The Impact of Spatialized Auditory Distraction on Visual Search Performance

A thesis submitted by

Sydni M. Nadler

in partial fulfillment of the requirements for the degree of Master of Science

in

Psychology

Tufts University

May 2025

Advisors:

Holly A. Taylor, PhD & Nathan Ward, PhD

Abstract

Visual search — the ability to find targets amongst distractors — is a skill used daily in personal and professional settings, alike. As an attention-guided skill, however, visual search is susceptible to both visual distractions and distractions in other modalities such as audition. This susceptibility is particularly notable if the distraction is spatialized (i.e., appearing to come from a specific direction), though this remains untested for auditory interruptions. With this in mind, the present study sought to investigate how spatialized auditory distractors impact visual search performance as a function of task requirements. Employing a modified version of the additional singleton task (Theeuwes, 1991, 1992) with an interrupting auditory identification task, data were collected from Tufts University community members across two experiments. Experiment 1 asked participants to identify a visual target and where an interrupting sound originated, while Experiment 2 asked participants to complete the same visual search task and identify what an interrupting sound represents. A 2x2 within-subjects design was used to compare visual search performance at the experiment-level during distraction present and distraction absent conditions, with distractors appearing in both visual and auditory modalities. Additional analyses then compared visual search performance across experiments to assess the influence of spatialized (rather than non-spatialized) auditory interruptions. Results indicate that visual search trials with auditory distractors produced significantly longer response times and lower target identification accuracy than trials without auditory distractors, regardless of spatial properties. This finding, congruent with our expectations for this study, serves to highlight the potential for cross-modal distractions impacting visual search performance in scenarios like driving and travel safety where distractibility may contribute to risky, real-life situations.

Keywords: Attention control, cross-modal distraction, spatial cognition, visual search

Table of Contents

List of Tables

List of Figures

The Impact of Spatialized Auditory Distraction on Visual Search Performance

Driving a motor vehicle is an attentionally demanding and highly visual task (Ma et al., 2020; Trick et al., 2004; Walker & Trick, 2018) that millions engage in daily (Steinbach & Tefft, 2023). Between checking the color of the traffic light ahead, scanning the upcoming sidewalks for pedestrians that may enter the road, and monitoring the distance between one's own car and the car in front of it, safe driving requires that a driver's visual attention remains focused on the road ahead. But what happens if the driver hears an emergency vehicle siren? Typically, drivers will take their eyes off the road and scan the traffic around themselves to determine where this sound is coming from and how to appropriately respond (i.e., continue driving or pulling over to the side). This is an instinctive response to the situation; humans naturally turn their heads when encountering spatialized sounds in order to change interaural time differences (i.e., the difference in time it takes for sound to reach each ear) and to gather visual cues, both of which improve localization accuracy (Lewald et al., 2000). While beneficial in identifying the direction from which the siren is coming, the tendency to reorient the head can also interrupt the visual vigilance necessary to drive safely. The driver may now know that they must pull over to allow an emergency vehicle through but may not have seen the distance between themselves and the car in front of them closing rapidly. Given that nearly 3,500 motor vehicle fatalities result from distracted driving in the United States each year (NCSA, 2023), it is essential to understand how spatialized auditory distractors impact visual search abilities. The purpose of this study, then, is to explore this relationship with the intent of informing policy and innovation to reduce these all-too-common tragedies.

Visual search — the process of identifying targets amongst distractors — is an attention-driven skill used in daily life (Wolfe, 1994; Wolfe & Horowitz, 2017). The study of this skill has been successfully applied both to academic and real-world contexts, informing research both in the lab and in settings that impact the everyday individual. Studies on radiology (Gandomkar & Mello-Thoms, 2019), driving (Chapman & Underwood, 1998), and airport security (Biggs et al., 2017), amongst others, elucidate the importance of understanding the mechanisms of visual search in professional and personal tasks. For instance, Travel Safety Administration (TSA) agents are familiar to travelers within the United States and represent a sector of professional visual searchers. In inspecting baggage scans to ensure no illicit items are being brought on board aircraft, TSA agents practice visual search on a daily basis as a requirement of their job. Studies of these professional searchers, investigating potential correlates to visual search performance like personality (e.g., Biggs et al., 2017) or rapid search assessments proficiency (e.g., Mitroff et al., 2018), serve to inform hiring practices and the broader academic literature. Thus, the study of visual search holds value not just to the field of psychology, but to the general public whose daily lives benefit from the insights gleaned from visual search literature.

Attention and Distraction

When investigating the mechanisms underlying visual search, whether in academic or applied settings, two prominent models emerge: top-down and bottom-up attention. These models offer complementary perspectives on how our attention is guided and influenced during the search process. Top-down attention, also referred to as goal-oriented attention, is user-driven and reflects an individual’s desire to locate a specific target in their immediate environment (Folk et al., 1992; Folk & Remington, 1998; Wolfe, 1994). This may manifest, for instance, as an individual intentionally searching a drawer for their car keys. Attention, though, is not an infallible resource, and rich visual environments often contain salient items that may draw visual attention regardless of intention — this is bottom-up attention (Schreij et al., 2008; Theeuwes, 1991, 1992; Treisman & Gelade, 1980). When searching the aforementioned drawer, objects of unique brightness, color, or motion, such as a tube of red lipstick rolling around, have the potential of capturing visual attention, despite the knowledge that the intended target is a set of car keys. These types of objects, those of salience or uniqueness, serve as distractors to the main goal of a given search, involuntarily pulling visual attention away from the user's goals. While both attentional models have a role in describing the complex processing required to interact with a dynamic environment, bottom-up features, drawing attentional resources away from intentional searching, may serve as an attentional impediment when a focused search is desired.

