Naval Submarine Medical Research Laboratory

NSMRL/F1902/TR--2022-1408 **November 28, 2022**

Focusing and Switching Attention During a Mixed Modality Auditory-Visual Vigilance Task

by:

Chad Peltier
Sylvia Guillory
LT Jennifer F. Louie
Justin D. Handy

Approved and Released by:
M. H. Jamerson, CAPT, MSC, USN
Commanding Officer
NAVSUBMEDRSCHLAB

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*: 28-11-2022

2. REPORT TYPE: Technical Report /Memorandum

3. DATES COVERED *(From - To)*: 09/2019 - 06/2022

4. TITLE AND SUBTITLE
Focusing and Switching Attention During a Mixed Modality Auditory-Visual Vigilance Task

5a. CONTRACT NUMBER: N0001420WX00127

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER: '0601152N

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER: F1902

6. AUTHOR(S)
Chad Peltier, Sylvia Guillory, Jennifer F. Louie, Justin D. Handy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Submarine Medical Research Laboratory
Box 900
Groton, CT 06349-5900

8. PERFORMING ORGANIZATION REPORT NUMBER: NSMRL/F1902/TR--2022-1408

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Navy Medicine and Development Enterprise
In-House Laboratory Independent Research Program
503 Robert Grant Avenue
Silver Spring, MD 20910

10. SPONSOR/MONITOR'S ACRONYM(S): ILIR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
There are many scenarios where it is advantageous to prioritize and selectively attend to a single sensory source as well as flexibly engage, disengage, and re-engage with multiple sensory modalities, such as when operating a vehicle. The influences of simultaneous stimuli from different sources and from multiple sensory modalities (e.g., visual and auditory) have been shown to divide attention, which degrades performance. In addition, time-on-task cognitive fatigue can compromise performance and the ability to ignore task-irrelevant information. The aims of this study were to examine whether vigilance-related declines in performance are attributable to supramodal resources that are shared across multiple sensory modalities and to characterize the main source of false alarm errors over time. Stimuli were presented visually and audibly, and participants were instructed to respond to either the visual targets only or the auditory targets only (i.e., focusing on a single modality) or both visual and auditory targets (i.e., alternating attention between modalities). The auditory and visual stimuli involved pairing letters and sounds (phoneme /b/ with the grapheme "b" and phoneme /p/ with the grapheme "p"), resulting in non-targets that could be either modality-relevant non-targets (i.e., the same modality as the target and not the target signal), modality-irrelevant non-targets (i.e., a different modality of the target and not paired with the critical target), or modality-irrelevant target-congruent (i.e., a different modality, but sharing a letter-sound correspondence with the critical signal). While significant effects of cognitive fatigue over time were not detected, there was reduced accuracy when switching attention between auditory and visual targets compared to focusing attention on a target from a single modality. Also, when focusing attention on a single modality, false alarms to modality-relevant non-target items were substantially higher than non-targets that were modality-irrelevant. These results highlight the vulnerabilities of selective attention to task-irrelevant information that are of the same modality as target signals.

15. SUBJECT TERMS
vigilance, switch-task, sustained attention, task-irrelevant, cognitive fatigue, auditory-visual attention

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	SAR	21	NAVSUBMEDRSCHLAB Commanding Officer

19b. TELEPHONE NUMBER *(Include area code)*: 860-694-3263

Reset

[THIS PAGE INTENTIONALLY LEFT BLANK]

Focusing and Switching Attention During a Mixed Modality Auditory-Visual Vigilance Task

By:

Chad Peltier[1]
Sylvia Guillory[1,2]
LT Jennifer F. Louie[2]
Justin H. Handy[2]

[1]Leidos, Inc.
[2]Naval Submarine Medical Research Laboratory

Approved and Released by:

JAMERSON.MATTHEW.HUNTER.1269308288 Digitally signed by JAMERSON.MATTHEW.HUNTER.1269308288 Date: 2022.12.05 11:17:48 -05'00'

M. H. Jamerson, CAPT, MSC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory
Submarine Base New London Box 900
Groton, CT 06349-5900

ADMINISTRATIVE INFORMATION

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, not the U.S. Government. This work was supported by funding work unit number F1902 and funding source In-house Laboratory Independent Research (ILIR) Program. 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.2019.0013.

Approved for public research: distribution unlimited. (November 28, 2022)

[THIS PAGE INTENTIONALLY LEFT BLANK]

Abstract

There are many scenarios where it is advantageous to prioritize and selectively attend to a single sensory source as well as flexibly engage, disengage, and re-engage with multiple sensory modalities, such as when operating a vehicle. The influences of simultaneous stimuli from different sources and from multiple sensory modalities (e.g., visual and auditory) have been shown to divide attention, which degrades performance. In addition, time-on-task cognitive fatigue can compromise performance and the ability to ignore task-irrelevant information. The aims of this study were to examine whether vigilance-related declines in performance are attributable to supramodal resources that are shared across multiple sensory modalities and to characterize the main source of false alarm errors over time. Stimuli were presented visually and audibly, and participants were instructed to respond to either the visual targets only or the auditory targets only (i.e., focusing on a single modality) or both visual and auditory targets (i.e., alternating attention between modalities). The auditory and visual stimuli involved pairing letters and sounds (phoneme /b/ with the grapheme "b" and phoneme /p/ with the grapheme "p"), resulting in non-targets that could be either modality-relevant non-targets (i.e., the same modality as the target and not the target signal), modality-irrelevant non-targets (i.e., a different modality of the target and not paired with the critical target), or modality-irrelevant target-congruent (i.e., a different modality, but sharing a letter-sound correspondence with the critical signal). While significant effects of cognitive fatigue over time were not detected, there was reduced accuracy when switching attention between auditory and visual targets compared to focusing attention on a target from a single modality. Also, when focusing attention on a single modality, false alarms to modality-relevant non-target items were substantially higher than non-targets that were modality-irrelevant. These results highlight the vulnerabilities of selective attention to task-irrelevant information that are of the same modality as target signals.

