# Neural evidence of visual-spatial influence on aural-verbal processes


**Elizabeth L. Fox**
elizabeth.fox.9@us.af.mil
Air Force Research Laboratory
WPAFB, OH 45433 USA

**Margaret Ugolini**
margaret.ugolini.ctr@us.af.mil
BAE Systems, Inc., Space & Mission Systems
Dayton, OH 45435 USA

**Ashley D. Cook**
ashley.cook.11.ctr@us.af.mil
Air Force Research Laboratory
WPAFB, OH 45433 USA

**Gregory Bowers**
gregory.bowers.ctr@us.af.mil
BAE Systems, Inc., Space & Mission Systems
Dayton, OH 45435 USA



## Abstract

Everyday tasks demand attentional resources to perceive, process, and respond to important information. Attempting to complete multiple tasks simultaneously, that is, multitasking, necessarily requires more resources than completing either task alone. Allocating common resources among two or more difficult tasks will lead to competition and result in performance deficits to one or more of the to-be-completed tasks. Multiple resource theory suggests separate pools for perceiving (aural, visual, tactile), processing (verbal, spatial), and responding (vocal, manual), but a common overarching resource pool still exists and is heavily taxed for the management of multiple ongoing tasks. We use the combination of neural activity and performance to estimate the degree to which the demands of a visual-spatial-manual (VSM) task impedes the performance of an auditory-verbal-vocal (AVV) task, where each taxes independent pools of attentional resources. We found AVV performance decreased when paired with a more difficult VSM task. Using components from group-level event related potentials (ERPs), we draw conclusions to estimate how and why cross-modal task performance changes, and diagnose resource bottlenecks and limitations. Specifically, we find auditory evoked potentials, P300, and Reorienting Negativity serve as fruitful indicators of not only high or low cross-modal load, but are predictive of (in)correct trial performance. Further, we discuss how these indicators provide insight to the underlying mechanisms driving misses, and whether crossmodal bottlenecks may occur at the perceptual, cognitive, or response stage.



**Keywords:** Cross-modal Influence; Multi-tasking; Event Related Potentials; Mental Workload


## Introduction

Humans have a limited attentional resource pool to meet critical task demands (Norman & Bobrow, 1975). Thoughtful multi-task design can tax separate input modalities (i.e., visual, auditory, tactile), processing codes (i.e., spatial, verbal), and response modes (i.e., verbal, manual) to alleviate some competition for perceptual, cognitive, and response resources (Wickens, 2002). However, cognitive interference often still exists and impedes performance in one or more tasks. Attentional overload often results in failures to sufficiently redirect attention to acoustic information, called inattentional deafness, and critical auditory signals in multi-modal environments go undetected (Macdonald & Lavie, 2011). Specifically, inattentional deafness increases when paired with a visual task with high (v. low) demands (Dehais, Roy, & Scannella, 2019a). We use a neurobehavioral approach to determine at what stage of processing visual-spatial demands impede the reallocation of attention to, and the cognitive processing of, auditory-verbal information. We show how future work may utilize this framework to determine the bounds of introducing or mediating the effects of inattentional deafness in multi-modal, multi-tasking environments.

## Auditory Evoked Potentials as a Marker of Selective Auditory Attention

Event Related Potentials (ERPs) have been used for decades to understand cognitive processes as they unfold on a millisecond-level timescale (Hillyard & Kutas, 1983). An ERP effect's scalp topography, timing, eliciting stimulus and overall task environment can be used to draw conclusions about the underlying cognitive processes involved in the task. Auditory Evoked Potentials (AEPs) include relatively early ERP effects in response to auditory stimuli. The N1-P2 (or N100 P200) complex is the most commonly studied portion of the AEP. This complex includes a negativity maximal at Cz (a centrally located electrode site) peaking around 100ms followed by a positivity peaking around 200ms. Occasionally, the earlier P1 (or P50) effect is included as an AEP. This effect is quite early, peaking around 50ms but is also quite low in signal-to-noise ratio (SNR) and therefore is not discernible in many auditory paradigms. The N1 of the AEP in particular is modulated by selective attention. Attended sounds elicit a greater N1 than unattended sounds. This is true when participants base selective attention on a variety of stimulus characteristics, such as pitch (Hansen & Hillyard, 1983), timing (Sanders & Astheimer, 2008; Astheimer & Sanders, 2009), and spatial location (Hillyard, Hink, Schwent, & Picton, 1973). Changes in the amplitude of AEPs can be used to draw conclusions about selective auditory attention in complex, realistic tasks (Fitzroy et al., 2020; Hölle, Blum, Kissner, Debener, & Bleichner, 2022), where the AEP is more robust for attended vs. unattended stimuli. AEPs to unattended, task irrelevant sounds are also informative. A larger AEP response is seen under lower workload, compared to high workload, while completing a variety of multimodal tasks (Xu et al., 2020; Ghani, Signal, Niazi, & Taylor, 2020; Wang et al., 2023; Solís-Marcos & Kircher, 2019).




