflect guesses or response blunders, violating the parametric assumptions of SDT. In cases like this, apparent sensitivity losses might be spurious.

A related concern is that performance in some tasks classified as cognitive might have actually been limited by sensory or perceptual processes. For example, a study by Mouloua and Parasuraman (1995) asked participants to monitor for occasional lowercase letters in a temporal stream of uppercase letters. Although the task was framed as a cognitive discrimination, vigilance losses were more pronounced when the task-relevant letters were surrounded by distractors, and when they were presented with spatial uncertainty in the visual periphery, than when they appeared alone in central vision. The vigilance decrement was thus largest when discriminability is likely to have been degraded by visual crowding and lateral masking (Bouma, 1970; Coates, Levi, Touch, & Sabesan, 2018; Loomis, 1978; Strasburger, Rentschler, & Juttner, 2011), implying that performance losses might have been sensory, not cognitive. Another experiment (Matthews et al., 1993) asked participants to identify targets in a stream of single digits presented in the central visual field, but degraded the stimuli with pixelated noise. Again, it seems possible that performance was limited more by sensory stimulus quality than by symbolic processing demands.

The goal of the current study was to test for vigilance effects in a cognitive monitoring task, isolating changes of response bias, sensitivity, mental lapse rate, and guess rate, with a task that meets the assumptions of signal detection analysis and is not limited by sensory or perceptual noise. To analyze performance, we develop a cognitive model motivated by the findings of psychometric function analyses reported in earlier studies (Gyles et al., 2023; McCarley & Yamani, 2021; Román-Caballero et al., 2022).

Modeling a Cognitive Vigilance Task

As described above, signal detection analysis of cognitive vigilance requires a task that meets at least two characteristics. First, discrete states of the world do not map onto discrete mental states, but are represented by variations in an internal decision variable that is continuous and contaminated by random error (Hautus et al., 2022; Wixted, 2020). Second, performance is not limited by sensory information quality, but by the quality of post-sensory processing. Toward that end, we adapted a numeric signal detection task (Healy & Kubovy, 1981) from earlier studies (Duncan-Reid & McCarley, 2021; Tikhomirov, Bartlett, Duncan-Reid, & McCarley, 2023) for use in a vigilance context. As shown in Figure 1, the stimulus each trial was a column of four three-digit readings sampled from one of two normal distributions, one that represented noise events and the other that represented signal events. Source distributions had a common standard deviation but differed in means. The participant's task each trial was to make a key press response if they judged that the displayed readings were drawn from the signal distribution. Accurate task performance thus required participants to mentally estimate the central tendency of the stimulus readings displayed each trial. Signals occurred randomly, with a probability of 0.20 each trial. Trials occurred at a pace of 40 per second, a rate that has been reported to be high enough to engender losses of sensitivity over time (Parasuraman, 1979).

For analysis, data were fit with a cognitive model built on SDT. As shown in Panel A of Figure 2, the model assumed that on some trials the participant rendered their decision in an attentive state, and that on the remaining trials, the participant lapsed into an inattentive state. The participant's state was determined randomly each trial, with the

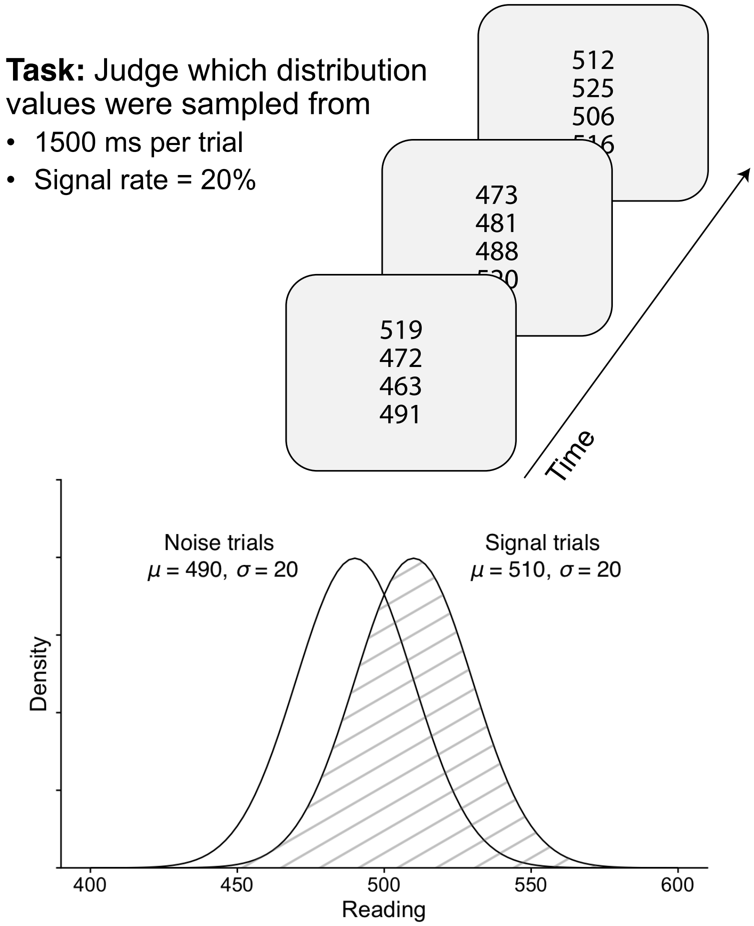

Figure 1. Schematic illustration of the stimuli and task.

lapse rate being the probability of an inattentive state. The assumption of discrete attentive and inattentive states is consistent with evidence from neural and behavioral-cognitive modeling (Hawkins, Mittner, Forstmann, & Heathcote, 2019; Mittner et al., 2014; Yamashita et al., 2021; Zanesco, Denkova, & Jha, 2021; Zeigenfuse & Lee, 2010) that processing on attention-demanding decision tasks can be modeled as a binary mixture distribution of identifiable latent states: a task-focused state generating efficient, stimulus-driven responses and an off-task state producing inefficient or contaminant responses. These binary attentive and inattentive states might themselves reflect a variety of underlying processing modes. The attentive state, for example, might be dissected into substates reflecting gradations of attentional focus (Zanesco, Denkova, Witkin, & Jha, 2020). Similarly, the inattentive state might be taken to subsume different forms of off-task thought (Unsworth & McMillan, 2014; Unsworth & Robison, 2016). To a reasonable approximation, though, attentive and inattentive states can be treated as binary.

Here, on attentive trials, the participant was assumed to make a stimulus-driven response using a conventional signal detection strategy, as illustrated in Panel B of Figure 2. Consistent with our earlier work (Duncan-Reid & McCarley, 2021; Tikhomirov et al., 2023), the model assumed that the participant used the estimated mean value of the displayed readings as a decision variable, but that the participant's estimate of the mean reading was contaminated by random error (Brezis, Bronfman, & Usher, 2015, 2018; cf., Brusovansky, Glickman, & Usher, 2018) modeled as zero-centered Gaussian noise. The participant rendered a judgment by comparing the estimated mean to a response cutoff. On the inattentive trials, the participant rendered a judgment by guessing, responding yes with a probability termed the *guess rate*.

Under the model, performance on the attentive trials was thus limited by the combination of variability in the stimulus source distributions and error in the participants' estimation of the mean readings. Performance on inattentive trials was determined by the guess rate; given the low frequency of signal events, a strategy of always guessing no would have maximized response accuracy. Attentive and inattentive trials were distinguished by the pattern of errors they produced. On attentive trials, errors would have been most common when the mean reading was very near the observer's response cutoff, and would have been rare when the mean reading was more extreme in either direction. On inattentive trials, on which responses were guessed, errors would have been equally common across all values of the stimulus readings.

To identify potential vigilance effects, lapse rate, estimation error, cutoff placement, and guess rate were allowed to vary between blocks of trials. Parameters were estimated using a Bayesian hierarchical procedure (Kruschke, 2015; Lee & Wagenmakers, 2013).