Keywords: vigilance, switch-task, sustained attention, task-irrelevant, cognitive fatigue, auditory-visual attention

Author Note: Correspondence should be addressed to Justin Handy, PhD, Naval Submarine Medical Research Laboratory, Submarine Base New London, Groton, CT 06349-5900, email: justin.h.handy.civ@health.mil.

Acknowledgements

The authors would like to acknowledge Ms. Bridget Wilson for her administrative assistance and Mr. Matthew Daley for generating the study stimuli.

Table of Contents

Table of Tables

Table of Figures

Introduction

The ability to selectively attend to information is essential to everyday activities. For example, when driving a car, a driver must focus on road elements that may involve adjacent cars while ignoring music playing on the radio. This is because the environment features more information than can be processed at any given time due to limited cognitive resources.[1] When selectively attending to information, individuals use bottom-up and top-down control. An example of bottom-up control is when attention is biased towards salient physical characteristics in the environment. Top-down control, on the other hand, refers to when an individual uses endogenous factors involving current goals and selection history to focus attention.[2-4] Goal-driven behavior relies on top-down attention to direct cognitive resources towards relevant information despite the presence of competing task-irrelevant information.[5] In general, for both bottom-up and top-down attention, research has shown that selective attention facilitates processing under unimodal conditions when information is from different sources. This leads to improved performance, as evidenced by faster reaction times and more accurate discrimination of task-relevant targets.[6-10] While modalities such as vision and audition are typically studied independently, the natural environment generally involves multimodal stimulation. The interactions between modalities can have important consequences on perception and behavior that may not be observable under laboratory-derived, unimodal conditions.

Multisensory situations can involve scenarios where information from different senses has spatial or temporal correspondence. Agreement or redundancy between sensory modalities, a case of multisensory integration, can help improve performance.[11,12] An example of this sensory agreement is when there is a match between the movement of a person's mouth, the visual information, and the speech heard, the auditory information. When there are disagreements between sensory modalities, the conflicting input can impair performance, which results in longer reaction times.[13] An example of this case is watching delayed speech during a video stream. In the case of ignoring a task-irrelevant modality, there is a cost to performance when compared to a unimodal task.[14] Dividing attention or task switching, the flexible control and shift of attention to disengage and reengage with multiple sources over time, is required for activities such as driving a vehicle, operating an aircraft, and other activities in military settings. For example, the instrument panels and dashboards of military vehicles or systems may require that service members prioritize attendance to visual information (e.g., lights and indicators) at some times and to auditory information (e.g., verbal instructions from a global positioning system, GPS) at other times. These modality shifts may occur over a short duration, such as when driving around a military base. Furthermore, some commands or instructions may be specific to a particular individual's duty. Here, the situation may require that an operator only respond to their auditory call signal, a unique identifier during radio communications, for positions of certain hostile vessels (target modality and target signal) while ignoring other auditory call signals (modality-relevant and irrelevant non-target) and disregarding monitors displaying the visual position of hostile (modality-irrelevant and target congruent) and merchant ships (modality-irrelevant and irrelevant non-target) when visual feeds may not update in real-time.

Prolonged periods of cognitive activity can cause fatigue and compromise performance,[15-17] as witnessed from subjective reports of difficulties concentrating and focusing attention.[18,19] According to cognitive resource theory, cognitive resources can be exhausted faster than they are replenished, which results in a decline in performance over time.[20] However, there is little consensus on whether decrements in performance are the result of depletions in attentional resources that are supramodal (i.e., irrespective of modality—auditory, visual, tactile, or gustatory—they are dependent on the same shared resources),[21-23] modality-specific (i.e., resources are modular and are not shared across modalities[7,24-26]), or a combination of both. This study examined whether switching between auditory and visual targets would deplete resources to a greater extent than focusing on a single sensory modality for targets, which would be evidence for a supramodal attentional resource pool. However, if performance is similar when focusing and switching attention, then this would support a resource pool that was modality specific.

In addition to examining the question of independent cognitive resources, this study investigated the source of errors when focusing on a single modality. Studies have shown that errors or false alarm rates for non-targets increase with time-on-task, implying that distinguishing between task-relevant and task-irrelevant information decreases with cognitive fatigue.[27,28] Research implies that, with respect to processing task-irrelevant information, changes in neural activity over time indicate that the suppression of irrelevant stimuli is less effective with cognitive fatigue.[29] It has been postulated that maintaining a representation of the task context may be attributed to the declines in performance.[30] Research indicates that false alarm errors are more likely for task-irrelevant stimuli that occurred in task-relevant spatial locations and are rare for task-irrelevant stimuli appearing in task-irrelevant spatial locations.[27] Other studies report similar results where manipulating the spatial context for task-irrelevant stimuli elicited fewer errors.[31] Although these studies found that errors were more likely when the stimuli shared features or attributes with the target, both studies used a unimodal task that included only visual stimuli. Extending these findings to multisensory conditions will further characterize errors and elucidate processes that are most vulnerable to time-on-task cognitive fatigue, and lead to the development of strategies that can mitigate its effects. For this reason, a secondary aim of the present study was to investigate errors that are most affected by maintaining vigilance.