In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

## The P300 as a Marker of Task-relevant Stimuli and Spare Capacity

The P300 is a positive ERP effect, centrally located, and peaking around 300ms after task relevant or rare stimuli (Sutton, Braren, Zubin, & John, 1965; Polich, 2007). P300 effects are commonly investigated using oddball paradigms, where participants listen to infrequent deviant stimuli, embedded among frequent standard stimuli, and are often required to respond behaviorally to the deviants, either by counting them or pressing a button when they are detected (Donchin, Ritter, McCallum, et al., 1978). The amplitude of the P300 is sensitive to workload, particularly related to demands on sensory information processing (Pritchard, 1981). The P300 is a robust effect that has been successfully used as a marker of the processing of task-relevant stimuli approximating real-world scenarios (Solís-Marcos & Kircher, 2019) and outside of the laboratory environment (Liebherr et al., 2021). Notably, this ERP effect has been used to address safety concerns related to aviation. The P300 could successfully predict auditory attentional deafness in-flight using real pilots as research participants (Dehais, Rida, et al., 2019; Dehais, Duprès, et al., 2019). The P300 is often used to understand how multimodal task demands influence workload, spare capacity, and perception of task relevant stimuli (Ghani et al., 2020).

## The Reorienting Negativity as a Marker of Reallocation of Attention

The Reorienting Negativity (RON) appears at fronto-central electrode sites about 400-900ms after a stimulus that requires attentional reallocation (Schröger & Wolff, 1998). The RON effect is stronger in working memory tasks. Munka and Berti (2006) argue that the RON can be divided into an early and a late portion, with the early portion (at roughly 400ms) only present when working memory is required to complete the task (v. sensory judgment). Presence and magnitude of the RON can provide information about the attentional reallocation demands imposed by a task characteristics and workload level; for instance, it has recently been used to a biomarker of impaired attentional processing in individuals with ADHD (Gumenyuk et al., 2023).

## Current Study

In the present study, we leverage these three distinct ERP effects to investigate attention and cognitive processing in a challenging multimodal multitask scenario. Specifically, we investigate how visual-spatial workload of varying levels influences auditory-verbal processing during a challenging task. The task of interest requires auditory pitch discrimination and processing of human utterances, maintenance and manipulation of verbal items in working memory, and the production of a vocal response. We specifically investigate these effects as a function of behavioral outcome in the auditory task, to see how workload influences cross-modal cognitive processing and performance.

# Method

Twenty people at or nearby a Midwestern university completed the study and were compensated $15/hour or at their normal hourly wage. Participants completed six 2-hour sessions.

## Apparatus

The lab was equipped with 2 desktop computers (MATLAB, Python), Fireface UCX audio card, ear buds, and a clip-on microphone. The EEG signals (1024 Hz) were recorded using 64 Ag-AgCl pintype active electrodes (ActiveTwo, BioSemi) mounted on an elastic cap (ECI) according to the extended 10-20 system, and from two additional electrodes placed on the right and left mastoids. Eye movement and blinks were monitored using 4 EOG electrodes (up/down, left/right). A ViewPixx LCD Display was placed at 110cm viewing distance to administer visual stimuli; a ViewPixx trigger cable was used to record labeled event onsets.

## Visual Task

Participants were shown continuously moving colored dots on a square (800 x 800 pixel) display. Black lines divide the display into four equal quadrants. One or two quadrants had a gray background, indicating target areas of interest (AOI), and the rest had a white background.

**Task Description** Participants were instructed to turn on and off alerts depending on the status of one (low visual workload; LVW) or two (high visual workload; HVW) AOI. If the ratio of red to pink dots (10px radius) in the AOI was greater than 1, an alert was raised; when the ratio was equal or less than 1 the alert was removed. Magenta dots were ignored. Interested readers can refer to Fox, Bowers, Capiola, and Stephenson (2023) for more details about the visual task.

## Auditory Task

**Stimuli** The auditory stimulus was a set of 4 random letters (A-Y; 500-1000ms). One of 150 random speakers was chosen from the Center for Spoken Language Understanding (CSLU) repository on each trial (Linguistic Data Consortium, 2002). The first letter was preceded by a 150 ms pure tone "Warning" at 50dB (Ventry, Woods, Rubin, & Hill, 1971), which were either sampled from a "high pitch" ($M$ = 1784 Hz, $SD$ = 54 Hz) or "low pitch" ($M$ = 588 Hz, $SD$ = 22 Hz) set of frequency distributions, which were chosen based on the Equal-Loudness contour (Suzuki & Takeshima, 2004).