However, visual distractors represent just one category of potential interruptions individuals may encounter in real-world settings. Research has begun to assess how auditory distractors affect visual search performance. When experiencing ongoing background noises that do not require user interaction, such as a pink noise tone, top-down visual attention remains strong (Mandal, Liesefeld, et al., 2024; Mandal, Röer, et al., 2024). Dynamics change, though, if that auditory distraction becomes task-relevant (e.g., an aspect of a recall test), as goal-directed attention becomes impaired when interaction is involved (Richard et al., 2002). Actively attending to both sights and sounds in one's environment is essential for navigating a multisensory world, but doing so may compromise the amount of focus that can be given to each modality.

Moreover, *spatialized* auditory distractors exacerbate the diversion of visual attention, further underscoring the intricate interplay between auditory and visual processing mechanisms. When the elevation of sound is varied, not only is visual search performance negatively impacted, but it can also be spatially guided by the auditory input (Santangelo et al., 2007; Spence & Driver, 1997a). While this attention guidance is adaptive and allows individuals to respond to the multi-sensory world around them, it also highlights the ease at which visual attention can be drawn towards, and even influenced by, spatialized sounds. This ease may be helpful in some situations (e.g., when one hears their keys jingle while searching a cluttered drawer for them). However, in other situations (e.g., driving, radiology, additional medical settings), the susceptibility of visual attention to cross-modal distraction may prove to be dangerous (Balint et al., 2014; Moher, 2020; Williams & Drew, 2017). It is with these potentials in mind that the study of spatialized distraction during visual search must be approached.

Neural Correlates of Audiovisual Integration

The audiovisual connection, particularly the processing required to identify and localize sounds and visual search targets, alike, rests in the abilities of two key perceptual streams in the brain. The dorsal pathway, known as the "where" pathway, is responsible for processing the location or motion of perceptual input, while the ventral pathway, known as the "what" pathway, processes the information necessary to identify a given input. Both the visual (Ingle et al., 1982; Milner & Goodale, 2006; Mishkin et al., 1983) and auditory (Rauschecker & Tian, 2000; Tian et al., 2001) processing systems have been noted to have independent versions of these two pathways. Further neuroimaging evidence suggests functional separation of these two pathways in the cross-modal, audiovisual processing system (Sestieri et al., 2006). While these two systems can and often do operate independently from one another, certain areas of the brain have been found where these systems combine their efforts. Namely, the lateral intraparietal sulcus (LIP) (Bushara et al., 1999; Macaluso et al., 2000; Macaluso & Driver, 2001) and superior temporal sulcus (STS) (Beauchamp et al., 2004; Calvert et al., 2000) have been noted as working to *integrate* information obtained cross-modally. How the modality-respective processing streams work when *interacting*, rather than integrating, though, is another story.

It is somewhat contested whether these pathways compete for attentional resources when receiving information that does not warrant integration from different modalities, such as an unexpected, cross-modal distraction. Evidence suggests that, in some circumstances, cross-modal auditory attention capture does not occur, thus preserving visual attention (Koelewijn et al., 2009b; Santangelo & Spence, 2007b). Contradictory evidence suggests that auditory attention capture may occur involuntarily, resulting in compromised visual attention (Koelewijn et al., 2009a; Lunn et al., 2019; Mazza et al., 2007). This debate has a role to play in various high-stakes, real-world situations and, therefore, requires further investigation.

Aims of Experiment 1

Despite the body of literature exploring the impact of auditory distraction on visual search performance, it is not well understood how spatialized auditory distractors during visual search might impact performance. Using both classic and novel cognitive tasks, the current study asks the question: “How do lateralized, spatially localized auditory distractors impact visual search performance as a function of task requirements and individual differences?” Given the established literature, it is expected that visual search conditions that expose participants to distraction in both visual and auditory forms will produce the greatest negative impact on visual search performance. In contrast, conditions without any distraction will produce the most accurate and fastest performance.

This study will include both visual and auditory modalities, having participants complete a visual task that may be interrupted by a spatialized auditory location identification task. While the two tasks are unrelated to one another, the methodology may unintentionally cue participants to search a returning visual array in the location from which they have just heard a spatialized auditory clip. This cueing effect has been noted in several studies concerning the effects of spatialized audio on a vision-based task (Klein, 1977; Spence & Driver, 1997b). As a result, an exploratory analysis will be conducted of the visual search trials that follow an interrupting audio task to assess if there is an error rate or reaction time difference between congruent (i.e., the final sound is perceived from the *right* side and the returning visual target is also present on the *right* side of the search array) and incongruent (i.e., the final sound is perceived from the *right* side and the returning visual target is present on the *left* side of the search array) trials. It is hypothesized that incongruent trials will result in slower response times and less accurate performance on the returning visual search task.

The influence of individual differences during visual search may also play a role in search performance (e.g. Biggs et al., 2017), and in the context of this study, individual differences in attentional processing become vital to understand. As such, participants of this study will complete an Attentional Control Scale questionnaire (Derryberry & Reed, 2002), which assesses individual differences in how one controls one's attention. The Scale produces an overall score as well as scores for two sub-scales — attention-shifting and attention-focusing. It is hypothesized that those who report stronger tendencies for attention-shifting will produce quicker reaction times to visual search trials that include visual or auditory distractors over those who do not have a pronounced attentional control strategy or those who report stronger attention-focusing, given their self-reported ease at shifting between tasks.

