Vigilance Performance Declines Faster When Monitoring for a Signal in Two Modalities Compared to One

Chad Peltier[1], Matthew Daley[1,2], Sylvia Guillory[1, 2], Justin D. Handy[1], and Bridget Wilson[1, 2]

[1] Naval Submarine Medical Research Laboratory, Submarine Base New London, Groton, CT 06349-5900

[2] Leidos, New London, CT, 06320

ABSTRACT

Vigilance is the maintenance of attention over prolonged periods, often required when attempting to detect infrequent and/or difficult to detect stimuli, such as in baggage screening or sonar monitoring. This type of attention is characterized by the vigilance decrement: longer reaction times and decreased accuracy as time-on-task increases. Research previously demonstrated the vigilance decrement in auditory and visual vigilance tasks. However, little research has compared the strength and onset of the vigilance decrement in unimodal (auditory or visual) versus bimodal (auditory and visual) modalities. This knowledge gap was investigated in an experiment that first equated the discriminability of stimulus type at ~80% to control for stimulus difficulty and then by tracking subjects' target identification rate and reaction time for a target intermixed with a nontarget across three conditions: auditory, visual, and audiovisual. Overall, accuracy was worse in the bimodal condition relative to the unimodal condition. Target detection accuracy in the auditory bimodal condition declined more over time relative to the auditory unimodal task, with reaction time data suggesting the decrease was not due to a speed-accuracy trade-off. Results indicate that monitoring for targets in two modalities is more difficult, resulting in a greater vigilance decrement than unimodal vigilance.

KEYWORDS

vigilance
unimodal
bimodal
sustained attention
audiovisual

INTRODUCTION

The Vigilance Decrement

Vigilance is the ability to sustain goal-directed attention over long periods (Davies & Parasuraman, 1982; Parasuraman, et al., 1987) and is an essential component of many tasks and jobs in the modern world. For example, vigilance is required to safely drive a vehicle, operate machinery, and monitor for security or safety hazards. In the military, vigilance is critical for many tasks, including sonar monitoring and satellite image surveillance. The maintenance of vigilance requires the application of a sufficient amount of cognitive resources to an often tedious task. Extensive evidence indicates that the ability to maintain vigilance declines in as little as ten minutes (Parasuraman & Mouloua, 1987; Warm et al., 2008). In military operations, any vigilance decrement may result in missed targets and/or slowed reaction times (Kamimori et al., 2005; Krueger, 1989; Mackworth, 1948, 1950), thus compromising warfighter safety and/or mission success. Despite multiple studies establishing the ubiquity of the vigilance decrement, it is still not fully understood. Specifically, while there has been research into unimodal vigilance (monitoring a single sensory modality for a signal), such as the auditory or visual psychomotor vigilance tasks, little research has compared the magnitude of the vigilance decrement in unimodal (auditory or visual) versus bimodal mixed vigilance tasks (both auditory and visual), which are common in modern real-world operations that require monitoring multiple sensory sources for a signal.

The Underload and Overload Theory of the Vigilance Decrement

Within an information processing framework, two competing models suggest vastly different outcomes of adding a second signal to unimodal tasks. One of these models, the overload theory, suggests that the effort required to maintain attention depletes cognitive resources faster than they can be replenished in a sustained setting. This leads to a decline in performance (Warm et al., 2008). Conversely, the underload theory (Manly et al., 1999) argues that lapses in sustained attention

Corresponding author: Sylvia Guillory, Naval Submarine Medical Research Laboratory, Submarine Base New London, Groton, CT 06349-5900.
Email: sylvia.b.guillory.ctr@mail.mil

DOI • 10.5709/acp-0347-7

result from under stimulation and not from strained capacity limits. According to this theory, under stimulation causes mind-wandering, which adversely impacts performance. Thus, increasing task difficulty by adding a second sensory modality would result in different outcomes depending on which of the above models is used. The underload theory predicts that the addition of a second sensory modality to monitor for signals might alleviate the monotony and boredom experienced when monitoring for a single signal, mitigating any decrements in performance that would be present in a unimodal sensory task. However, if the problem is a lack of resources, then adding a second sensory modality to monitor for signals may decrease resources faster, and to a greater extent, when compared to a unimodal task—an outcome consistent with overload theory.

Recent research assessing perceived mental workload and working memory load lend support to the overload theory of vigilance decline. Studies have found that when participants report on their subjective experience of a vigilance task, the task is characterized as effortful and demanding, contrary to what might be expected if participants were mind-wandering (Warm et al., 1996). Moreover, increasing working memory demands during a vigilance task amplifies the decrement in performance (Helton & Russell, 2011, 2013). Accordingly, monitoring for a signal in a second modality might then negatively impact performance in a similar manner.

Consecutive Multisource Mixed Presentation

Bimodal, mixed audiovisual vigilance tasks can be categorized into two broad types of paradigms. In one form of the bimodal task, stimuli are presented simultaneously. Most of the work investigating vigilance in audiovisual tasks has been to optimize performance via a cross-modal interaction where one modality affects perception in another, often facilitating signal detection (Craig et al., 1976; McGrath, 1965; Weinger & Smith, 1997). Another version of the bimodal sensory task presents stimuli from two modalities randomly and consecutively. In multimodal cueing, a promotion of modality dependence similar to simultaneous presentation shows performance improvements when spatial and temporal attention in one sensory modality involves a concurrent shift of attention in other modalities (Santangelo & Spence, 2008), but can also demonstrate impairments in performance in examples of the attentional blink (Dux & Marois, 2009) and inhibition of return (Klein, 2000). Performance improvements are shown to be consistent across ecologically valid studies (Gerdes et al., 2014) through real-world interface design (Baldwin et al., 2012). Situations where successive stimuli are not as obviously linked, a type of modality-switching, have been explored in recent research focusing on atypical processing in clinical populations such as attentional deficits or brain injury (Sandford & Turner, 1995; Tinius, 2003). These studies differ from cueing tasks as they do not promote the perception of a bimodal stimulus and are not expected to express the same multimodal enhancement. In modality switching, when a stimulus is preceded by the same modality, average accuracy and reaction times are often faster than when a stimulus is preceded by a stimulus from a different modality (Spence et al., 2001). The differences in response to the trials where there is a switch across modalities and trials that are repeated across modality is referred to as the switch cost. The overload theory might predict that with time-on-task cognitive fatigue, the switch-cost would worsen as the process of updating to a modality outside of the focus of attention would require resources from an already depleted system. Alternatively, if switching between modalities is considered more stimulating, it might reduce boredom and result in a steady switch-cost over time.

However, this line of modality-switching research has generally not compared the vigilance decrement between unimodal and bimodal mixed tasks. Although the ability to sustain and switch attention appears incompatible with one another, there is evidence that they rely on the same underlying brain systems (Duncan & Owen, 2000). In addition, potential differences in the vigilance decrement based on the number of senses being monitored is important from an applied perspective. Real-world vigilance tasks are rarely cleanly segregated into single sensory modalities. They are commonly multisensory efforts, with observers monitoring for both auditory and visual threats simultaneously. For example, in a sonar task, observers are listening for auditory signals to help identify a vessel while also searching for visual markers of vessel distance and speed. Similarly, security teams may be monitoring for auditory input in the form of radio calls (for example) while also watching a monitor for unsafe conditions or people. If vigilance declines faster in a bimodal task than a unimodal task, or vice versa, such a finding could have implications for these real-world tasks. A greater or faster decline in performance in one sensory modality over another could indicate that the tasks matching those requirements (e.g., baggage screening or radiology as unimodal tasks; driving, security monitoring, or operating sonars as bimodal/multimodal tasks) need more frequent interventions to mitigate the vigilance decrement and detect potential threats. Prior to designing a practical intervention to improve vigilance in one condition or another, an initial step is to determine if vigilance varies by unimodal versus bimodal task type.