Target "warning" cues (high frequency) had equal prevalence to distractor cues (low frequency). Cues were not always followed by letters (40%). When letters were not present, there was a variable inter-stimulus interval (ISI; 500-1000 ms) until the next cue was presented. Letters were presented 500 ms after 40% of cues, irrespective of whether the last cue was a target or distractor. Participants were instructed to only respond when letters were immediately preceded by a target cue (20%).

Table 1: ERP effects investigated and their associated time windows

| | Cue (tone) | Letter (speech) |
|---|---|---|
| AEPs | 10 - 300ms | 150 - 250ms |
| P300 | 300 - 500ms | 350 - 550ms |
| RON | 400 - 700ms | 450 - 750ms |

Note. AEP: Auditory Evoked Potential; RON: Reorienting Negativity.

**Task Description** If a target cue was followed by letters, the participant was instructed to decrypt the $1^{st}$ and $4^{th}$ letters and make a vocal response. Decryption was simply the letter that follows in the alphabet. For instance, "D, T, M, W" demands a response "E, X".

**Training** Participants were trained to correctly distinguish two sets of auditory "warning" cues before each session. Session 1 and 2 consisted of training of the visual task, both easy and hard conditions, and the aural task, respectively. Session 3 and 5 consisted of training on the easy and hard dual-tasks, respectively, where each was immediately followed by the session with EEG recordings (Session 4 and 6).

## Neural Data Cleaning, Processing, and Statistics

EEG data were filtered minimally based on guidelines from Delorme (2023). Data were rereferenced to the average of the mastoids, filtered with a 0.5Hz high pass filter. Artifact Subspace Reconstruction (ASR) was used for data cleaning (Mullen et al., 2015), followed by Independent Component Analysis (ICA) (Makeig, Bell, Jung, & Sejnowski, 1995), with the removal of all components identified by IC label (Pion-Tonachini, Kreutz-Delgado, & Makeig, 2019) to be eye or muscle contamination with a 90% degree of certainty or greater. Any channels removed by the ASR process were then interpolated. Data was then epoched around events of interest, starting 100ms before the event onset to 800ms after.

Statistical analysis was conducted using cluster based permutation analysis (Groppe, Urbach, & Kutas, 2011; Maris & Oostenveld, 2007; Sassenhagen & Draschkow, 2019), with a mass univariate technique with corrections for multiple comparisons. For all comparisons of interest, the difference score of the ERP evoked by the two conditions was submitted to a repeated measures, two-tailed cluster-based permutation test (Bullmore et al., 1999; Manly, 1997), $\alpha = 0.05$. Comparisons were made at electrode sites and time windows relevant in previous research. Time windows are summarized in Table 1.

Different montages of electrodes were tested, depending on the effect being investigated. The montage for AEPs and the RON included 12 centrally located electrodes (left; Figure 1). The montage for the P300 included 13 central and parietal electrodes (right; Figure 1). Both montages were selected based on prior literature (Dehais, Rida, et al., 2019; Dehais, Duprès, et al., 2019; Dehais, Roy, & Scannella, 2019b; Oray, Lu, & Dawson, 2002).

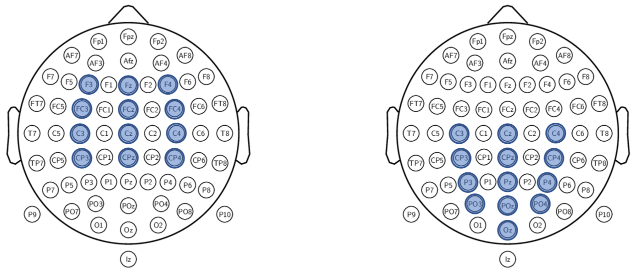


Figure 1: Montages used to investigate the Auditory Evoked Potential and Reorienting Negativity (left), and P300 (right).

Each ERP effect of interest was investigated in response to the types of events that traditionally elicit the effect. As such, we included responses to cues (AEP, RON), Letter 1 (AEP, P300, RON), and Letter 4 (RON). A target Letter 1 determined when the decryption task was required; Letters 2-4 were unlikely to show clear ERP effects since these events are continuous speech. Our investigation of the RON was exploratory; reorientation of attention could occur after cues or letters of interest.