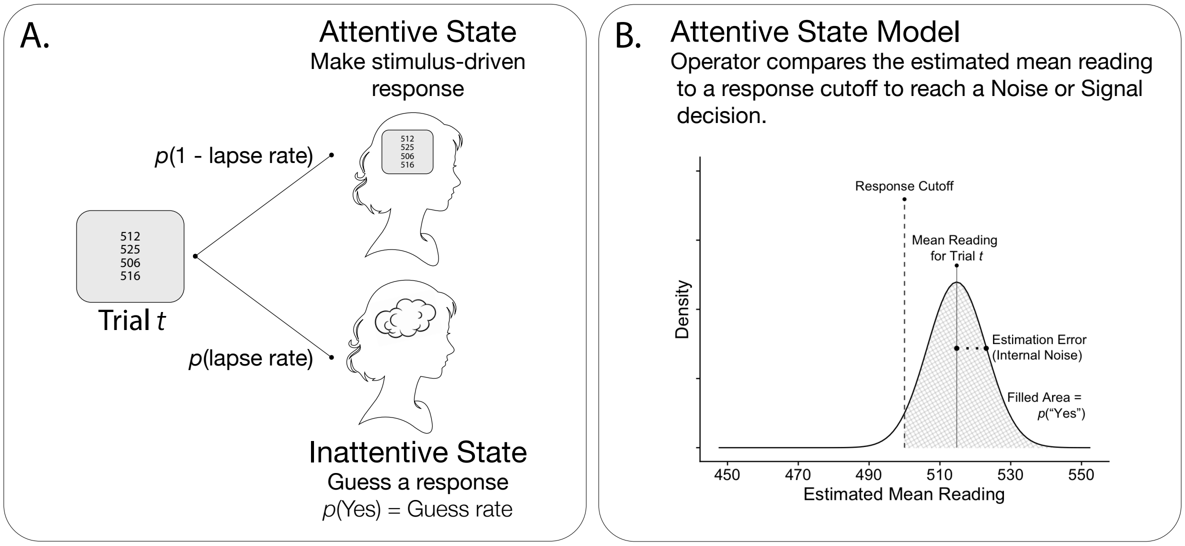

Figure 2. Schematic illustration of the model fit to the data.

Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The methods for this study were preregistered. Preregistration, data, and analytic code are available at https://osf.io/eytbz/?view_only=8dce3ec5d659450db3a5694435d21c8b. Deviations from the preregistered plan are noted below.

Participants

Two hundred participants were recruited from the online research platform Prolific (https://prolific.co/). Sample size was determined by a preregistered adaptive stopping rule. Under the stopping rule, we recruited an initial 125 participants then continued recruitment in increments of 25 until either, 1) the Bayes factors for the effects of block on response cutoff, estimation error, and lapse rate indicated an evidence ratio of at least 1:10 in either direction (i.e., in favor of or against the null), or, 2) sample size reached 200. Data collection ceased at 200 participants.

All participants gave informed consent and reported that they were fluent in English and had normal color vision and normal or corrected-to-normal visual acuity. Data exclusions described below left 180 participants for analysis (mean age = 22.58 years, gender = 73 females, 100 males, 4 non-binary, 3 not specified). Participants were reimbursed USD 5.00 for an experimental session lasting approximately 25 minutes.

Apparatus

Participants performed the experimental task online. The task was controlled by software written in PsychoPy (Peirce et al., 2019) and hosted on Pavlovia (https://pavlovia.org). Participation was restricted to participants using either laptop or desktop computers, not smartphones or tablet computers.

Procedure

Participants performed a numeric signal detection task after Healy and Kubovy (1978). The stimulus each trial was a set of four three-digit numeric readings, presented in a column in the center of the display. Readings were displayed in Arial font with a height 4% of the participant's display size.

On noise trials, the stimulus readings were sampled independently from a pseudorandom Gaussian distribution with $\mu = 490$ and $\sigma = 20$. On signal trials, they were sampled independently from a pseudorandom Gaussian distribution with $\mu = 510$ and $\sigma = 20$. Event type, noise or signal, was determined pseudorandomly each trial, with $p(\text{signal}) = 0.20$. We note that this signal rate (20%) is higher than in many traditional vigilance experiments, but matches that of our earlier studies of sensory vigilance (Gyles et al., 2023), which produced robust vigilance decrements. The participants' optimal strategy was to use the mean of the four readings each trial as the decision variable (Sorkin, Mabry, Weldon, & Elvers, 1991). Basing judgments on a single reading allowed a maximum sensitivity of $d' = 1.0$. Basing judgments on the mean of the four readings allowed a maximum sensitivity of $d' = 2.0$.

Experimental trials occurred at a forced pace of 1 every 1500 ms, producing an event rate of 40 trials per minute. The stimulus display appeared at the start of the trial and remained visible until the start of the next trial, giving an exposure duration of 1500 ms and an interstimulus interval of 0 ms. Participants were asked to press the space bar if they believed the readings on a given trial represented a signal state, and to withhold their response otherwise. A response was attributed to a given trial if it occurred any time between stimulus onset for that trial and stimulus onset for the subsequent trial. Participants did not receive post-trial feedback to indicate whether their judgments were correct or incorrect.

Written instructions were presented onscreen after the participant had indicated their consent to take part in the experiment. The instructions framed the task as cybersecurity monitoring. Participants were told, "For this task, imagine that you are a cybersecurity officer monitoring for malicious activity on your network. The system will provide a snapshot of network activity every 1.5 seconds. On every update, you will see four numbers, and each number represents the amount of activity on a single server. On average, normal network activity produces values below 500, and malicious activity produces values above 500. However, network activity is highly variable, meaning that there is no precise cut-off for detecting malicious activity. Normal activity will sometimes produce values above 500 and malicious activity will sometimes produce values below 500. For each trial, your job is to evaluate the set of four numbers and judge whether, collectively, they represent normal or malicious activity. If you think the set of numbers represents normal activity, you don't need to make a response. If you think the set of numbers represents malicious activity, you should press the space bar to report it. You will have 1.5 seconds to view the display and make a response before the system updates and new numbers appear."

After reading the instructions and indicating they were ready to proceed, the participant completed a practice vigil of 90 trials followed by a 12-minute experimental vigil. The vigil was limited to 12 minutes, a duration briefer than in many studies of sustained monitoring, in order to reduce the risk that online participants would withdraw before completing the task. Past work, including studies in our lab (Gyles et al., 2023), has found that 12 minutes on task is enough time to produce a detectable vigilance decrement in monitoring tasks (C. M. Craig & Klein, 2019; Neigel, Dever, Claypoole, & Szalma, 2019; Temple et al., 2000). The practice vigil was the same as the experimental vigil except that signal and noise events were equally probable, the first 25 trials occurred at a pace of 20 per minute (3000 ms/trial), and 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. Instructions in between the practice vigil and experimental vigil informed the participants, "You will now perform the task for a longer block and you will no longer receive feedback. Targets will also appear less frequently than they did during the practice."

To avoid potential end-spurt effects (Bergum & Lehr, 1963), participants were not told the exact length of the experimental vigil, but were aware that the entire session was expected to last less than 30 minutes.

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. The subscales were presented one at a time. Participants made their rating of each subscale by clicking a horizontal line anchored with text descriptions of the subscale endpoints. Ratings were scored on a scale of 0 to 100. Mental workload data were collected to characterize the task for comparison to earlier studies.

Data analysis

Data were analyzed using R (Version 4.4.1; R Core Team, 2023) and the R-packages *cowplot* (Version 1.1.3; Wilke, 2024), *dplyr* (Version 1.1.4; Wickham, François, Henry, Müller, & Vaughan, 2023), *forcats* (Version 1.0.0; Wickham, 2023), *ggplot2* (Version 3.5.1; Wickham, 2016), *jagsUI* (Version 1.6.2; Kellner, 2021), *magrittr* (Version 2.0.3; Bache & Wickham, 2022), *papaja* (Version 0.1.3; Aust & Barth, 2022), *purrr* (Version 1.0.2; Wickham & Henry, 2023), *tidybayes* (Version 3.0.7; Kay, 2023), *tidyr* (Version 1.3.1; Wickham, Vaughan, & Girlich, 2023), *tidyverse* (Version 2.0.0; Wickham et al., 2019) and *tinylabels* (Version 0.2.4; Barth, 2023).