Unimodal versus Bimodal Mixed Presentation

To date, few empirical studies have directly compared unimodal and mixed bimodal performance in a vigilance task, and the few that did investigate vigilance were published several decades ago (Baker et al., 1962; Binford & Loeb, 1963). Binford and Loeb (1963) compared the vigilance decrement in visual and mixed audiovisual vigilance conditions and found that even after controlling for event rate, mixed audiovisual vigilance had a greater time-on-task performance decrement and overall worse performance than visual vigilance alone, suggesting that mixed bimodal vigilance is more difficult and demanding than a unimodal vigilance task. In this study, the auditory and visual trials were presented separately, but were intermixed such that on any given trial, a participant would receive either an auditory or visual stimulus. There are, however, limitations to the conclusions one can draw from Binford and Loeb (1963). First, there was no condition where subjects only monitored for auditory signals. Thus, we do not have a baseline comparison point from which we can interpret the difference between unimodal conditions to mixed bimodal conditions. Second, the accuracy with which auditory and visual signals were detected was not equalized, introducing potential confounds. When combined with the lack of an auditory-only condition, it is impossible to determine if subjects prioritized attending the auditory signals, resulting in a decrease in visual signal detections in the mixed audiovisual condition, or if the auditory signals were significantly easier to detect. Regardless of which, or both, possibilities are true, a stronger test of mixed audiovisual vigilance would include an auditory-only condition and visual and auditory stimuli that are comparable in detection difficulty.

Baker et al. (1962) compared the vigilance decrement in auditory, visual, and mixed audiovisual tasks, which addressed one of the issues with the Binford and Loeb (1963) study. Baker et al. reported no interaction between the number of modalities monitored and the period of time on target detection performance. In other words, they found no evidence to suggest that the rate of vigilance decline is different in unimodal or mixed bimodal tasks. Though the results from Baker et al. (1962) appear clear, there are several reasons motivating a modern follow-up to their study. First, their results contradict those of Binford and Loeb (1963). Second, vigilance performance was analyzed in blocks of 30 min, while many vigilance studies show that the strongest decrement in performance occurs prior to the 30-minute mark (Helton et al., 1999; Teichner, 1974). Thus, it is possible there were substantial differences in the vigilance decrement in unimodal and mixed bimodal tasks, but it was masked by the window of analysis. Third, similar to Binford and Loeb (1963), performance across modalities was not controlled or set to a threshold. Finally, the mixed audiovisual data analyses were not examined by stimulus type. Thus, a decline in vigilance performance for one stimulus type may have been masked if there was an increase in performance for the other stimulus type. Therefore, the results of Baker et al. (1962) are not conclusive in determining if there are differences in performance between unimodal and mixed bimodal vigilance

Present Study

To help fill this knowledge gap, the present study compared a unimodal (auditory or visual) to a mixed bimodal (auditory and visual) task measuring the vigilance decrement and addressed the limitations in the Baker et al. (1962) and Binford and Loeb (1963) studies. To do this, (a) auditory and visual stimuli were developed that are equally discriminable and (b) the difficulty of discriminating signals and noise was increased relative to the high performances found in the previous studies, such that stimuli were accurately identified at a rate below ceiling to ensure that decrements in performance were detectable. A Go/NoGo continuous performance task (CPT) was selected as the vigilance paradigm to study the cognitive operations of inattention and speed of stimulus discrimination in conditions of auditory, visual, and mixed audiovisual vigilance over 28 min. We hypothesized that condition and time would interact such that vigilance performance would decrease faster and to a greater extent in the mixed bimodal audiovisual condition.

EXPERIMENT 1: PERCEPTUAL DISCRIMINABILITY STIMULI CALIBRATION

Stimuli Development

The stimuli used by Baker et al. (1962) had several limitations, namely, that the auditory and visual stimuli used were not equally discriminable (i.e., auditory stimuli were easier to detect than the visual stimuli) and that the auditory stimuli were too easy to detect. To equate for stimuli discriminability and to establish a performance threshold, we first identified a mixed audiovisual vigilance task that we could modify to suit our needs: the integrated visual and auditory continuous performance task (IVA-CPT, Sandford & Turner, 1995), which is used to identify individuals with attention deficits. The IVA-CPT allowed us to use, and modify, the phonemes /p/ and /b/ as auditory stimuli and the letters "b" and "p" as visual stimuli instead of the numbers 1 and 2 used by Baker et al. (1962). Numbers are easily discriminated, both auditorily and visually in normally functioning adults. In Baker et al.'s study, the numbers 1 and 2 commonly led to 100% accuracy rates, which could possibly prevent detection of a vigilance decrement.

The phonemes /p/ and /b/ and letters "b" and "p" share auditory and visual features that can be modified, affecting both categorical perception and increasing the difficulty of discriminability among the pairs. To make the auditory stimuli more difficult to discriminate, static noise can be played simultaneously, again changing the signal to noise ratio (SNR). To make the visual stimuli more difficult to discriminate, the tail length of the lowercase letters (ascender/descender) can be shortened, thus changing the SNR. We sought to identify the point at which the phonemes /b/ and /p/ and letters "b" and "p" could be discriminated at 80% accuracy when presented in only one modality at a time. This ensured that discrimination performance would be below ceiling and above chance, allowing us to detect a potential decrease in vigilance.

Method

PARTICIPANTS

Anonymous paid participants (N = 120) were recruited from Amazon Mechanical Turk (Mturk). The recruitment notification informed participants that computer speakers would be required to complete a 15 min task for which they would be compensated $5. Each participant completed one randomly assigned condition, either auditory or visual. No demographic or any other information was collected.

QUALITY CONTROL

To account for the lack of experimental control, participants' data were excluded from analyses if their overall accuracy, that is, correct target classification, was below 60%, as this would indicate chance level of performance, likely due to inattention or lack of effort. Our final sample size consisted of 45 participants in the auditory condition and 40 participants in the visual condition.

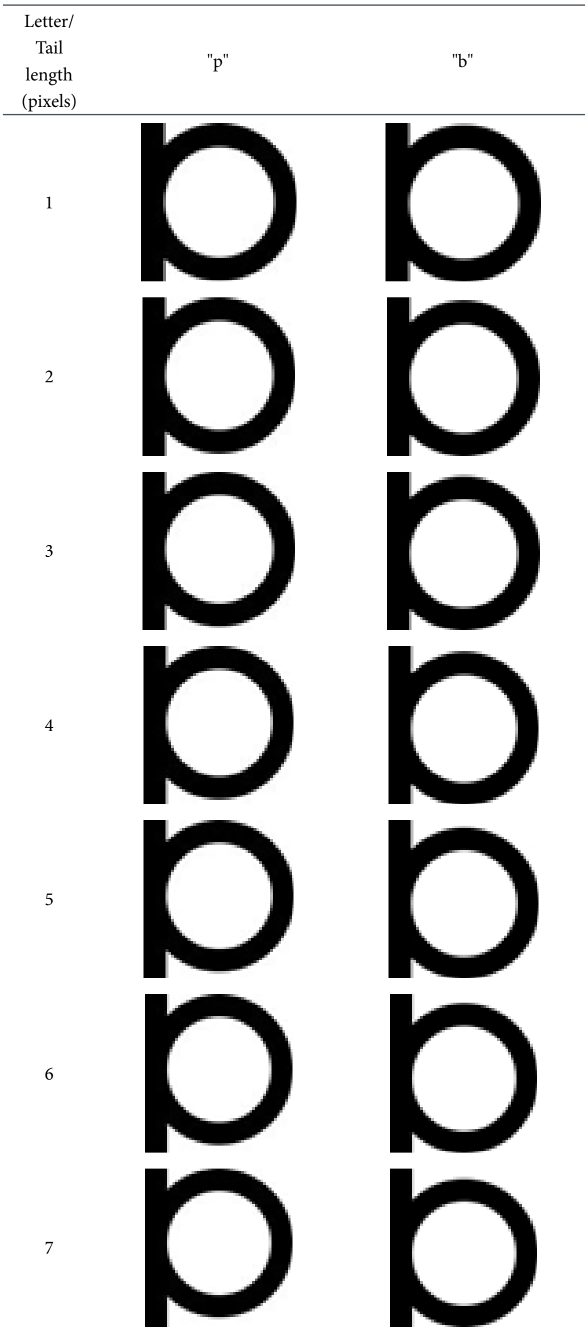

FIGURE 1.