# Results

We assessed Hits, Misses, Correct Rejections (CR), and False Alarms (FA) (Macmillan & Creelman, 2004). For target present trials (high frequency tone followed by a 4-letter string), a Hit was defined as a correct vocal response to *both* to-be-decrypted letters, and a Miss was when no vocal response was made. For target absent trials (low frequency tone followed by a 4-letter string), a CR was when no vocal response was made, and a FA was when a correct vocal response was made. For both trial types, incorrect or partial utterances were not analyzed beyond overall accuracy; FAs were infrequent and therefore excluded from neural interpretation.

## Behavioral Data

One participant was excluded from all analyses due to a low number of Hit trials. Any subset exclusion is noted in the corresponding subsection below.

**Visual Task** Participants' more slowly (response time [RT] in seconds) and less accurately (Acc) turned on alarms when the visual task difficulty was high ($M[SD]_{RT} = 1.55[1.38]$; $M[SD]_{Acc} = 63.2[6.4]$) versus low ($M[SD]_{RT} = 1.04[1.16]$; $M[SD]_{Acc} = 72.5[7.3]$), $t_{RT} = 19.02, p_{RT} < .001; t_{Acc}(37.4) = 4.31, p_{Acc} < .001$.

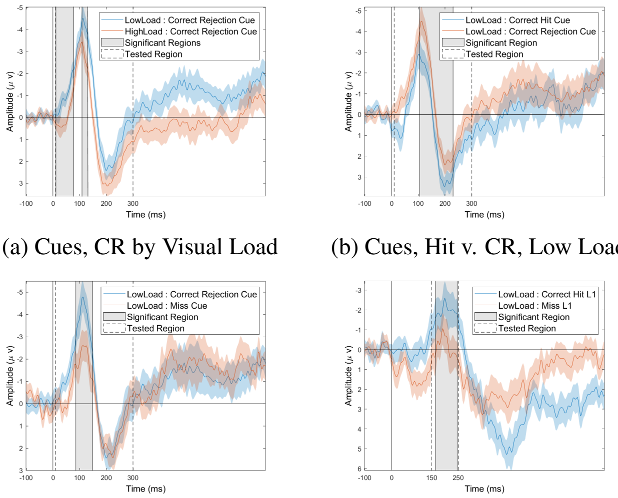
**Auditory Task** Participants' initiation of a correct vocal response slowed by 232 ms, $t = 8.06, p < .001$, and their hit/miss ratio decreased when attempting to simultaneously complete a visual task with high (0.98), versus low (1.28)

Table 2: Comparisons of Event Related Potentials as a function of workload and behavioral outcome.

| | AEP | | P300 | RON | | |
|---|---|---|---|---|---|---|
| Condition | Cue | L1 | L1 | Cue | L1 | L4 |
| CH, L vs H | – | – | * | – | – | – |
| CR, L vs H | * | – | – | *** | – | – |
| M, L vs H | – | – | – | – | – | – |
| CH vs M, L | – | * | * | – | – | – |
| CH vs M, H | – | – | – | – | – | * |
| CH vs CR, L | * | – | * | – | – | – |
| CH vs CR, H | – | – | * | – | – | * |
| CR vs M, L | * | – | – | – | – | – |
| CR vs M, H | – | – | – | – | – | – |

*$p < 0.05$ = *; $p < 0.01$ = **; $p < 0.001$ = ***; AEP = Auditory Evoked Potential; RON = Reorienting Negativity; CH = Correct Hit; CR = Correct Rejection; M = Miss; L = Low Visual Workload; H = High Visual Workload.*

spatial demands, where 1.0 is an equal number of hits and misses, and more (less) than 1.0 is greater (fewer) hits to misses.




(a) Cues, CR by Visual Load

(b) Cues, Hit v. CR, Low Load

(c) Cues, CR v. Miss, Low Load

(d) L1, Hit v. Miss, Low Load

Figure 2: Significant differences in Auditory Evoked Potentials. Colored shading represents standard error of the mean. CR: Correct Rejection; L1: Letter 1

## Auditory Evoked Potentials

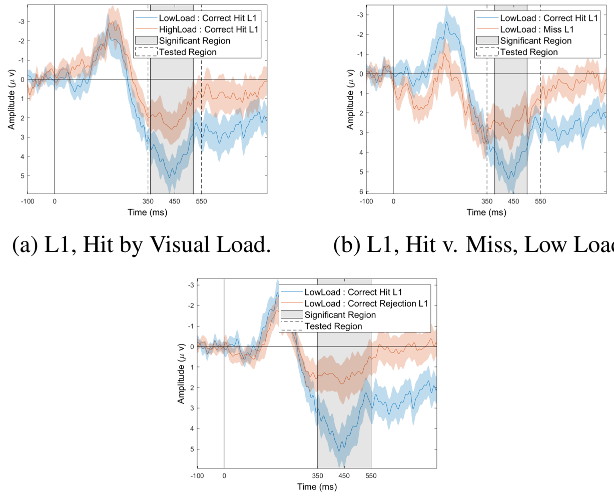
We investigated the AEP response to Cues and Letter 1 stimuli within and across workload and behavioral outcome. Table 2 depicts the results of all comparisons. Only significant results are summarized in the text below.