Signal Detection. Trials from the 12-minute experimental block were grouped into three consecutive, non-overlapping blocks of four minutes each. As in our earlier studies (Gyles et al., 2023; McCarley & Yamani, 2021), a duration of four minutes was chosen in order to minimize the risk that substantial vigilance losses might occur within the first block of trials; past work has shown that the vigilance decrement can begin after less than five minutes on task (Nuechterlein et al., 1983). A preliminary screening was conducted to identify participants who might have misunderstood or failed to follow task instructions. For this, binary responses were converted to the sensitivity measure d' using the log-linear correction (Hautus, 1995), and data were excluded from participants who failed to achieve a preregistered minimum sensitivity of $d' = 0.25$ in any of the three experimental blocks. Twenty participants were excluded on the basis of this screening.

For the main analysis, response choice data were fit with a hierarchical Bayesian model based on SDT and earlier analyses of sensory vigilance data (Gyles et al., 2023; McCarley & Yamani, 2021; Román-Caballero et al., 2022). Analysis was restricted to the first and last four-minute blocks of trials, allowing us to test for differences between blocks using the Savage-Dickey ratio (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010), as discussed below.

The model assumed that the participant performed the detection task each trial in one of two states, attentive or inattentive, where the inattentive state reflected an attentional lapse. On attentive trials, the participant made a stimulus-driven choice by estimating the mean of the displayed stimulus readings and comparing it to a response cutoff. The participant's estimated mean reading for a given trial was equal to the true mean contaminated by zero-centered Gaussian noise (Brusovansky et al., 2018; Tikhomirov et al., 2023) of standard deviation τ. Sensitivity, the participant's ability to correctly distinguish signal from noise events, therefore decreased as τ increased. The participant classified the stimulus as noise or signal by comparing the estimated mean to a cutoff, κ, rendering a positive judgment if the decision variable exceeded the cutoff.

In the inattentive state, the participant chose a response by guessing, responding "yes" with probability π independent of the displayed stimulus readings. The participant's state each trial, attentive or inattentive, was determined randomly each trial. The probability of being in an inattentive state, or lapse rate, was ϕ. Internal noise τ, cutoff κ, lapse rate ϕ, and guess rate π were all allowed to vary between participants and between experimental blocks. The probability of participant i in block j performing trial k responding "yes" was thus,

$$p_{i,j,k}(\text{"yes"}) = (1 - \phi_{i,j}) \times \Phi(\frac{R_{i,j,k} - \kappa_{i,j}}{\tau_{i,j}}) + \phi_{i,j} \times \pi_{i,j},$$

where $R_{i,j,k}$ represents the mean of that trial's displayed stimulus readings and Φ represents the standard normal transformation.

Note that the inclusion of the guess rate as a free parameter is a deviation from the pregistered analysis plan, which assumed that the participants never responded in the inattentive state. Guess rate was added as a free parameter in light of results obtained following the preregistration (Gyles et al., 2023; Román-Caballero et al., 2022). Fixing π to a value of 0 did not substantially change the patterns of effect in other model parameters described below.

Values of $\tau_{i,j}$, $\kappa_{i,j}$, $\phi_{i,j}$, and $\pi_{i,j}$ reflected additive effects of the subject mean parameter values and the subject-specific effects of block. To ensure that values of the lapse rate ϕ and guess rate π remained between 0 and 1, the model placed priors on probit-transformed proportions rather than on raw values (Rouder & Lu, 2005). Likewise, to ensure positive values for the standard deviation of estimation error, the model placed priors on the log of τ^2 (Pratte & Rouder, 2011) rather on τ directly. Thus,

$$\log \tau^2_{i,j} = \begin{cases} \log \tau^2_i - 0.5 \times \Delta_i^{\log \tau^2}, \ j = \text{first}, \\ \log \tau^2_i + 0.5 \times \Delta_i^{\log \tau^2}, \ j = \text{last}, \end{cases}$$

$$\kappa_{i,j} = \begin{cases} \kappa_i - 0.5 \times \Delta_i^{\kappa}, \ j = \text{first}, \\ \kappa_i + 0.5 \times \Delta_i^{\kappa}, \ j = \text{last}, \end{cases}$$

$$\Phi(\phi_{i,j}) = \begin{cases} \Phi(\phi_i) - 0.5 \times \Delta_i^{\Phi(\phi)}, \ j = \text{first}, \\ \Phi(\phi_i) + 0.5 \times \Delta_i^{\Phi(\phi)}, \ j = \text{last}, \end{cases}$$

and,

$$\Phi(\pi_{i,j}) = \begin{cases} \Phi(\pi_i) - 0.5 \times \Delta_i^{\Phi(\pi)}, \ j = \text{first}, \\ \Phi(\pi_i) + 0.5 \times \Delta_i^{\Phi(\pi)}, \ j = \text{last}. \end{cases}$$

Here, $\log \tau^2_i$, κ_i, $\Phi(\phi_i)$, and $\Phi(\pi_i)$ are subject-level mean parameter values, and $\Delta_i^{\log \tau^2}$, Δ_i^{κ}, $\Delta_i^{\Phi(\phi)}$, and $\Delta_i^{\Phi(\pi)}$ are subject-level effects of block.

To maintain consistency and facilitate comparisons across parameters, the model placed unit normal priors on the group-level means of the standardized mean difference between blocks (Lee & Wagenmakers, 2013), rather than placing priors on the raw effects of block. The standard deviation of the difference between blocks was assigned a uniform distribution between 0 and 100. For example, for the parameter κ, $\Delta_i^{\kappa} = \delta_i^{\kappa} \times \sigma_i^{\kappa}$, $\delta^{\kappa} \sim \text{Normal}(0, 1)$, $\sigma^{\kappa} \sim \text{Uniform}(0, 100)$, where δ^{κ} is the standardized mean difference in cutoff placement between the first and last blocks.

Finally, subject-level means τ_i, κ_i, ϕ_i, and π_i were sampled from group-level distributions with vague priors. Ggroup-level means of $\log \tau^2$ and κ were assigned normal prior distributions with a mean of 0 and variance of 1000. Group-level values of the mean probit-transformed lapse and guess rates were assigned normal priors with a mean of 0 and standard deviation of 1, corresponding to uniform distributions over the interval (0.0, 1.0) on the raw lapse and guess rates.

Stimulus readings were zero-centered for analysis. Estimation was performed using JAGS (Plummer, 2019). The model was run for four MCMC chains of 10,000 warmup steps and 25,000 estimation steps each, providing 100,000 total MCMC steps for analysis. All parameter estimates showed R-hat values of less than 1.02, indicating satisfactory convergence of MCMC chains (Gelman & Rubin, 1992).

We used the Savage-Dickey ratio (Wagenmakers et al., 2010) to estimate Bayes factors for or against an effect of block on each of the four parameters of interest. We describe Bayes factors using the evidence categories (anecdotal, substantial, strong, very strong, decisive) proposed by Wetzels et al. (2011). As a check of model fit, we calculated 95% posterior predictive equal-tail Bayesian credible intervals (BCIs) on the basis of a random sample of 1000 steps from the MCMC chains.

Parameter recovery tests using simulated data (Heathcote, Brown, & Wagenmakers, 2015) confirmed that the model estimated true parameter values well, though with a tendency to underestimate differences in lapse rate between blocks. Simulated data and outputs of parameter recovery tests are available at the OSF site linked above.