"p" and "b" visual stimuli at each tail length used in Experiment 1 to establish (a) equally discriminable stimuli across modalities aand (b) a performance threshold of 80% (see Figure 2 for accuracy rates at each level). The stimulus at the 5 px. Tail length was used in Experiment 2

STIMULI

All stimuli were presented, and responses recorded, with Inquisit Lab 6 (Millisecond® software). Each participant was presented with either 140 /p/ phonemes or "p" letters and 140 /b/ phonemes or "b" letters (depending on the modality), in random order, to identify the point at which the SNRs of auditory and visual stimuli could be discriminated at 80 percent. Auditory stimuli were audio clips of the phonemes /b/ and /p/. The SNR in the auditory condition was varied by masking the presentation of the phonemes with Gaussian noise. Twenty stimuli were presented with each of the following SNRs measured in decibels (dB): −7, −10, −13, −16, −19, −22, and −25. The visual stimuli were presented on a white square, outlined in grey, on a white background. The square was 2.75 in., as measured on a 15.5 in. LCD Dell computer monitor, filled with white (RGB: 255, 255, 255) and outlined in grey (RGB: 127,127,127). The visual stimuli were lowercase letters "p" and "b" in black (RGB: 0,0,0) and measuring 7⁄16 × 3⁄8 in.. Both the square and the letters were centered in the display. The text color was selected to maximize stimulus visibility and minimize image after-effects. The SNR was varied by manipulating both the length of the descender for the letter "p" and the ascender for the letter "b" on the portions above or below the circle of the letter (see Figure 1). Twenty stimuli were presented at each of the following tail lengths, measured in pixels on a 1024 × 768 display: 1, 2, 3, 4, 5, 6, and 7. Auditory stimuli were presented for 500 ms, and visual stimuli were presented for 167 ms (as in the original IVA-CPT).

EXPERIMENTAL PARADIGM

On each trial, participants were presented with a stimulus that required a response in order to continue to the next trial. In the auditory condition, the stimulus was either a /b/ or a /p/phoneme, and in the visual condition, the stimulus was either the letter "b" or "p." The stimulus presented was from one of the seven SNR categories. There were 140 trials in total.

PROCEDURES

Participants were tasked with identifying which letter was presented by either pressing the "Z" (for /b/ or "b") or "M" (for /p/ or "p") keys on the computer keyboard. The button mapping for these responses was included on the display to reduce errors. There was no time limit for responses. The participant's response initiated the start of the following trial. The computer equipment varied across participants, where the mean screen height was 8.86 ± 2.04 in. Over fifty percent (54%) of participants had a screen size that fit into one of three size categories (31% for 7.4 in., 14% for 6.72 in., and 10% for 11.3 in.).

STATISTICAL ANALYSIS

Mean accuracy was computed at each SNR level per participant. The data were analyzed using a regression model to determine the relationship between accuracy and the SNR of each modality separately and deriving a sigmoid equation to predict the appropriate SNR that would produce an 80% discriminability. The SNR was then approximated to the nearest tested level for each modality using Euclidean distance.

Results and Discussion

The results showed a decrease in discriminability as the SNR decreased (see Figure 2). In both sensory modalities, there was a significant decrease in discrimination accuracy as the dB level of the

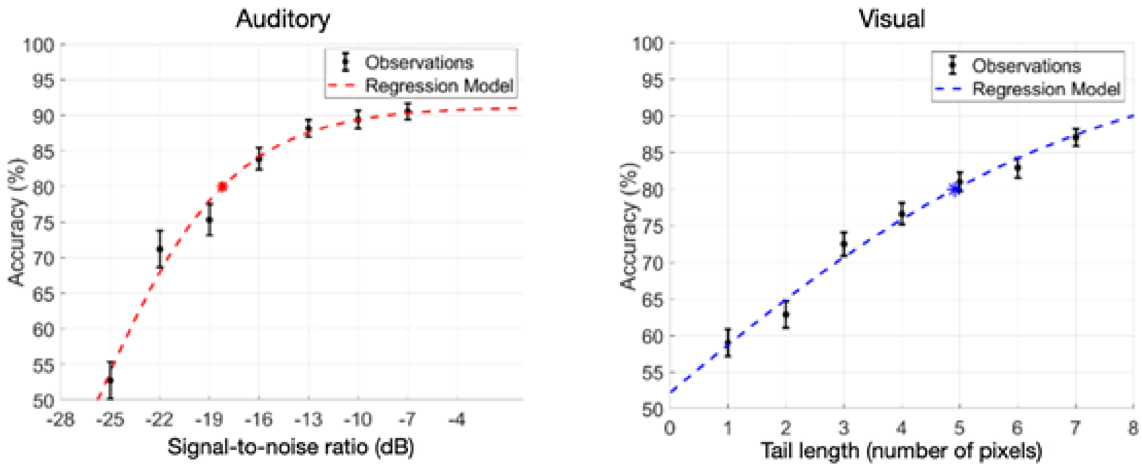

FIGURE 2.

Discrimination accuracy as a function of Gaussian noise (red: auditory stimuli) and tail length (blue: visual stimuli) testing levels. The dashed curve represents the fit of the regression model, with the asterisks being the points at which the discrimination accuracy equals the 80% threshold of performance. Error bars are ± 1 *SE*.

Gaussian noise (model fit auditory: $F[571, 574] = 4120$, $p < .001$, $R^2 = 0.35$) and as the ascender/descender length (model fit visual: $F[753, 756] = 5640$, $p < 0.001$, R2 = 0.27) decreased. With Gaussian noise of -18.2 dB and a tail length of 4.73 px, the auditory and visual stimuli were both accurately identified at 80%. As a consequence, Gaussian noise at a level of −16 dB and a tail length of 5 px were the stimuli parameters selected for use in the vigilance task as these levels were the nearest in Euclidean distance from the 80% values.

EXPERIMENT 2: UNIMODAL AND MULTIMODAL VIGILANCE TASK

Vigilance Task

Experiment 1 established the stimulus parameters that equated the auditory and visual stimulus discriminability and a performance threshold. To compare behavioral performance of a single sensory vigilance task to that for a mixed bimodal sensory task, Experiment 2 examined the vigilance decrement and decline in performance over time using performance measures of accuracy (correct target identification), average reaction time (latency to respond to the target), false alarms, and sensitivity measure A'.