**AEPs to Cues** AEPs to cues that led to CRs differed as a function of workload, such that the N1 was stronger under LVW ($p = 0.0328$; Figure 2 a). AEPs to Cues differed for Hits and CRs under LVW, with a greater N1 for CRs compared to Hits ($p$ = 0.0144; Figure 2 b). AEPs to Cues also differed for CRs and Misses under LVW, with a greater N1 for CRs compared to Misses ($p = 0.016$; Figure 2 c).

**AEPs to Letters** AEPs to the first letter stimulus differed for Hits and Misses under LVW, with a greater N1 for Hits compared to Misses ($p$ = 0.0416; Figure 2 d). This comparison includes 18 subjects (1 was excluded due to low trial numbers in the Miss condition). Electrodes FC3, C3, CP3 and CP4 were excluded from the average plot as no significant effects were found at these electrode sites for this comparison. The average is instead comprised of the remaining 8 sites in the AEP montage (left; Figure 1).

## P300 to Letter 1

P300 to Letter 1 (Table 2) differed for Hits by visual load, with a greater P300 for low load ($p$ = 0.0048; Figure 3 a). The P300 also differed in response to Letter 1 in Hits vs. Miss trials (LVW), with Hits resulting in a larger P300 ($p$ = 0.0248). Figure 3 b shows grand average ERPs for this effect, averaged across 18 subjects due to low trial numbers for one subject in the Miss condition. There was also a difference in P300 amplitude between Hits and CR under LVW in response to Letter 1, with a larger P300 for Hits compared to CRs ($p$ = 0.0224; Figure 3 c).




(a) L1, Hit by Visual Load.

(b) L1, Hit v. Miss, Low Load.

(c) L1, Hit v. CR, Low Load.

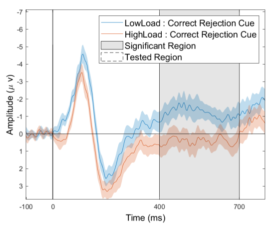
Figure 3: Significant differences in P300 response. Colored shading represents standard error of the mean. CR: Correct Rejection; L1: Letter 1

## Reorienting Negativity

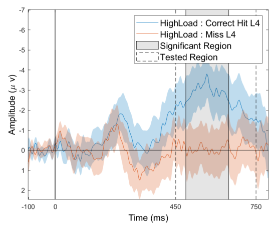
We investigated the RON in response to Cues, Letter 1 stimuli and Letter 4 stimuli within and across workload and behavioral outcome. Table 2 depicts the results of all comparisons. For all RON ERP plots, electrode C4 and CP4 are excluded from the montage (left; Figure 1) due to a lack of significant effects at these sites. This leaves 11 electrode sites remaining. Only significant results are summarized in the text below.

**RON to Cues** The RON to cues resulting in CRs differed as a function of workload, with a greater RON for LVW ($p$ = 0.0008; Figure 4 a).

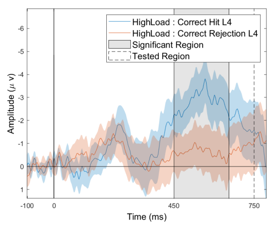
**RON to Letters** Under HVW, the RON to Letter 4 stimuli on trials resulting in Hits was stronger than for trials resulting in Misses ($p$ = 0.0424; Figure 4 b). Also under HVW, the RON differed for Hits and CRs, with a greater RON in response to Letter 4 for Hits ($p$ = 0.0288; Figure 4 c).



(a) Cues, CR by Visual Load.



(b) L4, Hit v. Miss, High Load.



(c) L4, Hit v. CR, High Load.

Figure 4: Significant differences in Reorienting Negativity response. Colored shading represents standard error of the mean. CR: Correct Rejection; L4: Letter 4

# Discussion

## Auditory Evoked Potentials

**AEPs to Cues** Investigating differences in AEPs as a function of visual load is informative for understanding how early auditory processing and attention is influenced by cross-modal load. Specifically, under LVW, cues leading to CRs elicited a stronger N1 response than cues leading to Hits or Misses. These effects were not predicted a priori, and particularly the stronger N1 to CRs vs. Hits was unexpected. However, it is likely that this difference in N1 amplitude is driven by physical stimulus characteristics. The N1 to Target Cues (Hit and Miss Cues), a higher frequency tone, is expected to be smaller than the N1 to Non-Target Cues (CR Cues), a lower frequency tone (Wunderlich & Cone-Wesson, 2001). What is perhaps most interesting about these findings is the fact that this difference is abolished under HVW. N1 response did not differ as a function of Cue type and behavioral outcome for any comparison investigated. Difference in number of trials retained as a function of workload was negligible, which eliminates "insufficient trials" as an explanation for the absence of an effect. We interpret the absence of this effect as a sign of degraded early sensory processing, where, under crossmodal load, the representation of the sound's physical characteristics is less veridical.