Response Times. A non-preregistered analysis estimated the difference in mean response time (RT) between the first and last blocks of trials in order to test for a change in response speed over time on task. Data were subject-mean RTs for correct responses, calculated separately for the first and last trial blocks. Scores were assigned a normal likelihood with mean $\overline{RT}_{i,j}$ and residual variance σ_ϵ. $\overline{RT}_{i,j}$ reflected the sum of a subject-level grand mean RT, $\overline{RT}_i$, and a fixed effect of block, $\overline{RT}_j$. Subject-level grand means were assigned a normal prior with mean $\mu_{\overline{RT}}$ and standard deviation $\sigma_{\overline{RT}}$. Finally, $\mu_{\overline{RT}}$ was assigned a normal prior with mean 0 and variance 1000, truncated below 0, and $\sigma_{\overline{RT}}$ was assigned a uniform prior between 0 and 100. The Bayes factor for an effect of block, versus a point-null hypothesis of 0, was estimated using the Savage-Dickey ratio (Wagenmakers et al., 2010).

Workload Ratings. Responses for the A-SWAT subscales were analyzed separately within a model that placed a normal likelihood function on observed ratings, and uniform priors between 0 and 100 on the group means and standard deviations of the ratings. The estimation procedure again ran four MCMC chains for 10,000 warmup trials then 25,000 estimation trials each. All parameter estimates showed R-hat values of less than 1.01, indicating satisfactory convergence of MCMC chains.

Results

Signal Detection

Figure 3 shows empirical and posterior predictive yes rates as a function of trial type and trial block. Yes rates for signal trials are hit rates and yes rates for noise trials are false alarm rates. Posterior predictive BCIs are narrow and contain the empirical means, implying a satisfactory model fit. Yes rates for signal trials are clearly higher than those for noise trials, confirming that participants could discriminate signal from noise events at levels above chance.

Yes rates showed no obvious changes between the first and last blocks, $M_{Diff} = -0.010,\ 95\%\ \text{BCI}[-0.025, 0.006]$ for signal trials, $M_{Diff} = -0.008,\ 95\%\ \text{BCI}[-0.016, 0.002]$ for noise trials. Raw response rates thus gave no clear evidence of a vigilance decrement.

In contrast, analysis of latent model parameters produced substantial or strong

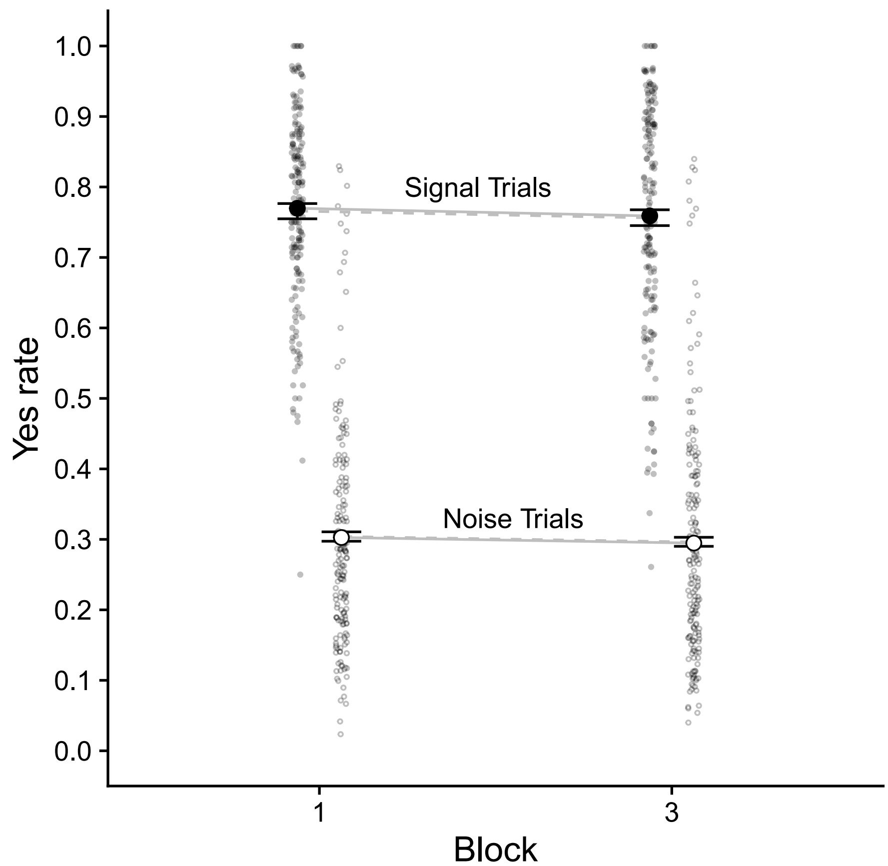

Figure 3. Empirical and posterior predictive yes rates, as a function of trial type and block. NOTE: Small symbols represent empirical means for individual participants, large symbols represent empirical group means, error bars represent 95% posterior predictive BCIs for group mean scores. Filled symbols correspond to hit rates, unfilled symbols correspond to false alarm rates.

evidence of changes in three out of four potential mechanisms of vigilance decrement. Figure 4 presents density plots of the standardized mean differences in decision model parameters between the first and last blocks on task. Note that the panel labeled Internal Noise depicts standardized differences in the log variance of participants' estimation error. The panels labeled Lapse Rate and Guess Rate present standardized differences in probit-transformed rates.

Data gave substantial evidence against a change in response cutoff over blocks, $M_{\text{Block 1}} = -1.39,\ 95\%\ \text{BCI}[-2.26, -0.53]$, $M_{\text{Block 3}} = -1.44,\ 95\%\ \text{BCI}[-2.30, -0.58]$, $M_{\text{Diff}} = -0.05,\ 95\%\ \text{BCI}[-0.59, 0.50]$, $B_{10} = \frac{1}{9.09}$. Data also gave substantial evidence for a decrease in internal noise–that is, an improvement in the ability to distinguish signal from noise–over blocks, $M_{\text{Block 1}} = 8.71,\ 95\%\ \text{BCI}[8.28, 9.16]$ for the standard deviation of the error in participants' estimates of the mean reading, $M_{\text{Block 3}} = 7.96,\ 95\%\ \text{BCI}[7.55, 8.39]$, $M_{\text{Diff}} = -0.75,\ 95\%\ \text{BCI}[-1.29, 0.21]$, $B_{10} = 6.01$.

Analyses of the remaining two parameters showed effects more consistent with a conventional vigilance decrement. Data gave strong evidence for an increase in the lapse rate between the first and last trial blocks, $M_{\text{Block 1}} = .05,\ 95\%\ \text{BCI}[0.03, 0.06]$, $M_{\text{Block 3}} = 0.07,\ 95\%\ \text{BCI}[0.06, 0.09]$, $M_{\text{Diff}} = 0.03\ 95\%\ \text{BCI}[0.01, 0.05]$, $B_{10} = 28.12$, and substantial evidence for a decrease in the guess rate, $M_{\text{Block 1}} = 0.70,\ 95\%\ \text{BCI}[0.57, 0.82]$, $M_{\text{Block 3}} = 0.56,\ 95\%\ \text{BCI}[0.45, 0.68]$, $M_{\text{Diff}} = -0.14\ 95\%\ \text{BCI}[-0.26, -0.01]$, $B_{10} = 3.17$. With time on task, that is, participants became more likely to lapse into inattentiveness and less likely to make a positive guess during a lapse.