Experiment 1 Methods

Design

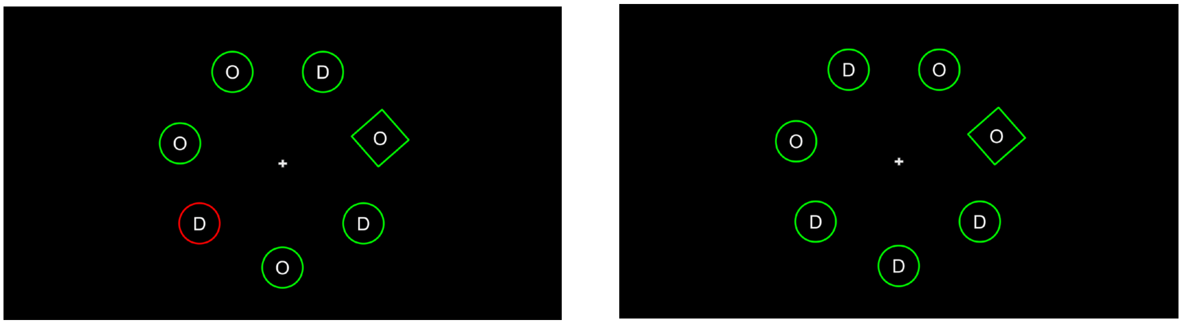
This study utilized a 2 x 2 within-subjects design, with factors of visual distraction (present or absent) and auditory distraction (present or absent). Distraction conditions occurred randomly with a 50% probability in both modalities. Participants were asked to complete 240 trials of the visual search task — 25% without any distractors, 25% with only a visual distractor in the search array, 25% with only an interrupting auditory localization task, and 25% with both a visual distractor and an interrupting auditory localization task (trial timing of each condition may be viewed in the figures in Appendix B). The methodology and analyses of this project were pre-registered and may be accessed at https://osf.io/yrjns.

Materials

Visual Search Performance

Figure 1.1

Example additional singleton task search arrays



Note. Subjects searched for the diamond, either with a distractor (left) or without (right)

The additional singleton task (Theeuwes, 1991, 1992) is a standard laboratory task that measures attention capture during visual search by asking participants to search for a shape singleton among non-target items. On some trials, a salient distractor (the additional singleton) is present to initially “capture” visual attention away from the target.

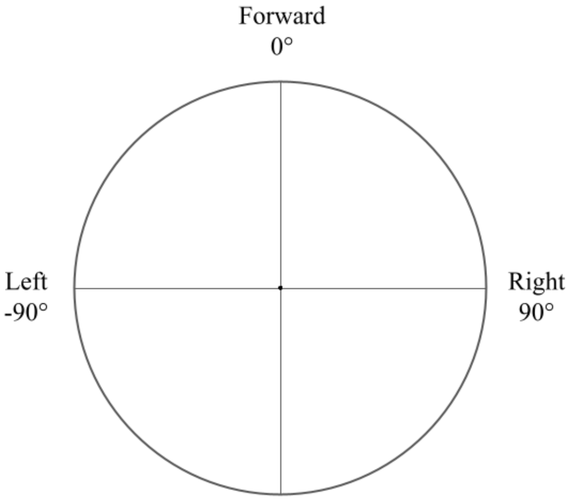
In our study, participants were asked to identify which letter, either a “D” or an “O,” was located inside a green diamond (the shape singleton target) using the computer keyboard (See Figure 1.1). To reduce the effect that gross motor movements between the “D” and “O” on a typical keyboard, stickers were placed on the “D” and “A” keys denoting “D” and “O,” respectively. Each additional singleton task array had seven items, and both the target and additional singleton changed location during each trial. The additional singleton, a red circle, appeared in 50% of additional singleton task trials, and the letters within the array items changed randomly in each trial. Visual search abilities were measured by response time to identify the target letter in each trial and error rate in identifying that letter.

Auditory Spatialization Performance

Three sequentially-presented tones (85.08 dB) were played over headphones during the interrupting auditory localization task. These tones each appeared to originate from a different direction, either in front of the participant, to the left of the participant, or to the right of the participant (See Figure 1.2). This was accomplished by manipulating the strength each tone had coming through both the left and right headphone channels (i.e., the left tone had a stronger signal coming from the left side of the headphone set than the right). Each tone was 1 second in length, not ramped, and had a frequency of 440 Hz (Boersma & Weenink, 2024).

Figure 1.2

Direction key for sound presentation



During each auditory interruption, three randomly selected tones would play and participants were asked to identify the direction from which each tone appeared to come (i.e., one tone would play, the participant would respond, then the next tone would play). Tone direction was identified by participants on the computer keyboard with the arrow keys; a left-originating tone was indicated with the left arrow, a right-originating tone was indicated with the right arrow, and a front-originating tone was indicated with the up arrow. These interruptions were fully randomized in terms of how often they appeared, so long as they occurred in 50% of the visual search trials. When an auditory interruption *did* occur during an additional singleton task trial, it would close that search trial after 75 milliseconds of being visible. The participants were then prompted to respond to three tones before the additional singleton task search array that had been presented immediately prior to the auditory interruption would return to the screen. Participants would then continue with their search. A reduced-speed example of this trial type can be viewed at https://osf.io/gvf45. Auditory spatialization performance was measured by response time to each tone and error rate in identifying the tone direction. Analyses pertaining to performance in this task may be viewed in Appendix C.