Further evidence is provided by the AEP response to CR cues, which differed when comparing HVW and LVW. Following the same logic as above, this difference may reflect a more accurate representation of the Non Target Cue's physical characteristics under LVW. Workload related effects may also be contributing to this amplitude difference. The CR cue is a task irrelevant auditory stimulus in that no decryption is required if letters are preceded by this cue type. Previous research has found the AEP to task irrelevant sounds to be predictive of spare capacity, across multiple task types, with lower workload leading to a stronger ERP (Xu et al., 2020; Wang et al., 2023). This difference could also be a replication of this effect. The conclusions outlined here do rely on the interpretation of some negative results (absence of effects under HVW when comparing cue types), and unfortunately we do lack the control condition of the ERP response to each tone type before training and without any additional workload. Due to these caveats, additional follow-on work is necessary to draw strong conclusions. However, the most logical interpretation is one where auditory attention and processing is degraded and yields a less accurate representation of physical stimulus characteristics under HVW.

**AEPs to Letter 1** AEP differences in response to Letter stimuli were also informative, which were present for just one condition. Under LVW, the N1 was stronger to Letter 1 Stimuli resulting in Hits compared to Misses. This difference suggests that poor sensory encoding and early auditory attentional allocation may contribute to errors (Misses) under LVW. Interestingly, this difference was absent for HVW. These effects hint at strategy differences across load, where under LVW, Misses are associated with changes in early auditory processing. Under HVW the picture is likely more heterogeneous and trial dependent, with no clear changes to early auditory processing evident at the grand average level.

## P300

**P300 to Letter 1** Consistent with previous literature, the P300 effect to task relevant stimuli requiring a response (in our case, Letter 1 following a Target Cue) was greater under low workload compared to high (Dehais, Duprès, et al., 2019). However, this was the case only for Letter 1 stimuli resulting in Correct Hits. Differences in the P300 were also present within LVW, where the P300 following Letter 1 was greater for trials resulting in Hits compared to Misses and compared to CRs. The P300 is therefore diagnostic for LVW, and could serve as a predictive tool to anticipate if targets will be missed or responded to appropriately. The P300 was similarly weak to Letter 1 stimuli across behavioral outcomes within HVW, and so its usefulness as a diagnostic cue under high levels of load is questionable.

## RON

**RON to Cues** The RON was investigated to understand attentional reallocation after receiving task-relevant information. The RON in response to cues differed for LVW and HVW but only for CRs. The RON was stronger under LVW, indicating that participants reoriented attention after accurately processing Non Target Cues. The interpretation of this effect is not straightforward, and relies on assumptions about differences in attentional state that were revealed by our AEP analyses. Under LVW, attention is allocated to the auditory modality as evidenced by the strong AEP in response to CR Cues. However, this cue signals that the upcoming letters will not require decryption and are therefore irrelevant. Attention then reorients to the visual task. This is reflected by the presence of a strong the RON after CR Cues in LVW. Conversely, in HVW, the weaker N1 to CR Cues suggests a failure to fully attend to the auditory modality. This would explain the absence of evidence of reorientation of attention back to the visual task after receiving a Non-Target Cue, since attention never shifted away from the visual task in the first place.

**RON to Letter 4** The RON is also useful for understanding the vocal decryption process under HVW. Specifically, the RON was stronger in response to Letter 4 on Hit, compared to CR or Miss, trials. This may mark the deployment of attentional resources to rehearse and decrypt Letter 1 and 4 that occurs only for Hit trials, while no RON was detected for Miss or CR trials (no decryption required). A very interesting conclusion can be drawn if considering where effects are absent. A pattern similar to that found for HVW was visible for LVW, bit did not reach statistical significance.

The less consistent difference in RON under LVW for these comparisons may provide neurophysiological evidence of a strategy described anecdotally by our participants. A number of participants claimed to be silently practicing the decryption task under LVW, as a self-directed means of gaining additional experience with the task. The RON effects align with these claims, where despite a lack of behavioral response for Correct Rejection trials, the cognitive operations related to the decryption task may occur on some trials, making the differences between these conditions less consistent from trial to trial. This leads to an interesting interpretation for Miss trials. Since the RON does not differ between Hits and Misses for LVW, one could argue that reorientation of attention to the decryption task is somewhat similar across the conditions despite the opposite behavioral outcomes. This suggests that other stages of processing are responsible for the failure to respond that occurs during Miss trials under LVW.