Method

PARTICIPANTS

Three hundred and forty-five anonymous paid participants were recruited using Amazon Mturk. As with Experiment 1, the recruitment notification informed participants that they would be required to use their computer speakers to complete a 30 min task for which they would be compensated $5. Each participant completed one randomly assigned condition. No demographic or any other information were collected.

STIMULI

The auditory stimulus (the phonemes /b/ or /p/) was presented as recorded speech tracks at approximately 65 dB (conversational level of noise) via the computer speakers for 500 ms. As with Experiment 1, a white square outlined in grey and centered at the vertical and horizontal meridian was constantly displayed during the task in all three conditions (auditory, visual, and audiovisual). The background color of the screen was white. The visual stimuli were presented to the participants for 167 ms and centered inside the outlined square (see Figure 3, Panel B). The computing equipment used by participants varied where the mean screen height was 9.85 ± 5.04 in. Over fifty percent (52%) of participants had a screen size that fit into one of three size categories (27% at 7.5 in., 13% for both 6.89 in., and 10.62 in.). The dimensions of the stimuli in Figure 3, Panel B features a screen height of 7.5 in.

EXPERIMENTAL PARADIGM

Participants completed a modified version of the IVA-CPT (Sanford & Turner, 1995) on one of the three vigilance conditions: auditory, visual, or audiovisual, using the Inquisit Lab 6 software. In each trial of the task, participants were presented with a stimulus. Participants responded if the stimulus was a target and withheld a response if the stimulus was a distractor. The trial duration was 2000 ms and the next trial would start immediately following the previous trial. The target (Go) to distractor (NoGo) ratio was 1:5.25 (Sanford & Turner, 1995). The design of requiring participants to withhold a response to a frequent stimulus was to increase the external validity of the task. In many situations, responses are required only for rare target events, such as with sonar operation where frequent marine biologics are task-irrelevant and ignored. The auditory unimodal condition presented only auditory stimuli, the visual unimodal condition presented only visual stimuli, and the mixed audiovisual bimodal condition presented a mix of auditory and visual stimuli chosen randomly without replacement from a list of half auditory and half visual distractors and targets (see Figure 3, Panel A).

PROCEDURE

In each condition, the same instructions were displayed on the screen until the participant indicated understanding with a button press. The instructions stated, ”You will be using the spacebar to make your responses. You will hear or see the letters ”b” and ”p.” If you hear or see the letter ”b,”” press the spacebar. If you hear or see the letter ”p,” do not respond.” These instructions were reversed if the participant was assigned to the ”p” target condition (50%). Participants were given sample pictures and audio clips of the ”b” and ”p” letters and/or phonemes before completing a short practice assessment containing ten trials with feedback in the sensory modality of the task in which they were assigned.

Participants were directed to perform the vigilance task using their computer monitor and speakers, though because data collection was performed online, it is acknowledged that participants may have used headphones. Participants were instructed to listen to the volume of the /p/ and /b/ phoneme stimuli and self-adjust their computer’s volume so that the auditory stimuli were at the loudness level of a quiet conversation. Participants were instructed to use the spacebar to make their responses with their dominant hand. Stimuli in all conditions were presented individually. The participants could respond to the stimulus at any point following the start of the trial (stimulus presentation) to 2000 ms after initial presentation.

Participants completed 1000 trials, which took approximately 26 minutes total. For analysis, the 1000 trials were then divided into four blocks of 250 trials each, but there was no break or any other indication of the change from block to block to the participant. The equal sized blocks were used to facilitate analyses.

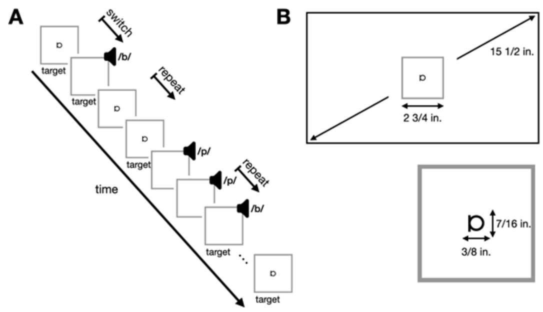

FIGURE 3.

Panel A: Task schematic of the bimodal mixed audiovisual task where the target signal is depicted as ”b” for the visual modality and /b/ for the auditory. The stimuli were presented in random order every 2000 ms with a stimulus duration of 167 ms for visual and 500 ms for auditory. For the analyses measuring the modality switch-cost, target-present trials were categorized as a switch trial when the trial preceding was of a different modality than the current trial (e.g., visual ”p” or ”b” → auditory /b/, auditory /b/ or /p/ → visual ”b”) or as a repeat when the preceding trial was from the same modality (e.g., auditory /p/ or /b/ → auditory /b/, visual ”b” or ”p” → visual ”b”). Panel B: Example display featuring the visual stimulus ”p” and the dimensions as measured using a 15.5 in. monitor.

STATISTICAL ANALYSIS

Differences in target identification rates (correct responses to the target), reaction times (latency to respond to the target), false alarms (incorrect responses to the distractor), and signal detection measure of sensitivity (A’) were determined using a 2 × 4 mixed analysis of variance (ANOVA) with the between-subjects factor of condition (unimodal and mixed bimodal) and the within-subject factor of block (1, 2, 3, and 4) for the auditory and visual conditions separately. Exploratory analyses were conducted to measure the sensory modality switch cost within the mixed bimodal task using a within-subjects design. A Greenhouse-Geisser correction was used when assumptions of sphericity were violated. An α criterion of 0.05 with two-tailed testing was used in all analyses, with the Bonferroni correction applied for post hoc tests.

QUALITY CONTROL

Due to the lack of experimental control, a two-step exclusion criterion was employed. First, a cutoff accuracy score of 60% or higher was applied. After visually inspecting the data, it was determined that participants committing excessive responding across all blocks and individuals with chance performance had performance at or below 60% total accuracy. Accordingly, as an informed judgement call by the authors, only participants with at least 60% accuracy were considered for further analysis. The second step in the exclusion process involved a cutoff for participants with two or fewer correct target identifications in their assigned modality (the audiovisual modality required two or fewer correct target identifications in each modality; this was necessary, as it was possible for a participant to make zero responses, but still have 84% percent overall accuracy due to the ratio of targets to distractors). The audiovisual condition in particular required over sampling; there was a high rate of noncompliance, such as participants responding to only one of the two stimulus types. There can be a large amount of variability in attention with online participation, and while this exclusion approach reduced some of the variability in attention, distractions and experimental setup are still a challenge. Thus, results should be viewed with these considerations in mind. This reduced our sample size in the auditory condition from 90 to 63, in the visual condition from 95 to 77, and in the audiovisual condition from 160 to 81.

Results

CORRECT TARGET IDENTIFICATION

A central question was whether performance in a mixed bimodal sensory vigilance task was worse than in a unimodal task. Figure 4 shows target accuracy (correct target identification) by block number for the auditory and visual stimuli, respectively.

A significant main effect of condition in both auditory, $F(0.79, 112.7) = 16.65$, $p < .001$, $\eta_p^2 = 0.105$, and visual, $F(0.749, 116.9) = 28.16$, $p < .001$, $\eta_p^2 = 0.153$, modalities revealed that performance overall was worse in the bimodal mixed task (auditory: 0.63±0.37; visual: 0.57±0.32) compared to the unimodal task (auditory: 0.84±0.25; visual: 0.79±0.24).