The current study investigated cognitive fatigue, the influence of focused attention on performance, and the sources of false alarm errors during a mixed modality auditory-visual task. Critically, participants were instructed to attend to one of the two modalities (focus) or to both modalities (switch) for a relevant stimulus. The conditions requiring the focus of attention allowed for the investigation of task context effects (e.g., sensory modality or sound-letter congruency) when processing task-irrelevant stimuli. The switching of attention task enabled the probing of whether vigilance is supported in part by supramodal resources. The two main hypotheses tested were: (1) when switching attention between two modalities in a manner similar to a dual task, performance would be worse compared to focusing attention on a single modality, lending support for a centralized, supramodal resource system, and (2) more false alarms errors would occur on stimulus trials that share features or attributes with the target (i.e., less errors on modality-irrelevant non-targets) with time-on-task.

Method

Participants.

Study participants included 494 anonymous individuals recruited from the crowdsourcing platform Amazon's Mechanical Turk (MTurk). Participants were compensated $5.00 for their participation in the study. Exclusion criteria for analysis included quitting the experiment before completing all 1000 trials or a hit rate of less than 20 percent during the first quarter of the task

(250 trials), which was used as an indication of inattentiveness.[32] This resulted in a final sample of 333 participants (focus auditory: n = 102; focus visual: n = 119; switch between both visual and auditory: n = 112). Demographic information about participants was not collected. All study procedures were approved by the Naval Submarine Medical Research Laboratory (NSMRL) Institutional Review Board (IRB).

Experimental task.

Participants completed the integrated visual and auditory continuous performance task (IVA-CPT)[33] described previously.[34] A CPT is designed to be monotonous and demand sustained attention. The CPT requires that a participant respond only to trials with a certain target or critical signal and not respond to trials featuring anything that is not the target. The paradigm comprised three conditions in which the presentation of the stimuli was balanced, i.e.,an equal number of trials, but the attentional demands varied. Participants received one of three instructions depending on the condition: (1) respond to a visual target (A−V+), (2) respond to an auditory target (A+V−), or (3) respond to an auditory and visual target (A+V+). Visual stimuli were lowercase characters from the Latin alphabet, “p” and “b,” presented in black font with an ascender/descender length of 5 pixels (**Figure 1A**). For the letter “p,” the descender was the length of 5 pixels that extended below the baseline of the loop. The letter “b” featured an ascender the length of 5 pixels that extended above the baseline of the loop. The auditory stimuli were audio clips of the phonemes /p/ and /b/, with Gaussian noise measured at -16 decibels (dB). These stimulus parameters (5 pixels and -16 dB Gaussian noise) were previously demonstrated to be equally discriminable across modalities with a performance threshold of 80% target identification accuracy.[34] The auditory and visual stimuli represented sound-letter correspondence in that the phoneme /p/ matched with the letter “p” and the phoneme /b/ matched with the letter “b.” Stimuli were presented individually on sequential trials, and participants were instructed to indicate whether the current stimulus was the target, i.e., task-relevant, such that the stimulus presented was both the critical signal and the correct modality presentation of the target (audio, visual, or both), with a button press to the space key on the computer’s keyboard with their dominant hand. Participants were instructed to withhold a response to non-targets.

On each trial, a single stimulus was presented. The display background was white, and a white square outlined in gray appeared at the center of the screen to serve as a fixation point, which remained on screen for the duration of the task. The trial length was fixed at 2000 ms with visual stimuli presented for 167 ms and auditory stimuli for 500 ms. The full task comprised 1000 trials. The trials were half auditory (500 trials) and half visual (500 trials) with the stimuli selected randomly.

There were three conditions: “switch” attention between auditory and visual targets (A+V+), “focus” attention to auditory targets (A+V−), and “focus” attention to visual targets (A−V+) (see **Figure 1B**). In the A+V+ condition, participants were instructed to respond to both a target letter and sound. In the A+V− and A−V+ conditions, participants were instructed to respond to a target in one modality (modality-relevant; i.e., a letter for the visual condition or a sound in the auditory condition) and to ignore the other modality (modality-irrelevant). In the “Focus” conditions, for example, if the target modality was visual then the modality-relevant target (T+) had the same number of trials (80 trials) as the modality-irrelevant target-congruent auditory stimulus (T−, 80 trials) with the same being true for the modality-relevant non-target (NT+, 420 trials) and modality-irrelevant non-target (NT−, 420 trials) stimulus trials. The frequency of each trial type resulted in a lower target/critical signal ratio in the “focus” (A+V−

and A−V+) condition (~9%) compared to the switch (A+V+) condition (~19%). Participants were randomly assigned to complete one of the three conditions. The task took approximately 45 minutes to complete.

Procedure.

Stimulus timing and recording of behavioral responses were controlled by Inquisit Lab 6 (Millisecond® software). Participants were directed to use their computer's speakers. Prior to the start of the task, participants completed ten practice trials to confirm task comprehension.