### Limitations and Future Directions

The impact of these data are narrowed by a few limitations of a within-subjects design, and careful control over the experimental factors and stimuli. First, the target and distractor tones were always pulled from the same frequency bands, limiting some conclusions about differences between the target and distractor tones. Future work could tease apart the effect of target v. distractor and high v. low frequency tones using a between-subjects factor where an equal number of subjects are assigned a high or low frequency target. Related, no EEG data were collected during training sessions; therefore, no comparisons of tones could be made prior to learning their task specific assignment (target/distractor). Future work could easily collect these data to form a representation for the underlying neural signal of each frequency band, fostering some interesting conclusions about shifts from baseline and potential individual differences. Further, although data were collected with a within-subjects design and allowed for individual-level assessment, data were only assessed at the group-level due to the low target prevalence in the auditory task (20% of trials), which was then unevenly split between Hits and Misses. Due to the difficult nature of these tasks, six 2-hour sessions were required (4 training); future work could consider alternative designs to collect enough trials to assess individual-differences, either by extending the number of sessions or altering the paradigm to lower the degree of training necessary to adequately understand the task(s) and settle on a multi-task strategy. Lastly, the target cues in this task were simple tones, where a real-world indicator that one should listen is typically a callsign or nickname; similarly, the decryption task included letters and knowledge of the English alphabet. These simple stimuli produced less variable neural responses, but future work should consider including more realistic stimuli to test the generalizability of our findings.

Future research should consider alternative/additional experimental choices, and a few are listed above. In addition, we plan to utilize additional assessment tools to evaluate the neural activity associated with trial types of interest. For instance, time-frequency analyses (TFA) over Hit, Miss, and CR trials can provide a larger picture about the activity in specific frequency bands for trial type, where the relative power spectral density to some bands are sensitive to cognitive state changes, such as mental workload (Ke et al., 2021). TFA may complement ERP analyses to provide early predictive information about cognitive state, performance, or identification of which task(s) are being attended to at any given time (Fox, Ugolini, & Houpt, 2022).

# Conclusion

In conclusion, we found degraded early processing (reduced AEP) under HVW, indicating a less accurate representation of physical stimulus features. Further, P300 amplitude to L1 was predictive of trial accuracy when spare capacity was high (LVW), providing evidence that insufficient letter processing is a likely cause for incorrect responses. Lastly, we found RON after L4 (HVW) led to a Hit (v. misses/CR), suggesting errors occur prior to decryption. Our data provide insight to the underlying processing of crossmodal tasks with varying difficulty and may inform future research questions.

## Acknowledgments

The views expressed are those of the author and do not necessarily reflect the official policy or position of the Department of the Air Force, the Department of Defense, or the U.S. government. This research was supported in part by an appointment to the Department of Defense (DOD) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the DOD. ORISE is managed by ORAU under DOE contract number DE-SC0014664. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DOD, DOE, or ORAU/ORISE.Distribution A. Approved for public release; distribution unlimited. AFRL-2024-0534; Cleared 30 Jan 2024.


## References


Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological psychology*, *80*(1), 23–34.

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural mr images of the brain. *IEEE transactions on medical imaging*, *18*(1), 32–42. doi: 10.1109/42.750253

Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R. N., & Lotte, F. (2019). Monitoring pilot's mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions. *Sensors*, *19*(6), 1324.

Dehais, F., Rida, I., Roy, R. N., Iversen, J., Mullen, T., & Callan, D. (2019). A pbci to predict attentional error before it happens in real flight conditions. In *2019 ieee international conference on systems, man and cybernetics (smc)* (pp. 4155–4160).

Dehais, F., Roy, R. N., & Scannella, S. (2019a). Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural brain research*, *360*, 51–59.

Dehais, F., Roy, R. N., & Scannella, S. (2019b). Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural brain research*, *360*, 51–59.

Delorme, A. (2023). Eeg is better left alone. *Scientific reports*, *13*(1), 2372.

Donchin, E., Ritter, W., McCallum, W. C., et al. (1978). Cognitive psychophysiology: The endogenous components of the erp. *Event-related brain potentials in man*, *349*, 411.

Fitzroy, A. B., Ugolini, M., Munoz, M., Zobel, B. H., Sherwood, M., & Sanders, L. D. (2020). Attention modulates early auditory processing at a real cocktail party. *Language, Cognition and Neuroscience*, *35*(6), 712–728.