Attentional Control

The 20-item self-report Attentional Control Scale (Derryberry & Reed, 2002) was employed to assess individual differences in attention abilities. The scale asks participants to assess on a Likert scale of 1 (almost never) to 4 (always) how strongly the series of questions relates to their patterns of working and concentrating. The ASC produces an overall score as well as scores for two sub-scales — attention-shifting and attention-focusing.

To determine a participant’s dominant attention tendency, their subscale scores were averaged according to the maximum value of each subscale, as the two subscales contained an unequal number of questions. The participant's attention-focusing average was then subtracted from their attention-shifting average. This yielded either a positive or negative value, where a positive value indicated more attention-shifting tendencies and a negative value indicated more attention-focusing tendencies. This value was then zero-centered for use in analyses.

Procedure

This study was approved by the Tufts University Institutional Review Board. Participants arrived at their scheduled slot and were welcomed into the Tufts University Spatial Cognition lab space, where they actively provided informed consent. Upon consent, they were guided to a computer and given headphones to adjust to a comfortable size.

The behavioral tasks were prepared using the PsychoPy Builder (Peirce et al., 2019) and presented to participants using that interface. Visual stimuli were displayed on a 15-inch 1920 x 1080 resolution screen approximately 18 inches from the participant. Head position was not experimentally controlled.

Instructions for both tasks were included in the opening of the experiment. Participants were told to place their left ring finger on the key marked with a sticker "O" (the "A" key) and their left pointer finger on the key marked with a sticker "D" (the "D" key). For the right hand, participants were asked to place their pointer finger on the left arrow key (to indicate a "left" tone), their middle finger on the up arrow key (to indicate a "center" tone), and their ring finger on the right arrow key (to indicate a "right" tone). It was noted that participants should fixate on the cross in the center of each array. The opportunity to address any questions a participant may have had regarding how to complete the experiment was then presented. Once these questions were answered, participants previewed the three tone directions and completed a series of practice trials before moving on to the main task. This was completed in a single block of 240 trials, each separated by an inter-trial interval of 100 ms. At the conclusion of the experiment, the researcher would direct participants to Qualtrics to complete the Attentional Control Scale and a demographics questionnaire. Participants were then thanked for their time and compensated accordingly. The typical experiment session took approximately 15 minutes in total.

Participants

Based on the sample sizes of previous studies concerning visual and/or auditory attention capture (e.g., $N_{\text{Exp 1}}$ = 28 and $N_{\text{Exp 2}}$ = 24 in Folk & Remington, 1998; $N_{\text{Exp 1}}$ = 27 and $N_{\text{Exp 2}}$ = 24 Santangelo et al., 2007; $N_{\text{Per Exp}}$ = 8-16 in both Theeuwes, 1991, 1992), we aimed to recruit approximately 40 participants for this study. Data were ultimately collected from 37 Tufts University community members. Recruitment occurred through the online study management website SONA, offering participants the option to complete the study for course credit in an introductory psychology class or for monetary compensation. Study sessions were approximately one half-hour in length, and participants were compensated either 0.5 course credit or $10 per half-hour. Eligibility requirements for this study included being 18 to 30 years of age, having normal or corrected-to-normal vision, having normal color vision, having normal or corrected-to-normal hearing, and being fluent in English. One participant was excluded for exceeding the required age range (N = 36).

A number of additional exclusion criteria were set in place for the behavioral and survey data that were collected. It was determined that participants who scored above a 20% error rate on either the visual search or auditory identification task would be removed from the task-respective analyses. Data from 2 participants was removed from all analyses for this reason, leaving a final sample of 34 (M_{Age} = 19.91, *SD* = 2.69 years. See additional demographic information in Appendix A). This criteria also meant that data from an additional 2 participants were removed only from the auditory task analyses. On a within-subjects level, data from visual search trials where the response time was quicker than 100ms or longer than 10 seconds were to be excluded. It was required that participants complete at least 50% of their trials in each of the four visual search conditions. Similarly, participants had to have answered at least half of the questions for each of the two Attention Control Scale subscales to have their data included in the exploratory analyses pertaining to attention style. No data was removed from analyses for either of these reasons.

Data Preparation and Analysis

All data cleaning procedures and analyses were performed using R Statistical Software (v4.3.3; R Core Team, 2024). Generalized mixed linear models, the primary statistical analysis performed, were assessed with the R package glmmTMB (Brooks et al., 2017) and tabulated with sjPlot (Lüdecke, 2023).

Generalized linear mixed models (GLMMs) were estimated to assess the influence visual and auditory distractors may have on the rate of error committed within-subjects during the additional singleton task, using both the participant and the stimuli as random effects. Model fit was determined with AIC values, marginal and conditional R^2 values, and root mean standard error values (RMSE). Intra-class correlation coefficient (ICC) values of the random-intercept-only models revealed that 17% of the variance in the data can be attributed to random effects. These models employed a binomial distribution family and a logit link function. The GLMMs used to evaluate the effects of the visual and auditory distraction conditions on additional singleton task response times was determined in the same way as that for error rates (See Table 1.1), though these models used a Gamma distribution family and a log link function.