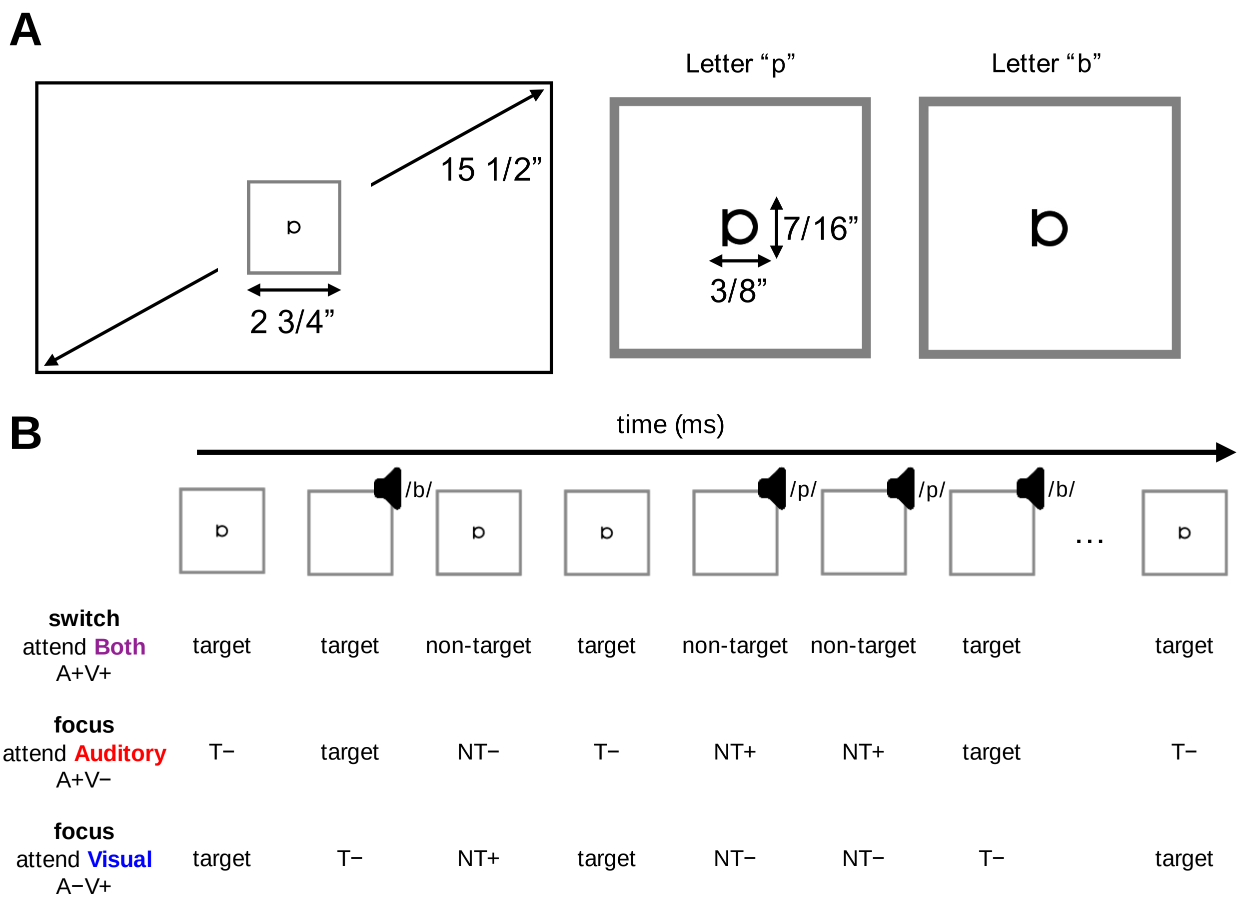

Figure 1. *Experimental Task*

(A) Example display featuring the visual stimulus "p" with the dimensions measured using a 15.5" monitor and the visual stimulus letters "p" and "b" with an ascender/descender length of 5 pixels. (B) Task schematic of the integrated visual and auditory continuous performance task (IVA-CPT) for each condition where the target/critical signal is "b" for the visual modality and /b/ for the auditory modality. The stimuli were presented in random order every 2000ms with a stimulus duration of 167ms for visual and 500ms for auditory. Participants were instructed to only respond to targets and withhold a response to non-targets (NT−, NT+, T−).

Statistical Analysis.

The trials were divided into four blocks of time with 250 trials per block. To test for differences between focusing attention to a single modality (A+V− or A−V+) and switching between both modalities (A+V+) (between-subject factor) and to evaluate changes in performance over time (Block I-IV) (within-subject factor), a 2×4 analysis of covariance

(ANCOVA) was run separately for the auditory and visual trials on correct target identification rates (hit rates) with target letter as a categorical covariate.

Next, to address the second hypothesis of whether the task context impacts errors over time during the "Focus" attention on a single modality (A+V− or A−V+), modality-relevant non-target (NT+), modality-irrelevant non-target (NT−), and modality-irrelevant target-congruent (T−) false alarms were used as dependent measures and submitted to a 2×4×3 ANOVA, with the four blocks (I-IV) and three error types (NT+, NT−, T−) as within-subject factors. The alpha level used was $p < 0.05$ with 2-tailed testing. Follow-up post-hoc tests of significant interaction effects were parsed into lower-order ANOVAs. Analyses were conducted with SPSS, version 23 (IBM Corp). Greenhouse-Geiser corrections were applied when the test of sphericity was violated and an inspection of sample sizes per condition with a threshold ratio of 1.5 was used when homogeneity of variance was violated.[35]

Results

The main dependent measure for determining differences between attending to one versus both modalities was hit rate: how many correct target identifications did participants make per number of stimuli presented. **Figure 2** shows hit rate by the number of modalities attended (both versus one). If attending to both modalities relied on the same cognitive resources, compared to focusing on one of the two modalities, then the switch condition should show a lower hit rate as resources are depleted with time-on-task fatigue. Analyses assessing hit rate effects for auditory and visual trials are summarized in **Table 1**. In comparing the auditory conditions of attending to both modalities (A+V+) and focusing on only the auditory trials (A+V−), hit rate performance was significantly reduced during A+V+ auditory trials (0.59 ± 0.029; Mean ± SE) compared to A+V− (0.79 ± 0.030). No other auditory effects reached statistical significance. This pattern was consistent for the visual condition. Namely, when switching between modalities, there was a significant decrease in hit rates for visual trials during A+V+ (0.57 ± 0.024) compared to A−V+ (0.69 ± 0.023). The main effect of Block and the Block×Condition interaction did not reach statistical significance.