Fox, E. L., Bowers, G., Capiola, A., & Stephenson, A. (2023). The design of an ostensible human teammate. In *Proceedings of the hawaii international conference on system sciences.*

Fox, E. L., Ugolini, M., & Houpt, J. W. (2022). Predictions of task using neural modeling. *Frontiers in Neuroergonomics*, *3*, 1007673.

Ghani, U., Signal, N., Niazi, I. K., & Taylor, D. (2020). Erp based measures of cognitive workload: A review. *Neuroscience & Biobehavioral Reviews*, *118*, 18–26.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields i: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725.

Gumenyuk, V., Korzyukov, O., Tapaskar, N., Wagner, M., Larson, C. R., & Hammer, M. J. (2023). Deficiency in re-orienting of attention in adults with attention-deficit hyperactivity disorder. *Clinical EEG and Neuroscience*, *54*(2), 141–150.

Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(1), 1.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual review of psychology*, *34*(1), 33–61.

Hölle, D., Blum, S., Kissner, S., Debener, S., & Bleichner, M. G. (2022). Real-time audio processing of real-life soundscapes for eeg analysis: Erps based on natural sound onsets. *Frontiers in Neuroergonomics*, *3*, 1.

Ke, Y., Jiang, T., Liu, S., Cao, Y., Jiao, X., Jiang, J., & Ming, D. (2021). Cross-task consistency of electroencephalography-based mental workload indicators: comparisons between power spectral density and task-irrelevant auditory event-related potentials. *Frontiers in Neuroscience*, *15*, 703139.

Liebherr, M., Corcoran, A. W., Alday, P. M., Coussens, S., Bellan, V., Howlett, C. A., … Bornkessel-Schlesewsky, I. (2021). Eeg and behavioral correlates of attentional processing while walking and navigating naturalistic environments. *Scientific reports*, *11*(1), 22325.

Linguistic Data Consortium. (2002). *Center for Spoken Language Understanding (CSLU) ISOLET Version 1.3.* `http://www.cslu.ogi.edu`. (University of Pennsylvania)

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73*, 1780–1789.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Makeig, S., Bell, A., Jung, T.-P., & Sejnowski, T. J. (1995). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, *8*.

Manly, B. (1997). Randomization, bootstrap and monte carlo methods. *Biology, 2nd edn. Chapman and Hall, London*.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, *164*(1), 177–190.

Mullen, T. R., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., ... Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry eeg. *IEEE transactions on biomedical engineering*, *62*(11), 2553–2567.

Munka, L., & Berti, S. (2006). Examining task-dependencies of different attentional processes as reflected in the p3a and reorienting negativity components of the human event-related brain potential. *Neuroscience letters*, *396*(3), 177–181.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive psychology*, *7*(1), 44–64.

Oray, S., Lu, Z.-L., & Dawson, M. E. (2002). Modification of sudden onset auditory erp by involuntary attention to visual stimuli. *International Journal of Psychophysiology*, *43*(3), 213–224.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197.

Polich, J. (2007). Updating p300: an integrative theory of p3a and p3b. *Clinical neurophysiology*, *118*(10), 2128–2148.

Pritchard, W. S. (1981). Psychophysiology of p300. *Psychological bulletin*, *89*(3), 506.

Sanders, L. D., & Astheimer, L. B. (2008). Temporally selective attention modulates early perceptual processing: Event-related potential evidence. *Perception & psychophysics*, *70*(4), 732–742.

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of meg/eeg data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335.

Schröger, E., & Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *Neuroreport*, *9*(15), 3355–3358.

Solís-Marcos, I., & Kircher, K. (2019). Event-related potentials as indices of mental workload while using an in-vehicle information system. *Cognition, Technology & Work*, *21*, 55–67. doi: 10.1007/s10111-018-0485-z

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.

Suzuki, J., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, *116*, 918-933.

Ventry, I., Woods, R., Rubin, M., & Hill, W. (1971). Most comfortable loudness for pure tones, noise, and speech. *The Journal of the Acoustical Society of America*, *49*, 1805-1813.

Wang, S., Gu, H., Yao, Q., Yang, C., Li, X., & Ouyang, G. (2023). Task-independent auditory probes reveal changes in mental workload during simulated quadrotor uav training. *Health Information Science and Systems*, *11*(1), 12.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, *3*(2), 159–177.

Wunderlich, J. L., & Cone-Wesson, B. K. (2001). Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *The Journal of the Acoustical Society of America*, *109*(4), 1526–1537.

Xu, J., Ke, Y., Liu, S., Song, X., Xu, C., Zhou, G., & Ming, D. (2020). Task-irrelevant auditory event-related potentials as mental workload indicators: a between-task comparison study. In *2020 42nd annual international conference of the ieee engineering in medicine & biology society (embc)* (pp. 3216–3219).