Table 1.1

Visual search GLMMs

Model Formula	*AIC*	*Marginal R^2/ Conditional R^2*	*RMSE*
Error Rate			
Visual Search Error Rate ~ 1 + (1|Participant)	2131.29	0.00 / 0.124	0.168
Visual Search Error Rate ~ 1 + (1|Trial Image)	2168.11	0.00 / 0.051	0.169
Visual Search Error Rate ~ Visual Distraction * Auditory Distraction + (1|Participant) + (1|Trial Image)	2121.98	0.01 / 0.181	0.167
Response Time			
Visual Search Response Time ~ 1 + (1|Participant)	1424.81	0.00 / 0.217	0.327
Visual Search Response Time ~ 1 + (1|Trial Image)	3128.69	0.00 / 0.036	0.355
Visual Search Response Time ~ Visual Distraction*Auditory Distraction + (1|Participant) + (1|Trial Image)	-673.43	0.156 / 0.402	0.295
Cueing			
Visual Search [Error Rate or Response Time] ~ Visual Distraction * Congruency + (1|Participant) + (1|Trial Image)			
Attention Tendency			
Visual Search [Error Rate or Response Time] ~ Visual Distraction * Attention Tendency Score + (1|Participant) + (1|Trial Image)			

Experiment 1 Results

Table 1.2

Means and standard errors by trial type and variable

	Response Time (sec)		*Error Rate (%)*	
Trial Type	*Avg*	*SE*	*Avg*	*SE*
Additional Singleton Task				
VD Absent + AD Absent	0.776	0.006	2.70	0.359
VD Present + AD Absent	0.817	0.007	2.25	0.329
VD Absent + AD Present	1.028	0.008	3.58	0.411
VD Present + AD Present	1.053	0.008	3.28	0.395
Auditory Identification Task				
Left	0.365	0.007	3.61	0.294
Center	0.413	0.007	8.40	0.44
Right	0.374	0.007	7.43	0.42

Note: VD = Visual Distractor, AD = Auditory Distractor.

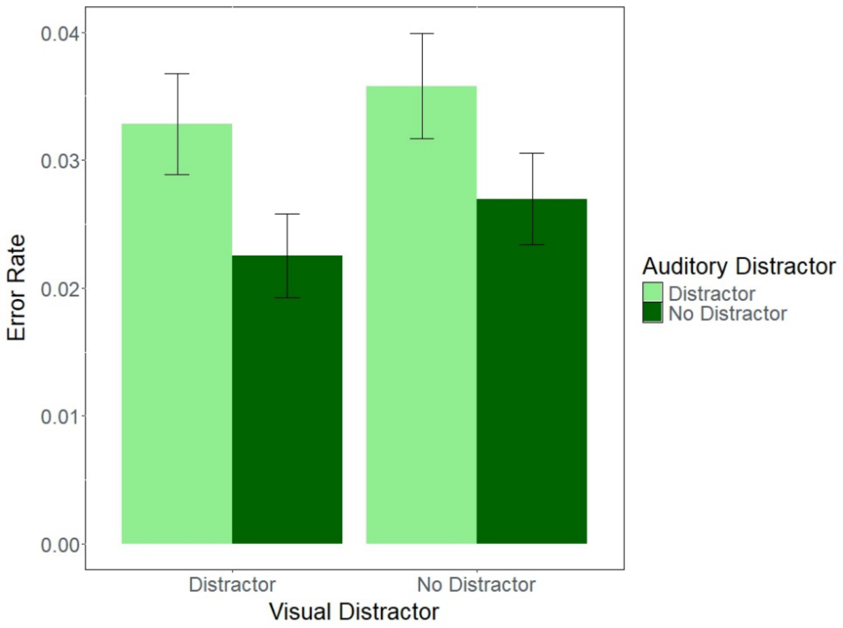
Primary Analyses

Error Rate.

We observed a main effect of auditory distraction on visual search error indicating that participants were 1.20 times more likely to incorrectly identify the additional singleton task target letter when their search resumed following an auditory interruption (β = 1.20, SE = 0.08, z(8160) = 2.69, *p* = 0.007) (See Figure 1.3). No significant effect was seen for visual distractors (β = 0.91, SE = 0.11, z(8160) = -0.83, *p* = 0.41) or for the interaction between visual and auditory distractors (β = 1.03, SE = 0.07, z(8160) = 0.43, *p* = 0.67) on error rates.

Figure 1.3

Visual search error rate



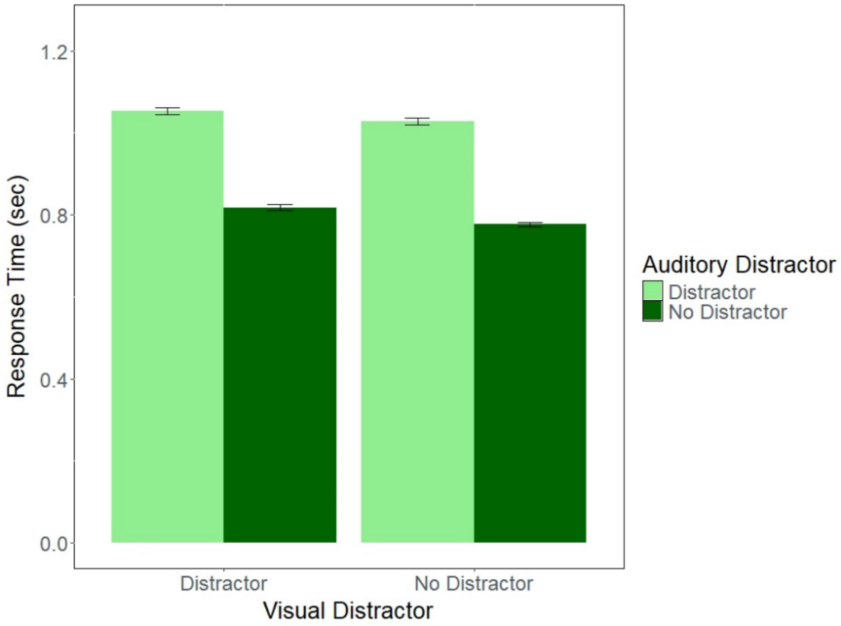
Note. Error bars are +/− SE.