Table 1. ANCOVA Summaries

Effects	F	df	p-value	η_p^2
Hits: Auditory Trials				
Condition (A+V−, A+V+)	10.21	1,211	**.002***	0.046
Block (I-IV)	2.26	2.51,529.77	.092	0.011
Condition×Block	2.06	2.51,529.77	.12	0.01
Hits: Visual Trials				
Condition (A−V+, A+V+)	12.62	1,228	**< .001***	0.052
Block (I-IV)	2.44	2.67,609.20	.07	0.011
Condition×Block	0.40	2.67,609.20	.073	0.002
False Alarm Errors				
Error Type (T−, NT+, NT−)	32.15	1.09,237.68	**< .001***	0.13
Condition (A+V−, A−V+)	1.74	1, 218	.19	0.008
Block (I-IV)	1.49	2.29, 498.50	.22	0.007

Error Type×Block	2.43	3.85, 839.68	**.048***	0.011
Error Type×Condition	3.19	1.09, 237.68	.072	0.014
Condition×Block	2.94	2.29, 498.50	**.047***	0.013
Error Type×Block×Condition	1.40	3.85, 839.68	.23	0.006

Significant correlations (p<.05) are indicated in bold with an asterisk (*)

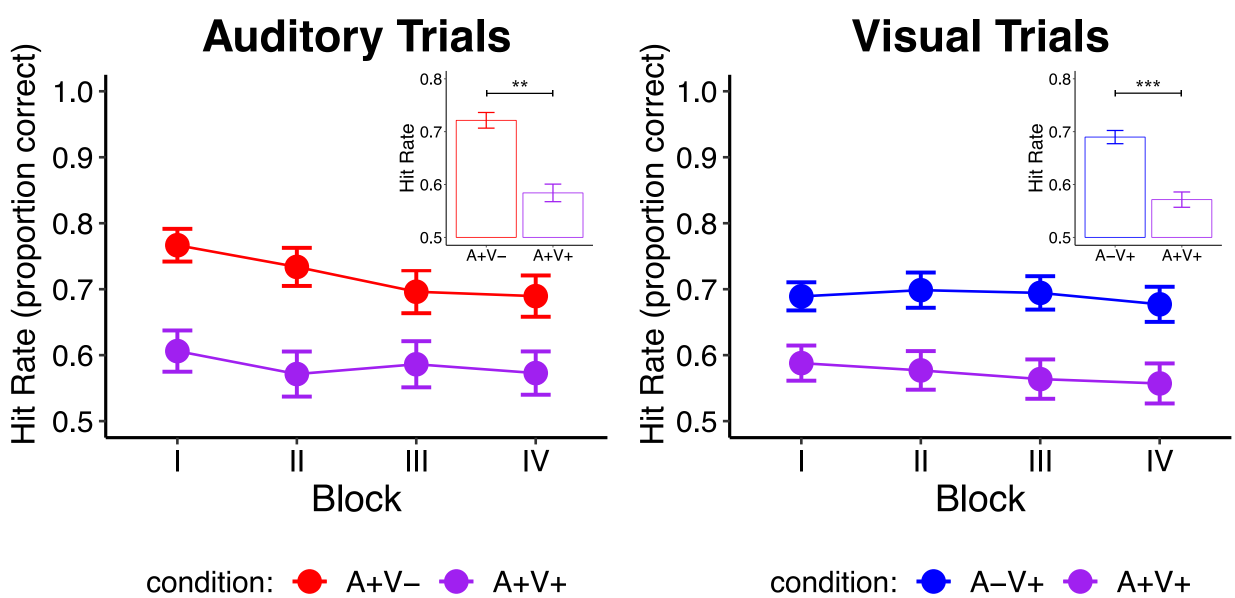

Figure 2. *Results: Hit Rates*
*Correct target identification (hit rate) for the auditory trials that compared the attend auditory and ignore visual (A+V−) shown in red and the attend to both auditory and visual condition (A+V+) in purple (Left). The right figure compares the attend visual and ignore auditory (A−V+) condition shown in blue to the visual trials of A+V+ condition. Inset show the main effects of condition collapsed across Blocks. Errors bars ± 1 standard error, ** p < 0.01, *** p < 0.001*

Results for false alarm errors are summarized in **Table 1**. To test for task context (modality and letter-sound correspondences) influencing errors, false alarm rates were compared for modality-relevant non-target trials (NT+; e.g., responding to the phoneme /b/ when the target was phoneme /p/), modality-irrelevant target-congruent trials (T−; e.g., responding to the letter “p” when the target was phoneme /p/), and task-irrelevant non-target trials (NT−; e.g., responding to the letter “b” when the target was phoneme /p/) of the “Focus” attention conditions (A+V− and A−V+). As shown in **Figure 3**, there was a main effect of error type indicating NT+ trials elicited a higher false alarm rate (0.44 ± 0.020) compared to the NT− trials (0.24 ± 0.020, $p < 0.001$) and T− trials (0.24 ± 0.021, $p < 0.01$). The NT− and T− (the task-irrelevant stimuli) were not detected as statistically different from each other ($p = 1.00$). The main effect of condition did not reach statistical significance, such that in terms of overall error rates, differences between the A+V− condition (0.33 ± 0.027) and A−V+ condition (0.28 ± 0.025) were not detected. Additionally, the main effect of Block on error rates did not reach statistical significance. There was a significant Block×Condition interaction and Error Type×Block interaction. The Error Type×Condition and Error Type×Condition×Block interactions did not reach statistical significance.

Parsing the significant Block×Condition interaction, the average overall false alarm errors by Block were subjected to an ANCOVA with a between subject factor of condition (A+V−, A−V+). Block I had a significant effect of condition ($F(1,218) = 4.54$, $p = 0.034$, $\eta_p^2 = 0.020$) with the A+V− having a higher false alarm rate than the A−V+ condition. There were no significant differences detected between the remaining Blocks ($p > 0.05$).