On the surface, the finding that estimates of internal noise, lapse rate, and guess rate all changed over time on task appears inconsistent with the finding that mean yes rates were roughly constant. At least two potential explanations seem plausible. One possibility is that the effects of parameter changes on raw yes rates were simply negligible; effects that are large and statistically credible in a standardized latent variable, like the parameter

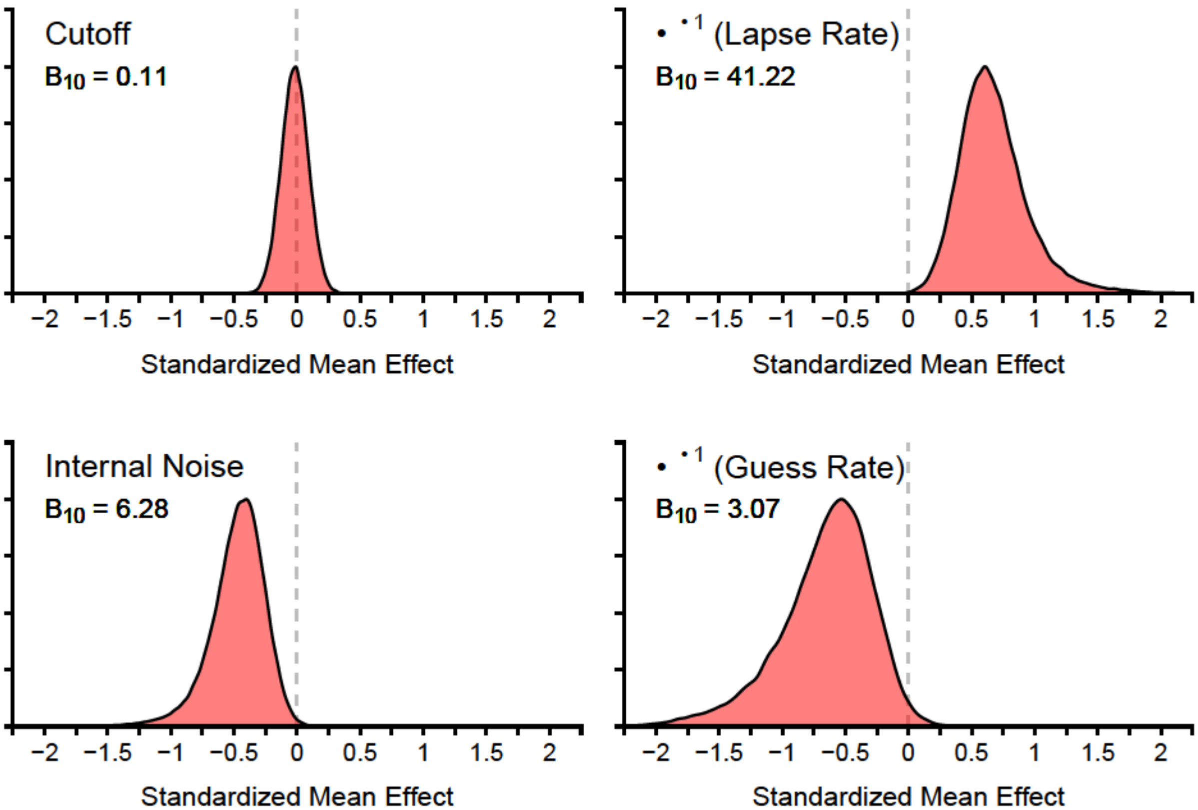

Figure 4. Posterior distributions of standardized mean differences in decision model parameters between the first and last blocks of trials. NOTE: Panel labeled *Internal Noise* depicts standardized differences in the log variance of participants' estimation error. Panels labeled *Lapse Rate* and *Guess Rate* present standardized differences in probit-transformed rates.

estimates shown in Figure 4, might not translate to large effects in an observed measure (Pek & Flora, 2018). A second possibility is that changes in internal noise and inattentiveness might have offset one another in the yes rates for signal trials. By itself, a decrease in internal noise would tend to improve signal-noise discriminability and raise signal detection rates. Conversely, an increase in lapse rate and a decrease in guess rate would tend to reduce target detection rates. This makes it possible that the changes in internal noise and lapse rate/guess rate might have effectively washed out within the yes rates to signal events.

To test these two possibilities, we conducted non-preregistered analyses to estimate the selective effects of internal noise changes and lapse and guess rate changes on yes rates. In the first case, we generated posterior predictive yes rates incorporating the effects of block on internal noise, but holding the cutoff, lapse rate, and guess rate fixed at their mean values for each participant. In the second case, we generated posterior predictive data incorporating the effects of block on lapse and guess rates, but holding the cutoff placement and internal noise estimate fixed at their mean values for each participant.

Results provide evidence for both the possibilities discussed above. As expected, changes in internal noise and inattention parameters had opposite effects on yes rates for signal events. In both cases, though, the effects of parameter changes on raw yes rates were small. In isolation, decreases in internal noise between blocks 1 and 3 would have increased hit rates by about 1 percentage point, $M_{\text{Block 1}} = 0.759,\ 95\%\ \text{BCI}[0.749, 0.769]$, $M_{\text{Block 3}} = 0.769,\ 95\%\ \text{BCI}[0.779, 0.780]$, $M_{\text{Diff}} = 0.010,\ 95\%\ \text{BCI}[-0.003, 0.022]$, and decreased false alarm rates by about the same amount, $M_{\text{Block 1}} = 0.304,\ 95\%\ \text{BCI}[0.299, 0.310]$, $M_{\text{Block 3}} = 0.295,\ 95\%\ \text{BCI}[0.289, 0.301]$, $M_{\text{Diff}} = -0.009,\ 95\%\ \text{BCI}[-0.017, -0.002]$. Conversely, changes in lapse and guess rates would have reduced hit rates by about 1 percentage point, $M_{\text{Block 1}} = 0.770,\ 95\%\ \text{BCI}[0.759, 0.780]$, $M_{\text{Block 3}} = 0.755,\ 95\%\ \text{BCI}[0.744, 0.767$, $M_{\text{Diff}} = -0.014,\ 95\%\ \text{BCI}[-0.030, 0.001]$, and reduced false alarm rates by less than 1 percentage point, $M_{\text{Block 1}} = 0.301,\ 95\%\ \text{BCI}[0.295, 0.308]$, $M_{\text{Block 3}} = 0.297,\ 95\%\ \text{BCI}[0.290, 0.304$, $M_{\text{Diff}} = -0.004,\ 95\%\ \text{BCI}[-0.014, 0.005]$. Altogether, results suggest that changes in internal noise and in lapse and guess rates had very small, and partially offsetting, effects on raw yes rates.

Robustness Checks. As a check on the robustness of the results reported above, we conducted two additional, non-preregistered analyses. The first was intended to confirm that results were not distorted by the exclusion of data from the middle four minutes. For this, the analysis described above was repeated, except that the Block 1 data included all trials from the first six minutes of the twelve-minute vigil and the Block 2 data included all trials from the second six minutes. Results again gave substantial evidence ($B_{10} = 0.11$) against a change in cutoff over time,strong evidence ($B_{10} = 13.85$) for an increase in lapse rate, and substantial evidence for a decrease in the guess rate ($B_{10} = 5.93\}$. Data trended again toward a decrease in internal noise between blocks 1 and 2, $M_{\text{Diff}} = -0.49$ 95% BCI$[-0.97, -0.00]$. However, the Bayes factor for this effect was now indifferent between the alternative and null hypotheses ($B_{10} = 1.08$), implying that the changes in internal noise might have largely occurred by roughly midway through the vigil.

As the second robustness check, we used deviance information criterion (DIC) (Spiegelhalter, Best, Carlin, & van der Linde, 2002) scores to compare versions of the model above selectively excluding effects of block on internal noise, cutoff, and lapse and guess rates. Consistent with the results described above, DIC values favored a model including effects of block on internal noise, lapse rate and guess rate, but excluding an effect of block on response cutoff placement.

Full results of the robustness checks are included at the OSF site linked above.

Response Times. Estimated group mean RT was $M_{\text{Block 1}} = 819$ ms, 95% BCI$[798, 839]$ for the first block of trials and $M_{\text{Block 3}} = 816$ ms, 95% BCI$[795, 836]$ for the third block of trials, $M_{\text{Diff}} = -3$ ms, 95% BCI$[-24, 15]$, $B_{10} = \frac{1}{3199}$. Data thus gave decisive evidence against a change in mean RT between blocks.

Subjective Workload. Estimated group mean ratings were $M = 23.41$, 95% BCI$[20.11, 26.73]$ for the time stress subscale of the ASWAT, $M = 80.48$, 95% BCI$[77.99, 82.97]$ for the mental effort subscale, and $M = 52.62$, 95% BCI$[49.13, 56.11]$ for the psychological stress subscale. Values were very similar to those observed in our earlier online experiments using a sensory vigilance task (Gyles et al., 2023) and are consistent with other evidence indicating that vigilance tasks are mentally demanding (Claypoole et al., 2019; e.g., Deaton & Parasuraman, 1993; Warm, Dember, & Hancock, 1996).