Response Time.

Similar to the error rate findings, we observed a main effect of auditory distraction (β = 1.14, SE = 0.00, z(8160) = 45.46, *p* < 0.001) with no statistically significant effect of visual distraction (β = 1.02, SE = 0.01, z(8160) = 1.33, *p* = 0.19). Trials interrupted by the secondary auditory task contributed to longer reaction times in the additional singleton task, suggesting that auditory distraction, but not visual distraction, individually contributes to reduced visual search performance in both measures. However, when the two variables interact (β = 0.99, SE = 0.00, z(8160) = -2.22, p = 0.026), it becomes apparent that the presence of both visual and auditory distractors led to participants requiring more time to respond to the trial (compared to when only one form of distractor is present; see Figure 1.4).

Figure 1.4

Visual search response time



Note. Error bars are +/− SE.

Exploratory Analyses

Cueing.

Given the established findings on the cueing effect that spatialized audio may have on visual tasks (Klein, 1977; Spence & Driver, 1997b), we chose to assess this effect in our own experiment. Additional singleton task trials were deemed to be "congruent" if the third of the three tones in an interrupting auditory task appeared to originate from the same direction that the target of the interrupted additional singleton task stimulus was located (e.g., the third tone appears to come from the right and the target of the reappearing additional singleton task image is located on the right of the screen). An "incongruent" trial meant that the third tone appeared to originate from a different direction than where the additional singleton task target was located (e.g., the third tone appears to come from the right and the target of the reappearing additional singleton task image is located on the *left* of the screen). We chose to focus on the third tone specifically because it was temporally closest to the resumption of the visual search task. Neither error rate (β = 0.70, SE = 0.18, z(4080) = -1.39, *p* = 0.16) nor response time (β = 0.99, SE = 0.02, z(4080) = 1.62, *p* = 0.11) indicated any statistically significant effects, giving us confidence that participants were not cued to search in a particular direction by the last tone played during the auditory task.

Attention Tendency.

Individual differences in attention have been linked to visual search performance (Adamo et al., 2017), so it was of interest to assess the potential relationship between self-reported attention tendencies (i.e., attention-shifting or attention-focusing) and additional singleton task performance in this study. Attention scores ranged from –0.199 (most attention-focusing) to 0.207 (most attention-shifting). While no main effects were found when attention style was added into the error rate model, we identified an interaction between *visual* distractor presence and attention tendency as well as an interaction between *auditory* distractor presence and attention tendency (See Table 1.3).

This first interaction indicates that, when compared to attention-focusers, the error rate of attention-shifters was more impacted by visual distractors during the AST (see Figure 1.5). The difference scores in error rate between visual distractor conditions for attention-shifters and attention-focusers further emphasizes this determination; at the most extreme attention-focusing score (-0.199), the difference in predicted error rate for visual distractor versus no visual distractor conditions equates to approximately 0.0151, while the difference at the most extreme attention-shifting score (0.207) equates to approximately 0.0159. The second interaction suggests the same conclusion: that attention-shifters were more affected by varying distractor conditions, specifically when auditory distractors were present versus absent (See Figure 1.6). When auditory distractors were present, as compared to when they were absent, the error rate of attention-shifters (at the most extreme attention-shifting score) increased by an average of 0.017. The error rate of attention-focusers (at the most attention-focusing score), in contrast, differed by only 0.0071 across the two auditory distractor conditions.

When attention tendency was added into the response time model, we also observed an interaction between auditory distractor presence and attention tendency (See Table 1.4). Attention tendencies impacted how quickly participants responded to the additional singleton task such that, when encountering a trial that included an auditory distractor, the response times of those with attention-focusing tendencies were less impacted (i.e., less slowed down) by the presence of this distractor than those with stronger attention-shifting tendencies. The difference in response time between the two auditory distractor conditions at the most extreme attention-focusing score was approximately 0.193 seconds, compared to the 0.276 second difference across conditions at the most extreme attention-shifting score.

Table 1.3

Visual search error rates and attention tendency

Predictors	*Odds Ratio*	*std. Error*	*CI*	*Statistic*	*p*
(Intercept)	0.02	0.00	0.02 – 0.03	-21.00	**<0.001*****
Visual Distractor	0.94	0.11	0.74 – 1.18	-0.56	0.578
Auditory Distractor	1.17	0.08	1.02 – 1.34	2.22	**0.026***
Attention-Shifting	0.37	0.57	0.02 – 7.29	-0.65	0.516
Visual Distractor × Auditory Distractor	1.06	0.07	0.92 – 1.21	0.82	0.413
Visual Distractor × Attention-Shifting	0.17	0.13	0.04 – 0.74	-2.37	**0.018***
Auditory Distractor × Attention-Shifting	4.36	3.26	1.01 – 18.84	1.97	**0.049***
Visual Distractor × Auditory Distractor × Attention-Shifting	0.64	0.47	0.15 – 2.74	-0.60	0.545
Marginal R^2 / Conditional R^2	0.026 / 0.193				

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Figure 1.5

Interaction between attention tendency and visual distractor presence on error rate