The Error Type×Block interaction was interrogated further with a repeated measures ANCOVA with block as the within-subject factor for each error type (NT+, NT−, T−) independently. None of the follow-up analyses reached statistical significance (NT+: $F(2.33, 509.84) = 2.83$, $p = 0.052$, $\eta_p^2 = 0.013$; NT−: $F(2.31, 505.98) = 0.82$, $p = 0.46$, $\eta_p^2 = 0.004$; T−: $F(2.51, 548.92) = 0453$, $p = 0.68$, $\eta_p^2 = 0.002$).

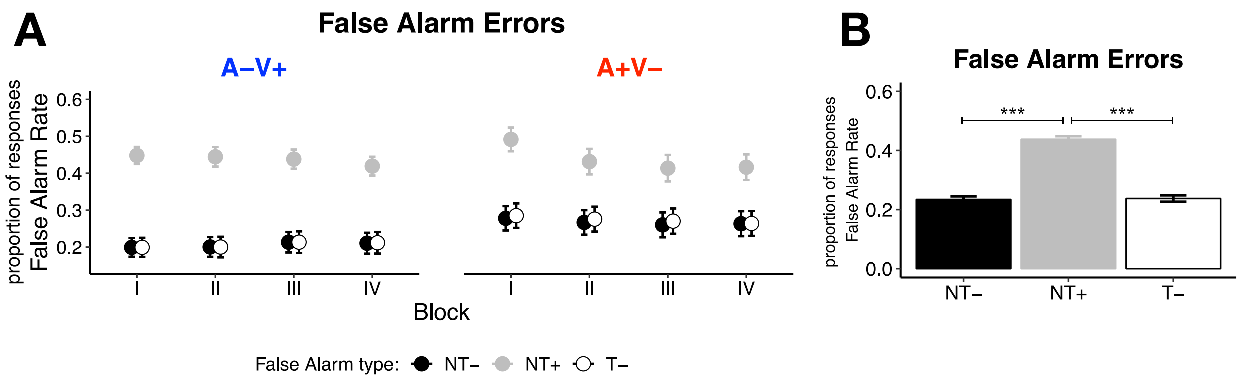

Figure 3. *Results: False Alarms Errors*
(A) Errors (false alarm response rate) for the (Left) attend visual and ignore auditory (A−V+) and (Right) attend auditory and ignore visual (A+V−) condition for trials that were non-target modality-relevant (NT+, filled gray), non-target modality-irrelevant (NT−, filled black), and the target modality-irrelevant stimuli (T−, unfilled black). (B) False alarm errors collapsed across Block and condition. Errors bars ± 1 standard error, *** $p < 0.001$.

Discussion

As environmental information is constantly inundating our senses, limitations in cognitive resources dictate the necessity for selectively attending to certain signals/stimuli. In this study, the basis for selective attention was extended in two regards by showing that: (1) when instructed to switch attention between modalities there was a reduction in target detections compared to when instructed to focus on a single modality, and (2) modality-relevant non-target stimuli are a more likely source of errors in comparison to modality-irrelevant stimuli for situations where the sources of information are considered conflicting and independent.

The first hypothesis predicted that monitoring for a target/critical signal in both auditory and visual modalities (A+V+) would result in worse performance than focusing on a single modality (A+V− or A−V+). The present findings support an account of supramodal resources by demonstrating overall lower hit rates in the auditory-visual “switching” attention condition when compared to the “focused” conditions. If resources were modality-specific, performance on the switch task would have reflected similar performance to the focused attention task as each modality would be relying on independent resources. Neuroimaging studies align in part with supramodal resources supporting attention whereby brain areas that have been implicated in visual vigilance are also activated during auditory vigilance.[36] These results compliment and are consistent with the dual task literature that demonstrate worse performance when completing multiple tasks simultaneously.[37,38] These results cannot be explained by the perceptual load (amount of sensory information required to process the current stimulus) or memory load (amount of working memory resources needed to process the current stimulus). Holding constant the discriminability of the target from non-target stimuli between auditory and visual modalities and the correspondence between the sound and letter required equal perceptual demands and reduced the memory demands during the attention switching condition. This allowed for the examination of performance changes that reflected the attentional resources in monitoring for a critical signal in two modalities. Studies on alternating attention, with the widely used neuropsychological instrument Trail Making Test,[39] report that inhibiting the prior task set impairs performance.[40,41] The results from the current study would be consistent with these reports, such that a supramodal center would be conducive to conditions that result in interference and inhibition challenges during switching. The stimuli characteristics differentiate the current study from others that have explored bimodal selective attention[26,42] or have stimuli presented concurrently.[43,44]

These results complement a prior study in which a mixed auditory-visual condition resulted in lower hit rates than both unimodal auditory and visual conditions.[34] In this prior study— unlike the current study that maintained a balance of the perceptual information, but differed in the number of target trials presented— the number of target trials were equated but the perceptual information differed. The tasks were either of one modality or a mix of two modalities. Under these conditions, participants would not have time to recover resources between relevant trials, if it were the case that resources were replenished between trials. We extend these findings by showing that monitoring for a signal in two modalities impacts performance by degrading performance to a greater extent than when monitoring for a signal in one modality, as demonstrated by the lower hit rate.