Discussion

Research on cognitive vigilance has produced inconsistent results, with some studies reporting vigilance losses, others showing null effects, and others reporting gains over time on task. The current experiment tested for a cognitive decrement in a task designed to minimize sensory performance constraints and conform to the assumptions of signal detection theory. Participants performed a vigilance task that asked them to monitor for infrequent signal events within a stream of numeric readings. The stimulus each trial was a set of four three-digit numbers, sampled from either a signal or non-signal distribution. Participants were asked to make a keypress response on any trial in which they judged that the readings were drawn from the signal distribution. Signals occurred with probability of 0.20.

Observed data of central interest were yes rates (i.e., positive responses rates) to signal and non-signal events. Data were fit with a model that assumed the participant made their judgment each trial in either an attentive or inattentive state. In the attentive state, the participant selected a response by estimating the mean of the stimulus readings and comparing it to a decision cutoff. In the inattentive state, they selected a response by guessing. Four parameters were allowed to vary across blocks of trials: internal noise corrupting the participant's estimates of the mean reading; cutoff placement; the probability of lapsing into the inattentive state; and the probability of guessing a positive response from inside the inattentive state.

Model fits suggested that time on task affected processing both by increasing noise in the participants' estimates of mean readings, and by increasing the lapse rate and decreasing the guess rate. Posterior predictive data confirmed that by itself, an increase in internal noise would have tended to increase hit rates and decrease false alarm rates. In contrast, the observed change in lapse and guess rates would have tended to drive hit and false alarm rates both downward. As manifest in observable data, however, these effects were small, however, and partially offset one another. As as result, yes rates showed little change between the first and last blocks of the task.

Results imply similarities and dissimilarities between cognitive and sensory vigilance decrements. One notable point of difference is the absence of a conservative shift of response bias in the current data. Conservative cutoff shifts have been nearly ubiquitous in sensory vigilance tasks (Broadbent & Gregory, 1965; A. Craig, 1987; Swets, 1977) and have also been reported in some cognitive vigilance tasks (Claypoole et al., 2019). In contrast, cutoff placement here was stable over blocks of trials. Two characteristics of the current experiments seem likely to have allowed participants to maintain a fixed cutoff (Kubovy, Rapoport, & Tversky, 1971). First, because stimuli were digital, participants could hold a deterministic cutoff value in verbal memory, rather than relying on an implicit and potentially noisy (Benjamin, Diaz, & Wee, 2009) cutoff representation in sensory memory. Second, because the task instructions explained that noise values were generally less than 500 and signal values greater than 500, participants were not required to discover an appropriate cutoff value through learning (Erev, 1998).

The current results also differ from those of sensory vigilance studies in indicating a reduction in observers' internal noise over time. Here, participants showed a small but credible decrease between blocks in the random error contaminating their estimates of the mean reading. This result mirrors the findings of Warm et al. (1984), who reported an improvement in signal detection rates over time in a cognitive vigilance task. The current data do not tell us the cause of this effect. One possibility, consistent with Warm et al. (1984)'s explanation of their own findings, is that participants became more motivated over time, investing greater effort in the task. However, this seems inconsistent with the finding that estimated lapse rates rose over over time.

A different possibility is that participants modified their task strategies over time.

One source of error in the participants' judgments, for instance, might have been a tendency to attend to fewer than all four of the readings displayed each trial. Thus, a participant who began the vigil attending to a subset of the readings each trial and gradually expanded their attention to incorporate a larger subset would have shown a decrease in estimation noise over time. A third possibility is that participants simply learned to estimate the mean reading more accurately with practice. But under any of these accounts, performance was robust against increases internal noise between blocks. Modeling gave no evidence that on the trials in which they made stimulus-driven responses, participants lost sensitivity over time. Notably, this result is inconsistent with the resource depletion theory of vigilance losses, which holds that the gradual consumption of processing resources over time on task reduces the observer's ability to discern signal from noise (Caggiano & Parasuraman, 2004; Parasuraman, 1979; Warm et al., 2015).

Finally, data showed evidence for a conventional vigilance decrement in a tendency for lapse rates to increase over time. This effect matches results seen in sensory vigilance (Gyles et al., 2023; McCarley & Yamani, 2021; Román-Caballero et al., 2022) and speeded response (Unsworth & Robison, 2016) tasks, and suggests that lapses are a very general mechanism of vigilance failure, showing up across various forms of monitoring tasks. The mean increase in estimated lapse rates between blocks 1 and 3 was modest, roughly 3%. It is possible that this value underestimates the true difference between blocks, since, as noted above, parameter recovery exercises indicated a tendency for the model to underestimates changes lapse rate. At best, though, the increase in lapse rate from the first to last block was too small to manifest as a detectable change in yes rates.

The current data do not reveal the nature of the participants' occasional lapses, for example, whether they reflected external distraction (Drody et al., 2023; Robison & Unsworth, 2015; Unsworth & McMillan, 2014), microsleeps (Buckley et al., 2016), or mind-wandering (McVay & Kane, 2009, 2012) or other breakdowns of task-set maintenance (Ariga & Lleras, 2011). However, given that off-task thoughts are common (McVay &

Kane, 2009) and increase with time on task (Kane et al., 2007; Zanesco et al., 2024), they seem very likely to be at least one source of the lapses observed here. Interestingly, the average estimated lapse rate here, collapsed across blocks, was roughly 6%. This value closely matches the rate at which participants in an early study reported themselves as being off-task in response to occasional thought probes (6.09%), and is similar to the rate at which a Markov chain model using probe response and choice RT data estimated those participants to be in a discrete state of full task disengagement (8.40%) (Zanesco et al., 2020). The correspondence between these values hints that the attentional lapses inferred from the current data might reflect the same state of mental disengagement identified by participants' self-reports. The current data do not indicate whether this potential increase in off-task thoughts was intentional, or was the result of attention control failures (Kane & McVay, 2012; McVay & Kane, 2009; Thomson et al., 2015).

Additional research will be necessary to test whether the pattern of effects seen here holds over longer vigils or variations in signal rate. Of note, the effects of time on task in the current data were strikingly smaller than those seen in some earlier work. In the study reported by Claypoole et al. (2019), for instance, hit rates dropped by roughly 35 percentage points from the first 6-minute block of trials to the second. In the experiments reported by Warm et al. (1984), hit rates in several variants of the task under study fell by 10 percentage points or more between the first and second 20-minute blocks on task. In contrast, lapse rates in the current data fell by only 2 percentage points over the course of a 12-minute vigil. Although all of these experiments used digits as stimuli, and required participants to perform mental arithmetic operations, they differed in multiple other ways. Participants in the current study were asked to estimate the mean of a set of 3-digit numbers, whereas those in Claypoole et al. (2019) and Warm et al. (1984) were asked to find the sum and difference of pairs of single digits. Maybe more notably, signal rate in the current task (20%) was substantially higher than in either Warm et al. (1984) (10% or lower) or Claypoole et al. (2019) (2% or lower). These comparisons suggest that the magnitude of the cognitive vigilance decrement might be highly sensitive to task demands, and that increases in lapse rate might not be the sole mechanism of cognitive vigilance loss.

Statements and Declarations

Funding

This work was supported by a grant from the National Science Foundation (2240256) to Y.Y. and J.S.M. and a grant from the Human Factors and Ergonomics Society Perception and Performance Technical Group to S.G.

Conflicts of interest/Competing interests

The authors report no conflicts of interest.

Availability of data and material

Data are available at https://osf.io/eytbz/?view_only=8dce3ec5d659450db3a5694435d21c8b.

Code availability

Analytic code is available at https://osf.io/eytbz/?view_only=8dce3ec5d659450db3a5694435d21c8b.

Authors' contributions

The authors made the following contributions. Shannon Gyles: Conceptualization, Investigation, Software, Formal Analysis, Methodology, Writing - Original Draft Preparation, Writing - Review & Editing, Funding Acquisition; Yusuke Yamani: Conceptualization, Writing - Review & Editing, Funding Acquisition; Jason S. McCarley: Conceptualization, Investigation, Formal Analysis, Methodology, Writing - Original Draft Preparation, Writing - Review & Editing, Funding Acquisition.