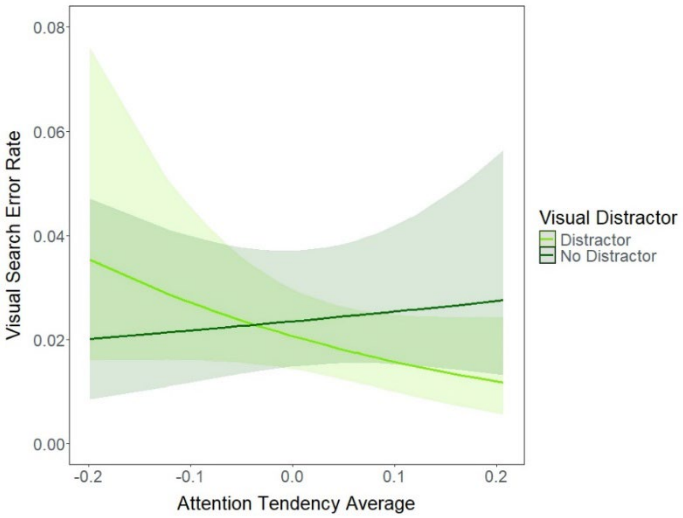
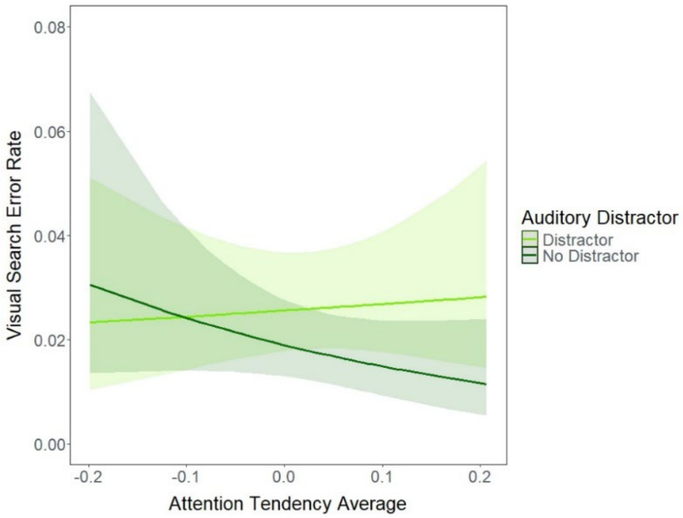


Figure 1.6

Interaction between attention tendency and auditory distractor presence on error rate



Note. Attention tendency averages < 0 = attention-focusing tendencies and averages > 0 = attention-shifting tendencies

Table 1.4

Visual search response time and attention tendency

Predictors	*Estimates*	*std. Error*	*CI*	*Statistic*	*p*
(Intercept)	0.90	0.03	0.85 – 0.95	-3.70	**<0.001*****
Visual Distractor	1.02	0.01	0.99 – 1.05	1.32	0.185
Auditory Distractor	1.14	0.00	1.13 – 1.15	45.48	**<0.001*****
Attention-Shifting	1.44	0.40	0.84 – 2.48	1.32	0.188
Visual Distractor × Auditory Distractor	0.99	0.00	0.99 – 1.00	-2.21	**0.027***
Visual Distractor × Attention-Shifting	0.98	0.03	0.92 – 1.04	-0.70	0.486
Auditory Distractor × Attention-Shifting	1.07	0.03	1.00 – 1.13	2.11	**0.035***
Visual Distractor × Auditory Distractor × Attention-Shifting	1.04	0.03	0.98 – 1.10	1.26	0.209
Marginal R^2 / Conditional R^2	0.167 / 0.403				

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Experiment 1 Discussion

The goal of this study was to explore the impact of spatialized auditory distractors on visual search performance, maintaining an interest in the role individual differences in attention may have on this relationship. We had initially hypothesized that the study condition that exposed participants to both visual and auditory distractors would result in the most impeded performance, while the condition with neither distractor would result in the quickest and most accurate search. This hypothesis was confirmed by the interaction we identified between auditory distractor presence and visual distractor presence in the case of visual search response time; the presence of distractors in both modalities resulted in significantly longer response times during the additional singleton task. Outside of this interaction, our findings indicate that error rate and response time were *only* impacted by auditory distractors, *not* visual distractors. Regardless of the presence of a visual distractor, error rates and response times increased when the additional singleton task was interrupted by the spatialized auditory identification task. This finding is explored in depth in our Limitations, as it mirrors the results of other studies investigating the nature of top-down and bottom-up visual attention.

We also conducted several exploratory analyses examining the role additional factors might have when visual search interacts with auditory distraction effects. Unlike previous studies (Klein, 1977; Spence & Driver, 1997b), we did not observe any influence of auditory cueing on error rate or response time on interrupted visual search trials. Though this contradicts our hypothesis that incongruent trials (i.e., trials where the final sound originates from the right and the returning visual target is on the left side of the search array) would result in slower response times in identifying the visual target or lead to more errors in identification, this is a result we are pleased to have seen. We can more confidently attribute changes in performance to the intended experimental manipulations rather than the tone direction in the spatialized auditory task knowing that cueing does not appear to mediate interrupted additional singleton task outcomes.