As vigilance has modulatory effects on attention, it impacts the immediate processing of information, it is vital to understand the role of vigilance when alternating attention. The results from the present study did not find any changes in performance over time. The lack of detecting a significant finding could be due to several factors. While vigilance decrements have been reported to occur within the first 15 minutes[45], and in some cases as early as 5 minutes,[46] (the present study required over 30 minutes to complete), this decline may depend and interact with task parameters. Studies have found an enhancement in the vigilance decrement as a function of signal/stimulus frequency is that detection rate decreases as the signal/stimulus frequency decreases.[47] In the current study, the target or critical signal to non-critical signal ratio was higher in the conditions where participants were attending to both visual and audio stimuli. However, while focusing on a single modality had less target/critical signals to detect, a vigilance decrement was not observed in these conditions. Many studies have used a critical signal probability between 3% and 30%.[48,49] Here, the critical signal probability was 19% for attending to both and 9% for attending to a single modality. These signal ratios are within the range of observing a vigilance decrement, though one that might be less pronounced than a lower critical signal probably. Another task parameter that is known to influence detection rates is the difficulty of signal detection, where highly degraded signals produced a larger vigilance decrement.[48] While the stimuli used in the current study were comparable in difficulty (i.e., a similar hit rate) to the highly degraded signal in the Neuchterlein et al., (1993) study (~85% accuracy), the participants had less practice than those of Neuchterlein et al. (1993). This might indicate that learning effects were occurring simultaneously with the vigilance decrement, resulting in a cancelation effect, so that as learning increased performance over time there was a corresponding magnitude decrease related to the maintenance of vigilance. This is supported by the higher errors rates in Block I in comparison to the other blocks, as participants were becoming more familiar with the stimuli.

The second hypothesis examined the extent to which certain errors are more vulnerable to time-on-task cognitive fatigue. As noted above, we did not detect any time-on-task performance degradations. However, we did find that errors associated with modality-relevant non-target (NT+) stimuli were overall more common (for example, during the task to focus attention on the visual stimuli, responding to the letter "p" when the target letter was "b"). Previous studies have found greater errors on non-target stimuli presented on spatially relevant locations[27] and targets at spatially irrelevant locations[31] where these trials shared features or attributes with the critical target signal. However, these studies of selective attention have involved only a single modality. In the current study's design, task context is represented by either the same modality or a letter-sound congruency. The results revealed that a more likely source of error occurs for stimuli that share the same task context of modality as the critical target signal. Research has found a global facilitation of feature-based attention that interacts with temporal perception.[50] Individuals tend to have bias processing of objects/stimuli that have features or attributes consistent with what they are selectively attending. The strength of this bias depends on the extent to which the irrelevant object activates the same neural structures as the relevant object.[51] As a consequence of feature-based attention, the results suggest that interference from the target signal overlapped in features with the modality-relevant non-target stimulus (NT+) to a greater degree than the modality-irrelevant target-congruent stimuli (T−). This interference elevated the errors associated with those trial types. There is a possibility that participants focusing attention on a single modality condition could have disregarded task instructions and altered the delivery of the stimuli such that participants focusing on the visual stimuli may have muted their speakers. Conversely, participants in the focus attention to the auditory stimuli cases could have closed their eyes. However, if this were true false alarm errors to the task-irrelevant stimuli would have been absent or rare.

In conclusion, our results demonstrate that monitoring two modalities for a critical abstract display signal is cognitively more resource demanding compared to selectively attending to abstract information via a single sensory modality that are not necessarily linked or generated from the same source, as evidenced by decreased accuracy for two modalities. While the stimuli, auditory and visual, were generated from the same computer there was obvious link that the auditory and visual were paired together such as with a mouth that generating speech. In the mouth-speech example, the two are related signals. In this study the signals were presented sequentially to separate the stimuli from one another. This finding extends the empirical evidence of the dual task literature that consistently reports reduced performance of concurrent stimuli. The results also indicate that irrelevant features that overlap with a target object interfere with processing and contribute to the source of errors during a selective attention task. These results have potential applications for environments and situations that require the management of information from multiple sensory sources (e.g., sonar operators). Specifically, when sensory sources overlap in modality, decreasing task similarity may help reduce conflicts and errors. Limitations of this study are the inherent difficulties in a non-laboratory-based environment where there is less control of the test conditions and the participants enrolled. Nonetheless, despite the added noise as a result of these factors, these analyses revealed remarkably clear patterns in the data. Future research is needed to fully characterize these sources of errors, which will ultimately contribute to our understanding of the limitations when switching attention.

Reference

1. Pashler H, Johnston JC, Ruthruff E. Attention and performance. *Annu Rev Psychol.* 2001;52:629-651.
2. Pinto Y, van der Leij AR, Sligte IG, Lamme VA, Scholte HS. Bottom-up and top-down attention are independent. *J Vis.* 2013;13(3):16.
3. Posner MI. Orienting of attention. *Q J Exp Psychol.* 1980;32(1):3-25.
4. Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model for visual search. *J Exp Psychol Hum Percept Perform.* 1989;15(3):419-433.
5. Chun MM, Golomb JD, Turk-Browne NB. A taxonomy of external and internal attention. *Annu Rev Psychol.* 2011;62:73-101.
6. Bonnel AM, Hafter ER. Divided attention between simultaneous auditory and visual signals. *Percept Psychophys.* 1998;60(2):179-190.
7. Spence C, Driver J. On measuring selective attention to an expected sensory modality. *Percept Psychophys.* 1997;59(3):389-403.
8. Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu Rev Neurosci.* 1995;18:193-222.
9. Payne L, Guillory S, Sekuler R. Attention-modulated alpha-band oscillations protect against intrusion of irrelevant information. *J Cogn Neurosci.* 2013;25(9):1463-1476.
10. Huang J, Sekuler R. Distortions in recall from visual memory: two classes of attractors at work. *J Vis.* 2010;10(2):24 21-27.
11. Diederich A, Colonius H. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept Psychophys.* 2004;66(8):1388-1404.
12. Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res.* 2003;17(2):447-453.
13. Donohue SE, Appelbaum LG, Park CJ, Roberts KC, Woldorff MG. Cross-modal stimulus conflict: the behavioral effects of stimulus input timing in a visual-auditory Stroop task. *PLoS One.* 2013;8(4):e62802.
14. Tulving E, Lindsay PH. Identification of simultaneously presented simple visual and auditory stimuli. *Acta Psychol (Amst).* 1967;27:101-109.
15. Parasuraman R, Mouloua M. Interaction of signal discriminability and task type in vigilance decrement. *Percept Psychophys.* 1987;41(1):17-22.
16. Warm JS, Parasuraman R, Matthews G. Vigilance requires hard mental work and is stressful. *Human factors.* 2008;50(3):433-441.
17. Mackworth NH. The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology.* 1948;1(1):6-21.
18. Bartlett FC. Ferrier Lecture-Fatigue following highly skilled work. *Proceedings of the Royal Society of London Series B-Biological Sciences.* 1943;131(864):247-257.
19. Brown ID. Driver fatigue. *Hum Factors.* 1994;36(2):298-314.
20. Davies DR, Parasuraman R. *The psychology of vigilance.* Academic Press; 1982.
21. Wahn B, Konig P. Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Front Psychol.* 2015;6:1084.
22. Arnell KM, Jenkins R. Revisiting within-modality and cross-modality attentional blinks: effects of target-distractor similarity. *Percept Psychophys.* 2004;66(7):1147-1161.