The main effect of block did reach statistical significance in the auditory condition, $F(2.37, 336.5) = 12.45$, $p < .001$, $\eta_p^2 = 0.081$. Post hoc tests of the auditory condition found a statistically significant accuracy decline between Block 1 (0.79±0.29) and Block 2 (0.74±0.32, $p = .03$), Block 1 and Block 3 (0.71±0.32, $p < .001$), Block 1 and Block 4 (0.71±0.30, $p < .001$), and Block 2 and Block 3 ($p = .02$), indicating that across conditions, a vigilance decrement was observed after approximately 6.5 minutes on task and was sustained until the task ended. No other pairwise comparisons between blocks reached statistical significance ($p > .05$).

The main effect of block was not significant in the visual condition, $F(2.225, 350.5) = 1.85$, $p = .15$, $\eta_p^2 = 0.012$. There was a significant interaction within the auditory modality between block and condition (unimodal, bimodal), $F(2.37, 336.5) = 4.63$, $p = .007$, $\eta_p^2 = 0.032$. The difference in performance between the unimodal and bimodal auditory conditions continued to diverge across time, suggesting that the occurrence of the vigilance decrement might vary as a function of the number of modalities monitored. The interaction effect in the visual modality was not statistically significant, $F(2.25, 350.5) = 1.42$, $p = .24$, $\eta_p^2 = 0.009$.

To dissect the statistically significant auditory results further, the condition × block interaction was decomposed into separate repeated-measures ANOVAs for the auditory unimodal and bimodal conditions independently. The analysis across blocks showed a significant effect of block in the bimodal auditory group, $F(2.36, 188.42) = 14.90$, $p < .001$, $\eta_p^2 = 0.157$, such that performance was comparable between Block 1 (0.70±0.33) and Block 2 (0.65±0.38, $p = .078$), but then declined during Block 3 (0.59±0.38, $p < .001$) and Block 4 (0.58±0.38, $p < .001$); the vigilance decrement manifested after roughly 13 minutes on task. The performance decline was also significant between Block 2 and both Block 3 ($p = .004$) and Block 4 ($p = .017$), with no further declines between Block 3 and Block 4 ($p = 1.00$). Differences in performance across blocks in the auditory unimodal group were not statistically significant, $F(2.34, 150.97) = 1.53$, $p = .22$, $\eta_p^2 = 0.024$, suggesting that the main effect of block was driven in part by the bimodal condition.

LATENCY TO IDENTIFY TARGETS

Reaction time as a correlate of information processing speed provides another approach to measure changes in performance that may not be captured with accuracy measures alone. Here, a delayed or longer reaction time across blocks would be associated with time-on-task fatigue. Latency measures were confined to an upper limit of 2000 ms to control for a reaction time distribution with a long tail.

Within the auditory modality, there were no statistically significant main effects of condition, $F(0.79, 97.6) = 1.87$, $p = .17$, $\eta_p^2 = 0.015$; unimodal: 796.6±143.6 ms, bimodal: 833.5±222.5 ms), block, $F(2.38, 292.9) = 0.70$, $p = .50$, $\eta_p^2 = 0.006$; nor an interaction, $F(2.38,292.9) = 1.91$, $p = .13$, $\eta_p^2 = 0.015$ (see Figure 4, Panel C).

Within the visual modality (see Figure 4, Panel D), there was a statistically significant main effect of condition, $F(0.94, 130.2) = 14.3$, $p < .001$, $\eta_p^2 = 0.094$. Reaction times in the unimodal task (622.93±140.07 ms) were faster than the visual component of the bimodal task (706.19±199.38 ms). The main effect of block was not statistically significant, $F(2.83, 390.1) = 0.65$, $p = .57$, $\eta_p^2 = 0.005$, such that reaction times overall were similar over time. There was a significant interaction between block and condition on mean latency, $F(2.83, 390.1) = 3.23$, $p = .025$, $\eta_p^2 = 0.023$, where the greatest differences in reaction times between the unimodal and bimodal visual conditions occurred during the initial block and then began to converge.

Interrogating the significant condition × block effect in the visual modality, the interaction was broken down into repeated-measures ANOVAs by group, where in both the unimodal, $F(2.63, 200.00) = 2.02$, $p = .12$, $\eta_p^2 = 0.026$, and bimodal, $F(3, 240) = 2.07$, $p = 0.11$, $\eta_p^2 = 0.025$, tasks, there were no statistically significant differences across blocks.

FALSE ALARMS

False alarm rates were examined to determine wherther there was an increase in responding to the distractors in unimodal versus the mixed bimodal condition.

Within the auditory modality, the bimodal condition elicited significantly more false alarms than the unimodal condition (mean difference: 0.09 ± 0.02, $p < .001$). There was also a significant change in false alarm rates across blocks, where Block 1 had significantly higher false alarms compared to Blocks 2 through 4 ($ps < .001$). All other block comparisons did not reach statistical significance ($ps > .05$).

In the visual condition, there were significantly more false alarm errors in the bimodal condition compared to the unimodal condition (mean difference: 0.06 ± 0.02, $p = .006$). There was no main effect of block ($p > .05$) or interaction effect ($p = .37$).

SENSITIVITY INDEX

Participant sensitivity for discriminating the target from the distractor stimuli was higher in the unimodal condition compared to the mixed bimodal condition (visual: mean difference of 0.18 ± 0.03, $p < .001$; auditory: mean difference of 0.18 ± 0.04, $p < .001$). There was no significant effect of block for either visual or auditory stimuli ($ps > .05$).

AUDIOVISUAL EXPLORATORY ANALYSIS: MODALITY SWITCH COST

Switching from one modality to another often results in a performance cost such that worse performance is observed when there is a change in stimulus modality between successive trials compared to when the stimulus modality remains constant (Driver & Spence, 1998). If this were the case, then the worse performance in the mixed bimodal condition could be specific to the trials where participants had to switch between sensory modalities rather than the overall cognitive demand of monitoring two modalities that would result in similar performance across the trial types. To examine whether any potential modality switch cost in the bimodal mixed audiovisual task contributed to the vigilance decrement, accuracy was calculated for target stimuli (Go) that followed either a trial of the same sensory modality (repeat) or from a different sensory modality (switch), and was submitted to a repeated-measures ANOVA with block (1,2,3, and 4) and trial type (switch and repeat) as within-subject factors.