Ethics approval

This project was approved by the Institutional Review Board of Oregon State University (protocol #8691).

Consent to participate

All participants consented to participate.

Consent for publication

NA

References

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

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. Retrieved from https://github.com/crsh/papaja

Bache, S. M., & Wickham, H. (2022). *Magrittr: A forward-pipe operator for r*. Retrieved from https://CRAN.R-project.org/package=magrittr

Barth, M. (2023). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, *116*(1), 84–115. https://doi.org/10.1037/a0014351

Bergum, B. O., & Lehr, D. J. (1963). End spurt in vigilance. *Journal of Experimental Psychology*, *66*(4), 383–385. https://doi.org/10.1037/h0044865

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

Brezis, N., Bronfman, Z. Z., & Usher, M. (2015). Adaptive Spontaneous Transitions between Two Mechanisms of Numerical Averaging. *Scientific Reports*, *5*(1), 10415. https://doi.org/10.1038/srep10415

Brezis, N., Bronfman, Z. Z., & Usher, M. (2018). A Perceptual-Like Population-Coding Mechanism of Approximate Numerical Averaging. *Neural Computation*, *30*(2), 428–446. https://doi.org/10.1162/neco_a_01037

Broadbent, D. E., & Gregory, M. (1963). Vigilance considered as a statistical decision. *British Journal of Psychology*, *54*(4), 309–323. https://doi.org/10.1111/j.2044-8295.1963.tb00886.x

Broadbent, D. 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

Brusovansky, M., Glickman, M., & Usher, M. (2018). Fast and effective: Intuitive processes in complex decisions. *Psychonomic Bulletin & Review*, *25*(4), 1542–1548. https://doi.org/10.3758/s13423-018-1474-1

Buckley, R. J., Helton, W. S., Innes, C. R. H., Dalrymple-Alford, J. C., & Jones, R. D. (2016). Attention lapses and behavioural microsleeps during tracking, psychomotor vigilance, and dual tasks. *Consciousness and Cognition*, *45*, 174–183. https://doi.org/10.1016/j.concog.2016.09.002

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

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

Claypoole, V. L., & Szalma, J. L. (2018a). Facilitating Sustained Attention: Is Mere Presence Sufficient? *The American Journal of Psychology*, *131*(4), 417–428. https://doi.org/10.5406/amerjpsyc.131.4.0417

Claypoole, V. L., & Szalma, J. L. (2018b). Independent Coactors May Improve Performance and Lower Workload: Viewing Vigilance Under Social Facilitation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *60*(6), 822–832. https://doi.org/10.1177/0018720818769268

Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018). Foveal Crowding Resolved. *Scientific Reports*, *8*(1), 9177. https://doi.org/10.1038/s41598-018-27480-4

Colquhoun, W. P., & Baddeley, A. D. (1964). Role of pretest expectancy in vigilance decrement. *Journal of Experimental Psychology*, *68*(2), 156–160.

https://doi.org/10.1037/h0042875

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

Craig, A. (1978). Is the vigilance decrement simply a response adjustment towards probability matching? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *20*(4), 441–446. https://doi.org/10.1177/001872087802000408

Craig, A. (1987). Signal detection theory and probability matching apply to vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(6), 645–652. https://doi.org/10.1177/001872088702900604

Craig, C. M., & Klein, M. I. (2019). The abbreviated vigilance task and its attentional contributors. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(3), 426–439. https://doi.org/10.1177/0018720818822350

Cunningham, S., Scerbo, M. W., & Freeman, F. G. (2000). The electrocortical correlates of daydreaming during vigilance tasks. *Journal of Mental Imagery*, *24*(1 & 2), 61–72.

Deaton, J. E., & Parasuraman, R. (1993). Sensory and cognitive vigilance: Effects of age on performance and subjective workload. *Human Performance*, *6*(1), 71–97.

Dillard, M. B., Warm, J. S., Funke, G. J., Nelson, W. T., Finomore, V. S., McClernon, C. K., … Funke, M. E. (2019). Vigilance Tasks: Unpleasant, Mentally Demanding, and Stressful Even When Time Flies. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(2), 225–242. https://doi.org/10.1177/0018720818796015

Drody, A. C., Pereira, E. J., & Smilek, D. (2023). A desire for distraction: Uncovering the rates of media multitasking during online research studies. *Scientific Reports*, *13*(1), 781. https://doi.org/10.1038/s41598-023-27606-3

Duncan-Reid, J., & McCarley, J. S. (2021). Strategy use in automation-aided decision making. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 96–100. https://doi.org/10.1177/1071181321651259

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review*, *105*(2), 280–298.

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

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

Greenlee, E. T., DeLucia, P. R., & Lui, T. G. (2022). Modality changes in vigilance displays: Further evidence of supramodal resource depletion in vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872082210997. https://doi.org/10.1177/00187208221099793

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

Gyles, S. P., McCarley, J. S., & Yamani, Y. (2023). Psychometric curves reveal changes in bias, lapse rate, and guess rate in an online vigilance task. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02652-1

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Hautus, M. J., Macmillan, N. A., & Creelman, C. D. (2022). *Detection theory: A user's guide* (Third edition). New York, NY: Routledge.

Hawkins, G. E., Mittner, M., Forstmann, B. U., & Heathcote, A. (2019). Modeling distracted performance. *Cognitive Psychology*, *112*, 48–80.

https://doi.org/10.1016/j.cogpsych.2019.05.002

Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition*, *6*(5), 544–553. https://doi.org/10.3758/BF03198243

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), 344–354.

Heathcote, A., Brown, S. D., & Wagenmakers, E.-J. (2015). An Introduction to Good Practices in Cognitive Modeling. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (pp. 25–48). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-2236-9_2

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

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*.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For Whom the Mind Wanders, and When: An Experience-Sampling Study of Working Memory and Executive Control in Daily Life. *Psychological Science*, *18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348–354. https://doi.org/10.1177/0963721412454875

Kay, M. (2023). *tidybayes: Tidy data and geoms for Bayesian models.* https://doi.org/10.5281/zenodo.1308151

Kellner, K. (2021). *jagsUI: A wrapper around 'rjags' to streamline 'JAGS' analyses.*

Retrieved from https://CRAN.R-project.org/package=jagsUI

Kingdom, F. A. A., & Prins, N. (2016). *Psychophysics: A practical introduction* (2nd ed.). Amsterdam: Elsevier Science.

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: The Journal of the Human Factors and Ergonomics Society*, *31*(1), 45–62. https://doi.org/10.1177/001872088903100104

Krimsky, M., Forster, D. E., Llabre, M. M., & Jha, A. P. (2017). The influence of time on task on mind wandering and visual working memory. *Cognition*, *169*, 84–90. https://doi.org/10.1016/j.cognition.2017.08.006

Kruschke, J. K. (2015). *Doing Bayesian data analysis* (2nd ed.). London: Academic Press/Elsevier.

Kubovy, M., Rapoport, A., & Tversky, A. (1971). Deterministic vs probabilistic strategies in detection. *Perception & Psychophysics*, *9*(5), 427–429. https://doi.org/10.3758/BF03210245

Kuss, M., Jäkel, F., & Wichmann, F. A. (2005). Bayesian inference for psychometric functions. *Journal of Vision*, *5*(5), 8. https://doi.org/10.1167/5.5.8

Lee, M. D. (2018). Bayesian methods in cognitive modeling. In J. T. Wixted & E.-J. Wagenmakers (Eds.), *The Stevens' handbook of experimental psychology and cgnitive neuroscience, volume 5: Methodology.* (4th ed., Vol. 5, pp. 37–84). Hoboken, N.J.: John Wiley & Sons, Inc.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* New York: Cambridge University Press.