23. Jolicoeur P. Restricted attentional capacity between sensory modalities. *Psychon Bull Rev.* 1999;6(1):87-92.
24. Alais D, Morrone C, Burr D. Separate attentional resources for vision and audition. *Proc Biol Sci.* 2006;273(1592):1339-1345.
25. Duncan J, Martens S, Ward R. Restricted attentional capacity within but not between sensory modalities. *Nature.* 1997;387(6635):808-810.
26. Spence C, Nicholls ME, Driver J. The cost of expecting events in the wrong sensory modality. *Percept Psychophys.* 2001;63(2):330-336.
27. Boksem MA, Meijman TF, Lorist MM. Effects of mental fatigue on attention: an ERP study. *Brain Res Cogn Brain Res.* 2005;25(1):107-116.
28. Lorist MM, Klein M, Nieuwenhuis S, De Jong R, Mulder G, Meijman TF. Mental fatigue and task control: planning and preparation. *Psychophysiology.* 2000;37(5):614-625.
29. Faber LG, Maurits NM, Lorist MM. Mental fatigue affects visual selective attention. *PLoS One.* 2012;7(10):e48073.
30. Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol Rev.* 1992;99(1):45-77.
31. Hahn M, Wild-Wall N, Falkenstein M. Age-related differences in performance and stimulus processing in dual task situation. *Brain Res.* 2011;1414:66-76.
32. Neigel AR, Claypoole VL, Hancock GM, Fraulini NW, Szalma JL. Ceci n'est pas un walrus: lexical processing in vigilance performance. *Exp Brain Res.* 2018;236(3):867-879.
33. Sandford JA, Turner A. Intermediate visual and auditory continuous performance test interpretation manual. *Richmond, VA: Braintrain.* 1995.
34. Peltier C, Daley M, Guillory S, Handy JD, Wilson B. Vigilance Performance Declines Faster When Monitoring for a Signal in Two Modalities Compared to One. *Advances in Cognitive Psychology.* 2022;18(1).
35. Blanca MJ, Alarcon R, Arnau J, Bono R, Bendayan R. Effect of variance ratio on ANOVA robustness: Might 1.5 be the limit? *Behav Res Methods.* 2018;50(3):937-962.
36. Langner R, Eickhoff SB. Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. *Psychol Bull.* 2013;139(4):870-900.
37. Koch I, Poljac E, Muller H, Kiesel A. Cognitive structure, flexibility, and plasticity in human multitasking-An integrative review of dual-task and task-switching research. *Psychol Bull.* 2018;144(6):557-583.
38. Fougnie D, Cockhren J, Marois R. A common source of attention for auditory and visual tracking. *Atten Percept Psychophys.* 2018;80(6):1571-1583.
39. Battery AIT. Manual of directions and scoring. In: Washington, DC: War Department, Adjutant General's Office; 1944.
40. Kortte KB, Horner MD, Windham WK. The trail making test, part B: cognitive flexibility or ability to maintain set? *Appl Neuropsychol.* 2002;9(2):106-109.
41. Arbuthnott K, Frank J. Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J Clin Exp Neuropsychol.* 2000;22(4):518-528.
42. Sutton S, Hakerem G, Zubin J, Portnoy M. The effect of shift of sensory modality on serial reaction-time: A comparison of schizophrenics and normals. *The American Journal of Psychology.* 1961;74(2):224-232.

43. Lupker SJ. The semantic nature of response competition in the picture-word interference task. *Memory & Cognition.* 1979;7(6):485-495.
44. Pashler H. Dual-task interference in simple tasks: data and theory. *Psychological bulletin.* 1994;116(2):220.
45. Teichner WH. The detection of a simple visual signal as a function of time of watch. *Hum Factors.* 1974;16(4):339-353.
46. Temple JG, Warm JS, Dember WN, Jones KS, LaGrange CM, Matthews G. The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Hum Factors.* 2000;42(2):183-194.
47. Craig A, Davies DR, Matthews G. Diurnal variation, task characteristics, and vigilance performance. *Hum Factors.* 1987;29(6):675-684.
48. Nuechterlein KH, Parasuraman R, Jiang Q. Visual sustained attention: image degradation produces rapid sensitivity decrement over time. *Science.* 1983;220(4594):327-329.
49. Baddeley AD, Colquhoun WP. Signal probability and vigilance: a reappraisal of the 'signal-rate' effect. *Br J Psychol.* 1969;60(2):169-178.
50. Ono F, Yamada K, Chujo K, Kawahara J. Feature-based attention influences later temporal perception. *Percept Psychophys.* 2007;69(4):544-549.
51. Fias W, Lauwereyns J, Lammertyn J. Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Brain Res Cogn Brain Res.* 2001;12(3):415-423.