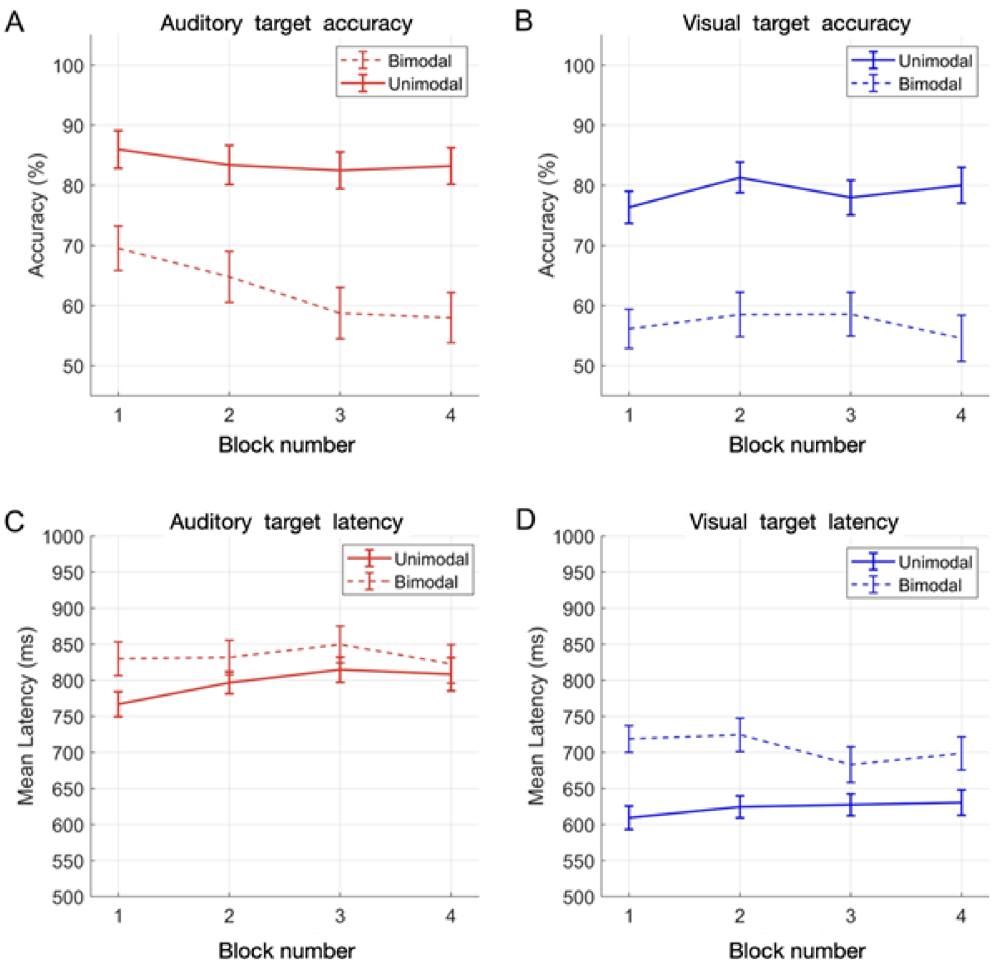

FIGURE 4.

Target Identification Rate for the auditory (Panel A) and visual (Panel B) modalities. Solid lines represent performance on the single sensory unimodal task and the dashed lines represent the mixed sensory bimodal task across the four blocks (150 trials each). Latencies to identify target are shown for auditory (Panel C) and visual (Panel D) conditions. Error bars are ± 1 *SE*.

The analysis detected no statistically significant difference in performance between trial type, $F(0.79, 126.9) = 0.18$, $p = .61$, $\eta_p^2 = 0.011$). Performance was similar between switch (0.59±0.31) and repeat trials (0.61±0.30). Across blocks, there were statistically significant differences in performance, $F(2.38, 380.8) = 6.67$, $p < .001$, $\eta_p^2 = 0.041$, with participants less accurate in later blocks reflecting similar trends as the main analyses (Block 2: 0.62±0.31 vs. Block 4: 0.56±0.33; $p = .02$). No other block (Block 1: 0.63±0.27; Block 3: 0.59±0.33) comparisons were statistically significant ($p > .05$). The trial type × block interaction did not reach statistical significance, $F(2.38, 380.8) = 0.14$, $p = .90$, $\eta_p^2 = 0.007$). The reaction time data yielded no statistically significant differences between trial type ($p = .40$) and block ($p = .16$), or a statistically significant interaction ($p = .69$).

GENERAL DISCUSSION

We investigated two questions: (a) is a mixed bimodal vigilance task involving abstract symbol recognition more difficult than a comparable unimodal vigilance task, and (b) does the mixed bimodal vigilance task result in a greater vigilance decrement than the unimodal vigilance task? The probing of these questions serves to better understand the limits and use of cognitive resources during sustained attention tasks and the interaction with attentional switching. To properly address these questions, we designed a task and selected stimuli that equally tested auditory, visual, and audiovisual vigilance while overcoming past research limitations and confounding factors. Our first step was identifying auditory and visual stimuli that were equally discriminable at a subceiling rate (~80%) because the task would be sufficiently difficult that a subject who experienced the vigilance decrement would exhibit an observable decrease in identification rate. We chose stimuli that were similarly defined in the auditory (/p/ and /b/ phonemes) and visual ("p" and "b" letters) domains to prohibit potential explanations of performance changes due to overtaxing the working memory with too many signal/noise representations. Equally discriminable stimuli would also control for the ease or difficulty of detecting one stimulus over another. Equally discriminable stimuli were identified in Experiment 1 by testing a range of difficulties for each stimulus and selecting for ~80% positive identification rate. Finally, we designed the task based on a preexisting vigilance task that controlled the proportion of auditory and visual targets. An equal proportion of each stimulus type ensures that the task did not bias subjects to attend to one modality more than another. Implementing these methodological changes helped to elucidate performance differences between unimodal and mixed bimodal tasks.

The results presented here support the hypothesis that monitoring for a signal in a mixed bimodal condition is more difficult than monitoring for a comparable signal in unimodal conditions, as shown by the significantly lower target identification rate of both auditory and visual targets in the bimodal mixed audiovisual condition relative to the unimodal conditions, regardless of block. The elevated false alarm errors during the mixed bimodal condition adds further support to the claim that monitoring for two signals is more difficult. Indeed, the sensitivity measures indicate that participants in the mixed bimodal condition had greater difficulty discriminating between the target and nontarget stimuli in comparison to the unimodal tasks. Responses from the correct target identification and the false alarms allude to an absence of a response bias that was neither too liberally marked by high false alarm and high correct target identification rates or too conservatively defined by low false alarm and low correct target identification rates.

The finding of a statistically significant interaction between condition and block within the auditory modality supports the second hypothesis that monitoring for a signal in a mixed bimodal task also results in a stronger vigilance decrement relative to monitoring in a unimodal condition, at least within the auditory modality. These data are consistent with the overload theory of the vigilance decrement that argues for worse performance in the mixed bimodal task (Helton & Warm, 2008) in contrast to the underload theory, which predicts worse performance in the unimodal condition, the less cognitively demanding task. With a limit on the amount of cognitive resources available, the distribution of these resources over multiple sensory modalities resulted in their faster utilization, impacting performance compared to the single sensory setting. The reaction time data reinforce these results to some degree. Reaction times were overall longer (i.e., slower) within the visual modality of the mixed audiovisual task compared to the single task, suggesting greater cognitive demands impacting processing speeds.

It is perhaps not unexpected that correctly identifying a signal using two modalities results in a decrease in overall performance. Dual-task literature, including distracted driving research, has shown that monitoring two modalities decreases performance (Gherri & Eimer, 2011; Levy & Pashler, 2008; McKnight & McKnight, 1993). However, such research often has confounding factors. For instance, in distracted driving literature, the performance metrics for the two modalities are often on different scales and the task requirements are different for the two modalities, resulting in an unequal number of events in each modality. Additionally, a subject may naturally prioritize the visual/driving condition because that is how they would normally drive in real life. Our task controls for these factors, reducing confounds in analyses and the interpretation of data. In equating the difficulty of the unimodal tasks, the current study removed differences in the performance in the single sensory task such that any differences between unimodal and mixed bimodal performance is attributable to the cognitive cost of monitoring for a signal in the two modalities.