Loeb, M., Noonan, T. K., Ash, D. W., & Holding, D. H. (1987). Limitations of the Cognitive Vigilance Increment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(6), 661–674. https://doi.org/10.1177/001872088702900606

Loomis, J. M. (1978). Lateral masking in foveal and eccentric vision. *Vision Research*, *18*(3), 335–338. https://doi.org/10.1016/0042-6989(78)90168-2

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

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*(1), 6–21. https://doi.org/10.1080/17470214808416738

Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of "nonparametric" measures of sensitivity and response bias. *Psychonomic Bulletin & Review*, *3*(2), 164–170. https://doi.org/10.3758/BF03212415

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*, 661–670. https://doi.org/10.1016/S0028-3932(98)00127-4

Matthews, G., Davies, D. R., & Holley, P. J. (1993). Cognitive predictors of vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(1), 3–24. https://doi.org/10.1177/001872089303500101

Matthews, G., Warm, J. S., Reinerman-Jones, 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), 187–203. https://doi.org/10.1037/a0019572

McCarley, J. S., & Yamani, Y. (2021). Psychometric curves reveal three mechanisms of vigilance decrement. *Psychological Science*, *32*(10), 1675–1683. https://doi.org/10.1177/09567976211007559

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

Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., & Forstmann, B. U. (2014). When the brain takes a break: A model-based analysis of mind wandering. *The Journal of Neuroscience*, *34*(49), 16286–16295. https://doi.org/10.1523/JNEUROSCI.2062-14.2014

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, *38*(2), 311–322.

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

Neigel, A. R., Claypoole, V. L., Smith, S. L., Waldfogle, G. E., Fraulini, N. W., Hancock, G. M., … Szalma, J. L. (2020). Engaging the human operator: A review of the theoretical support for the vigilance decrement and a discussion of practical applications. *Theoretical Issues in Ergonomics Science*, *21*(2), 239–258. https://doi.org/10.1080/1463922X.2019.1682712

Neigel, A. R., Dever, D. A., Claypoole, V. L., & Szalma, J. L. (2019). Task Engagement and the Vigilance Decrement Revisited: Expanding Upon the Work of Joel S. Warm Using a Semantic Vigilance Paradigm. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(3), 462–473. https://doi.org/10.1177/0018720819835086

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. (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., & 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

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological Methods*, *23*(2), 208–225. https://doi.org/10.1037/met0000126

Plummer, M. (2019). *Rjags: Bayesian graphical models using MCMC.*

Pratte, M. S., & Rouder, J. N. (2011). Hierarchical single- and dual-process models of recognition memory. *Journal of Mathematical Psychology*, *55*(1), 36–46. https://doi.org/10.1016/j.jmp.2010.08.007

Proctor, R. W., & Vu, K.-P. L. (2023). *Attention: Selection and control in human information processing.* Washington, DC: American Psychological Association.

R Core Team. (2023). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reinerman-Jones, L., Matthews, G., & Mercado, J. E. (2016). Detection tasks in nuclear power plant operation: Vigilance decrement and physiological workload monitoring. *Safety Science*, *88*, 97–107. https://doi.org/10.1016/j.ssci.2016.05.002

Robison, M. K., & Unsworth, N. (2015). Working memory capacity offers resistance to Mind-Wandering and External Distraction in a Context-Specific Manner. *Applied*

Cognitive Psychology, *29*(5), 680–690. https://doi.org/10.1002/acp.3150

Román-Caballero, R., Martín-Arévalo, E., & Lupiáñez, J. (2022). Changes in response criterion and lapse rate as general mechanisms of vigilance decrement: Commentary on McCarley and yamani (2021). *Psychological Science*, 095679762211213. https://doi.org/10.1177/09567976221121342

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *12*(4), 573–604.

Schumann, F., Steinborn, M. B., Kürten, J., Cao, L., Händel, B. F., & Huestegge, L. (2022). Restoration of Attention by Rest in a Multitasking World: Theory, Methodology, and Empirical Evidence. *Frontiers in Psychology*, *13*, 867978. https://doi.org/10.3389/fpsyg.2022.867978

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), 230–349.

Seli, P., Risko, E. F., & Smilek, D. (2016). On the Necessity of Distinguishing Between Unintentional and Intentional Mind Wandering. *Psychological Science*, *27*(5), 685–691. https://doi.org/10.1177/0956797616634068

Shaw, T. H., Warm, J. S., Finomore, V., Tripp, L., Matthews, G., Weiler, E., & Parasuraman, R. (2009). Effects of sensory modality on cerebral blood flow velocity during vigilance. *Neuroscience Letters*, *461*(3), 207–211. https://doi.org/10.1016/j.neulet.2009.06.008

Sorkin, R. D., Mabry, T. R., Weldon, M. S., & Elvers, G. (1991). Integration of information from multiple element displays. *Organizational Behavior and Human Decision Processes*, *49*(2), 167–187. https://doi.org/10.1016/0749-5978(91)90047-W

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*(5), 13–13. https://doi.org/10.1167/11.5.13

Swets, J. A. (1977). Signal detection theory applied to vigilance. In R. R. Mackie (Ed.), *Vigilance* (pp. 705–718). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4684-2529-1_34

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. (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

Tikhomirov, L., Bartlett, M. L., Duncan-Reid, J., & McCarley, J. S. (2023). Identifying inefficient strategies in automation-aided signal detection. *Journal of Experimental Psychology: Applied*, *29*(4), 869–886. https://doi.org/10.1037/xap0000484

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica*, *150*, 14–25. https://doi.org/10.1016/j.actpsy.2014.04.001

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 601–615. https://doi.org/10.3758/s13415-016-0417-4

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, *60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In *Human Factors in Transportation. Automation and human performance: Theory and applications.* (pp. 183–200). Hillsdale, NJ: CRC Press.

Warm, J. S., Finomore, V. S., Vidulich, M. A., & Funke, M. E. (2015). Vigilance: A perceptual challenge. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman, & J. L. Szalma (Eds.), *The Cambridge Handbook of Applied Perception Research* (pp. 241–283). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511973017.018

Warm, J. S., Howe, R., Dember, W. N., & Sprague, L. (1984). Cognitive Demand and the Vigilance Decrement. In *Trends in Ergonomics/Human Factors* (pp. 15–20). Elsevier.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 *t* tests. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2023). *Forcats: Tools for working with categorical variables (factors).* Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools.* Retrieved from https://CRAN.R-project.org/package=purrr

Wickham, H., Vaughan, D., & Girlich, M. (2023). *Tidyr: Tidy messy data.* Retrieved from https://CRAN.R-project.org/package=tidyr

Wilke, C. O. (2024). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.* Retrieved from https://CRAN.R-project.org/package=cowplot

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 201–233. https://doi.org/10.1037/xlm0000732

Yamashita, A., Rothlein, D., Kucyi, A., Valera, E. M., Germine, L., Wilmer, J., . . . Esterman, M. (2021). Variable rather than extreme slow reaction times distinguish brain states during sustained attention. *Scientific Reports*, *11*(1), 14883. https://doi.org/10.1038/s41598-021-94161-0

Zanesco, A. P., Denkova, E., & Jha, A. P. (2021). Self-reported mind wandering and response time variability differentiate prestimulus electroencephalogram microstate dynamics during a sustained attention task. *Journal of Cognitive Neuroscience*, *33*(1), 28–45. https://doi.org/10.1162/jocn_a_01636

Zanesco, A. P., Denkova, E., & Jha, A. P. (2024). Mind-wandering increases in frequency over time during task performance: An individual-participant meta-analytic review. *Psychological Bulletin.* https://doi.org/10.1037/bul0000424

Zanesco, A. P., Denkova, E., Witkin, J. E., & Jha, A. P. (2020). Experience sampling of the degree of mind wandering distinguishes hidden attentional states. *Cognition*, *205*, 104380. https://doi.org/10.1016/j.cognition.2020.104380

Zeigenfuse, M. D., & Lee, M. D. (2010). A general latent assignment approach for modeling psychological contaminants. *Journal of Mathematical Psychology*, *54*(4), 352–362. https://doi.org/10.1016/j.jmp.2010.04.001