The results revealed that the reaction time comparisons between the unimodal and mixed bimodal auditory conditions were comparable, whereas in the visual condition, participants responded slower to targets in the bimodal condition. The split results suggest that the reaction time measure may be more sensitive to modality differences than accuracy—a measure where both auditory and visual modalities experienced a decline. Research indicates that attention can be selective to a single modality at the expense of a secondary modality (Sandhu & Dyson, 2012; Spence et al., 2001). While the task equated the number of stimuli from each modality, and no directions were given to prioritize one modality over another, the similar reaction times in the auditory modality not observed in the visual comparison suggest that the auditory stimulus might have been the preferred modality dominating processing during the mixed bimodal task. Studies have reported an auditory bias during temporal tasks, including a processing dominance that delays visual responses (Dunifon et al., 2016; Robinson et al., 2018). Also, visual task demands were shown to influence auditory information processing (Ciaramitaro et al., 2017). The linguistic properties of the stimulus may have contributed to an auditory bias where it may be more habitual to process the phonemes (Aramaki et al., 2010). Here, the visual stimuli may not have been represented as letters but decomposed into the component parts, with a focus on the ascender/descender tail length affecting categorization, conceptual priming, and speed of discriminating. Another possible explanation is that the mixed bimodal task has fewer trials than the unimodal task by modality, resulting in less exposure to accrue the same level of familiarity and expertise at discriminating the stimuli. Due to the difficulty in differentiating these stimuli, one might expect that with increased stimulus exposures, performance would increase with practice. Instead, in the mixed bimodal condition, there was no change in reaction time across blocks and when taken together with the accuracy results, there was a decrease in performance over time. This suggests that either a practice effect is not occurring, or that the vigilance decrement is so strong that its negative effect on performance overcomes any benefits of practice. Lastly, the sudden onset of the auditory stimulus could have resulted in a startle effect that oriented attention to that modality (Valls-Sole et al., 1995), which could explain the lack of a reaction time difference in the auditory condition and a potential processing bias. Further research is needed to investigate and characterize the extent of these modality differences.

When switching between tasks or modalities, as the mixed bimodal condition necessitates, inter-reference can occur, as there is competition for resources with negative consequences on performance (Nieznanski et al., 2015). An overload account of a switch cost would predict that over time, the switch cost would increase as more cognitive resources have been depleted without the opportunity of replenishing those resources with a break. The underload theory might argue that the switching between modalities alleviates the monotony of the task that might otherwise be cause for participants to disengage. Here, performance on switch trials might have reoriented participants that recovered any loss in attention and would have resulted in higher accuracy. The results are in broad agreement that switching between modalities does not significantly contribute to the performance difference or the vigilance decrement. The relatively long interstimulus interval of 2000 ms may have diminished any within-modality facilitation on repeat trials associated with a switch cost (Cuppini et al., 2020). With the overall worse performance in the mixed bimodal condition compared to the unimodal tasks, the lack of a switch cost (i.e., no decrease in target identification accuracy or reaction time following a switch between modalitites in the mixed bimodal condition, Blotenberg et al., 2018; Sutton et al., 1961) provided a clearer understanding of the extent to which the control process associated with coordinating between modalities contributed to the performance decline. These results suggest that performance declines observed in the mixed bimodal task were driven in part by diminished cognitive resources from monitoring two sensory modalities unrelated to switching attention.

The interaction between task condition and block on target detection accuracy in the auditory modality is a novel finding. Little research has investigated the effects of monitoring two modalities on the vigilance decrement. Though Binford & Loeb (1963) and Baker et al. (1962) did so, their work had several confounds. The lack of an overall decrease in reaction time suggests that the decrease in accuracy was not due to a speed-accuracy tradeoff, further supporting our interpretation. This finding suggests that participants struggled when attempting to attend to two or more modalities simultaneously, particularly when they attended to both over long periods of time, which aligns with an overload account of performance. This has broad implications in applied settings, such as sonar monitoring, security surveillance, and air traffic control, where there is a requirement to maintain attention on multiple modalities over long periods. Performance in these bimodal tasks may be initially lower and decrease faster than if they were performed under unimodal conditions.

However, these results come with limitations. First, though we found the predicted interaction between condition and block in the mixed audiovisual condition, we found no vigilance decrements in the single modality conditions. This suggests that our task may not be optimized for detecting any potential decrement in such conditions and/or that the task duration was not long enough to induce any decrement. Studies have found an enhancement in the vigilance decrement as a function of signal frequency. Detection rate decreases as the signal frequency decreases (Craig et al., 1987). Many studies have used a critical signal probability between 3% and 30% (Baddeley & Colquhoun, 1969; Nuechterlein et al., 1983). Here, the critical signal probability was 19%, within the range of observing a vigilance decrement, though one that might be less pronounced. It is also possible that the subjects in the unimodal conditions did experience a vigilance decrement, but it simultaneously occurred with a learning effect, where they were becoming more adept at discriminating between stimuli. If this were true, we would expect that if we first trained subjects to reach a plateau in stimulus discrimination, we could then detect a vigilance decrement as the task proceeded. Unfortunately, the high dropout and noncompliance rates with online data collection make such a hypothesis difficult to evaluate. Second, the online nature of our recruitment and data collection leads to potential statistical noise in our experiment. With no demographic information or recruitment filters, subjects with attention or sensory deficits or whose native language is not English (and, therefore, who may be less familiar with the English letters and phonemes) may have participated in our study. With no controls over the equipment used, participants may have had different quality of stimuli presentations (e.g., some speakers filter out static noise built into our stimuli), leading to performance fluctuations for reasons other than task difficulty. We also detected a high rate of noncompliance, with numerous subjects' performances at or below chance level, suggesting that subjects either gave responses randomly or started the task then failed to attend to it. Environment and physical setup factors can influence our response measures and a future lab study can address these concerns. Nonetheless, while these factors may have reduced the reliability or effect size of our findings, the random assignment of a large sample to the three conditions using a diverse community of subjects through Amazon Mturk (Huff & Tingley, 2015), the control over stimulus detection difficulty, and the control over event rate all support the generalizability of these results despite a lack of control over the testing environment.

We investigated the potential performance decrements associated with monitoring for a critical signal across two modalities, auditory and visual, simultaneously for a signal relative to monitoring for only a visual signal or only an auditory signal. Our results indicate that monitoring two modalities simultaneously is more difficult and decreases performance more over time relative to unimodal tasks. We conclude that those who must perform tasks requiring their attention to be devoted to two modalities simultaneously may struggle to perform at a high level, particularly when performing the task over prolonged periods. Methods to decrease the strength and onset speed of the vigilance decrement, such as transcranial direct current stimulation to increase attentional resources (McIntire et al., 2017) or adding additional feedback on task performance (Schwark et al., 2012), should be considered to improve performance in these attentionally demanding tasks.

ACKNOWLEDGEMENTS

The views expressed in this report reflect the results of research conducted by the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

The study protocol was approved by the Naval Submarine Medical Research Laboratory Institutional Review Board in compliance with all applicable Federal regulations governing the protection of human subjects, protocol number NSMRL.2019.00013.

The corresponding author is a military service member or federal/contracted employee of the U.S. Government. This work was prepared as part of official duties. Title 17, U.S.C., §105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C., §101 defines a U.S. Government work as work prepared by a military Service member or employee of the U.S. Government as part of that person's official duties.

This research was supported in part by the Office of Naval Research, work unit number F1902 (CP).

All authors declare that they have no conflict of interest.

DATA AVAILABILITY

The data generated and analyzed during the current study may contain proprietary information or restricted access information. Data may be available upon written request to Naval Medical Research Center (NMRC) Legal subject to approval.

REFERENCES

Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., & Besson, M. (2010). Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *Journal of Cognitive Neuroscience, 22*, 2555–2569. doi: 10.1162/jocn.2009.21398

Baddeley, A. D., & Colquhoun, W. P. (1969). Signal probability and vigilance: a reappraisal of the 'signal-rate' effect. *British Journal of Psychology, 60*, 169–178. doi: 10.1111/j.2044-8295.1969.tb01189.x

Baker, R. A., Ware, J. R., & Sipowicz, R. R. (1962). Vigilance: A comparison in auditory, visual, and combined audio-visual tasks. *Canadian Journal of Psychology, 16*, 192–198.

Baldwin, C. L., Spence, C., Bliss, J. P., Brill, J. C., Wogalter, M. S., Mayhorn, C. B., & Ferris, T. K. (2012). Multimodal cueing: The relative benefits of the auditory, visual, and tactile channels in complex environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*, 1431–1435. doi: 10.1177/1071181312561404

Binford, J. R., & Loeb, M. (1963). Monitoring readily detected auditory signals and detection of obscure visual signals. *Perceptual and Motor Skills, 17*, 735–746. doi: 10.2466/pms.1963.17.3.735

Blotenberg, I., Stephan, D., & Koch, I. (2018). Consistent shifts of stimulus modality induce chunking in sequence learning. *Advances in Cognitive Psychology, 14*, 101–111. doi:10.5709/acp-0242-8

Ciaramitaro, V. M., Chow, H. M., & Eglington, L. G. (2017). Cross-modal attention influences auditory contrast sensitivity: Decreasing visual load improves auditory thresholds for amplitude- and frequency-modulated sounds. *Journal of Vision, 17*, 20. doi: 10.1167/17.3.20

Craig, A., Colquhoun, W. P., & Corcoran, D. W. J. (1976). Combining evidence presented simultaneously to the eye and the ear: A comparison of some predictive models. *Perception & Psychophysics, 19*, 473–484.

Craig, A., Davies, D. R., & Matthews, G. (1987). Diurnal variation, task characteristics, and vigilance performance. *Human Factors, 29*, 675–684. doi:10.1177/001872088702900607

Cuppini, C., Ursino, M., Magosso, E., Crosse, M. J., Foxe, J. J., & Molholm, S. (2020). Cross-sensory inhibition or unisensory facilitation: A potential neural architecture of modality switch effects. *Journal of Mathematical Psychology, 99*, 102438. doi: 10.1016/j.jmp.2020.102438

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

Driver, J., & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology, 8*, 245–253. doi: 10.1016/s0959-4388(98)80147-5

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475–483. doi: 10.1016/s0166-2236(00)01633-7

Dunifon, C. M., Rivera, S., & Robinson, C. W. (2016). Auditory stimuli automatically grab attention: Evidence from eye tracking and attentional manipulations. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1947–1958. doi: 10.1037/xhp0000276

Dux, P. E., & Marois, R. (2009). The attentional blink: a review of data and theory. *Attention, Perception, & Psychophysics, 71*, 1683–1700. doi: 10.3758/APP.71.8.1683

Gerdes, A. B., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology, 5*, 1351. doi: 10.3389/fpsyg.2014.01351

Gherri, E., & Eimer, M. (2011). Active listening impairs visual perception and selectivity: An ERP study of auditory dual-task costs on visual attention. *Journal of Cognitive Neuroscience, 23*, 832–844. doi: 10.1162/jocn.2010.21468

Helton, W. S., Dember, W. N., Warm, J. S., & Matthews, G. (1999). Optimism, pessimism, and false failure feedback: Effects on vigilance performance. *Current Psychology, 18*, 311–325. doi: 10.1007/s12144-999-1006-2

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. Exp Brain Res, 212(3), 429–437. doi:10.1007/s00221-011-2749-1

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

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

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics, 2*, 2053168015604648. doi: 10.1177/2053168015604648

Kamimori, G. H., Johnson, D., Thorne, D., & Belenky, G. (2005). Multiple caffeine doses maintain vigilance during early morning operations. *Aviation, Space, and Environmental Medicine, 76*, 1046–1050.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Science, 4*, 138–147. doi: 10.1016/s1364-6613(00)01452-2

Krueger, G. P. (1989). Sustained work, fatigue, sleep loss and performance: A review of the issues. *Work & Stress, 3*, 129–141. doi: 10.1080/02678378908256939

Levy, J., & Pashler, H. (2008). Task prioritisation in multitasking during driving: Opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing. *Applied Cognitive Psychology, 22*, 507–525. doi: 10.1002/acp.1378

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

Mackworth, N. H. (1950). *Researches on the measurement of human performance. (Med. Res. Council, Special Rep. Ser. No. 268.).* His Majesty's Stationery Office.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia, 37*, 661–670. doi: 10.1016/s0028-3932(98)00127-4

McGrath, J. J. (1965). Performance sharing in an audio-visual vigilance task. *Human Factors, 7*, 141–153. doi: 10.1177/001872086500700206

McIntire, L. K., McKinley, R. A., Nelson, J. M., & Goodyear, C. (2017). Transcranial direct current stimulation versus caffeine as a fatigue countermeasure. *Brain Stimulation, 10,* 1070–1078. doi:10.1016/j.brs.2017.08.005

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention, 25*, 259–265. doi: 10.1016/0001-4575(93)90020-W

Nieznanski, M., Obidzinski, M., Zyskowska, E., & Niedzialkowska, D. (2015). Executive resources and item-context binding: Exploring the influence of concurrent inhibition, updating, and shifting tasks on context memory. *Advances in Cognitive Psychology, 11*, 106–117. doi: 10.5709/acp-0176-9

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

Parasuraman, R., & Mouloua, M. (1987). Interaction of signal discriminability and task type in vigilance decrement. *Perception & Psychophysics, 41*, 17–22. doi:10.3758/bf03208208

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

Robinson, C. W., Moore, R. L., Jr., & Crook, T. A. (2018). Bimodal presentation speeds up auditory processing and slows down visual processing. *Frontiers in Psychology, 9*, 2454. doi: 10.3389/fpsyg.2018.02454

Sandford, J. A., & Turner, A. (1995). *Intermediate visual and auditory continuous performance test interpretation manual*. Braintrain.

Sandhu, R., & Dyson, B. J. (2012). Re-evaluating visual and auditory dominance through modality switching costs and congruency analyses. *Acta Psychologica, 140,* 111–118. doi:10.1016/j.actpsy.2012.04.003

Santangelo, V., & Spence, C. (2008). Is the exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Consciousness and Cognition, 17*, 989–1015. doi: 10.1016/j.concog.2008.02.006

Schwark, J., Sandry, J., Macdonald, J., & Dolgov, I. (2012). False feedback increases detection of low-prevalence targets in visual search. *Attention, Perception, & Psychophysics, 74*, 1583–1589. doi: 10.3758/s13414-012-0354-4

Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics, 63*, 330–336. doi: 10.3758/bf03194473

Sutton, S., Hakerem, G., Zubin, J., & Portnoy, M. (1961). The effect of shift of sensory modality on serial reaction-time: A comparison of schizophrenics and normals. *The American Journal of Psychology, 74*, 224–232. doi: 1419407

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors, 16*, 339–353. doi: 10.1177/001872087401600402

Tinius, T. P. (2003). The Integrated Visual and Auditory Continuous Performance Test as a neuropsychological measure. *Archiches of Clinical Neuropsychology, 18*, 439–454. doi: 10.1093/arclin/18.5.439

Valls-Sole, J., Sole, A., Valldeoriola, F., Munoz, E., Gonzalez, L. E., & Tolosa, E. S. (1995). Reaction time and acoustic startle in normal human subjects. *Neuroscience Letters, 195*, 97–100. doi: 10.1016/0304-3940(94)11790-p

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

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50*, 433–441. doi: 10.1518/001872008X312152

Weinger, M. B., & Smith, N. T. (1997). Vigilance, alarms, and integrated monitoring systems. In: J. Ehrenwerth, J. B. Eisenkraft, & J. M. Berry (Eds.), Anesthesia equipment: Principles and applications (pp. 350–384). Mosby Year Book.

RECEIVED 04.05.2021 | ACCEPTED 12.02.2022