The other factor we were curious to investigate was the potential moderating influence that individual differences in attention style could have in these outcome measures. Knowing that individual differences could influence search performance (Biggs et al., 2017), particularly those concerning attention (Adamo et al., 2017), participants completed the Attentional Control Scale as part of this experiment. As we hypothesized, attention tendency *did* play a role in additional singleton task response times. Our results indicate that those who reported attention-focusing tendencies, compared to attention-shifting tendencies, were less impacted by the presence of an auditory distractor in terms of their response times. The error rates of attention-focusers were also less impacted by the presence of distractors in either modality than the error rates of their attention-shifting peers. Collectively, attention-focusing appeared to play a protective role in search performance when encountering both visual and auditory distractors. Attention-focusing tendencies, therefore, could be a consideration when making personnel selections in professional settings that necessitate strong visual search abilities, like airport security.

In all, it was clear through our primary and exploratory visual search analyses that auditory distractors have a debilitating effect on search response time and accuracy, and two potential reasons for this stand out. The first of these phenomena is the competition that may emerge between the auditory and visual systems over attentional resources during sensory interaction. It has been suggested that perceptually demanding tasks may allocate attentional resources to one modality over the other (Santangelo & Spence, 2007a), supported by event-related potential evidence that attention may serve a modulatory role in multisensory processing (Talsma & Waldorff, 2005). Our study is one of several that has noted impaired visual task performance as a result of auditory distraction (Koelewijn et al., 2009a; Lunn et al., 2019). In the case of the present study, the error rates of the auditory task compared to the error rates of the additional singleton task may indicate that the auditory task was the more demanding of the two (See Table 1.2). The cognitive resources granted to this secondary task could have left less available for the returning additional singleton task trial, contributing to the longer response times and lower accuracy observed in the interrupted visual trials than the uninterrupted ones.

This, then, paves the way for the second potential cause of our visual search results: "distraction hangover," or resumption lag. Originally associated with distracted driving, this phenomenon explains the impaired primary task performance that may emerge following an interruption or distraction by a secondary task due to the cognitive load attributed to task switching (Borowsky et al., 2016; Snider et al., 2023). This detriment can be exacerbated by the complexity of the primary task and the mental fatigue induced by the secondary task (Chen et al., 2022). As we observed impaired accuracy and response times during the additional singleton task, specifically when the trial was interrupted by the secondary auditory task, we believe that participants could have been experiencing a "distraction hangover" when returning to the additional singleton task. The impairments in performance that resulted could also have been amplified by the nature of our novel task, switching both between modalities (e.g., visual and auditory) and the type of information that needed to be responded to (e.g., letter identity and spatial location). Spatial interruptions in particular have been noted to impair resumption performance more than non-spatial interruptions (Ratwani & Trafton, 2008), providing more evidence to suggest that resumption lag could explain our findings.

Limitations

This study was informative in a number of ways but still encountered limitations in its execution and findings. The primary methodological limitation we encountered was the use of headphones to present the auditory task to participants. Doing so limited the auditory directions that could be reliably distinguished. We were unable to include a tone that appeared to originate from behind the participant because, by using the headphones, we eliminated the spectral cues that typically come from head-related transfer functions; essentially, because the "behind" tone did not have to travel around the participant's head and pinna, it became indistinguishable from the “in front of” tone. However, as the nature of this study focuses on attention and distraction, we found that headphones provided a great advantage in reducing the possibility that participants may be distracted by audition not relevant to the study (e.g., students talking in the hallway outside of the laboratory). As such, the decision was made to forgo a fourth directional tone despite its relevance to real-world situations where auditory distractions may originate from behind an individual while conducting a daily visual search.

Perhaps one of the most interesting outcomes of this study is the nature of the results from the additional singleton task during uninterrupted trials. This task was chosen specifically for its propensity to elicit bottom-up visual attention capture by including a salient additional stimulus to the visual search array (Theeuwes, 1991, 1992), but we did not observe this finding. Rather, the main effects we observed on both visual search response time and error rate only pertained to the presence of auditory distraction, not visual. While we had expected that auditory distraction would impact search performance, this finding is contrary to our initial hypothesis that having both visual and auditory distractors would yield the worst performance. However, the indication that visual singletons did not impact visual search performance measures is not unprecedented and may be due to the perseverance of top-down attentional control.

Given that the targets and distractors in our study were two different colors (green and red, respectively) within an array of green stimuli, participants may have been able to “filter out” singleton features that were irrelevant to them finding the target object and identifying the letter contained within. This “top-down control setting” (Folk et al., 1992; Folk & Annett, 1994) would allow participants to stay directed on the goal of the task and ignore the red circle distractor as it did not contain any features similar to the green diamond target. Essentially, participants would have the opportunity to learn to suppress the red distractor, regardless of salience, and focus exclusively on the green diamond (Cosman & Vecera, 2013; Gaspelin & Luck, 2018; Turatto & Galfano, 2001; Zehetleitner et al., 2012). When specifically considering color singletons, this finding has been observed numerous times in relation to both general attention (Folk et al., 1992; Folk & Remington, 1998) and visual search (Folk & Annett, 1994; Jonides & Yantis, 1988). There has been extensive debate on the effect a discontinuous color distractor may have on maintaining goal-directed attention and search; we had relied on findings pertaining to bottom-up attention capture, but the result we observed was also a known possibility. This is no fault of the additional singleton task itself but rather a preview into the dynamic nature of attentional control when experiencing salient but ultimately goal-irrelevant stimuli.

Summary

This study affirms the idea that spatialized auditory distractors have a distinct, negative impact on the ability to successfully carry out visual search. Whether for radiologists, airport security officers, or the average driver, this is a finding that can prove troublesome in many real-world scenarios. As such, it is our goal to spur innovations in both policy and technological design to minimize the negative consequences that spatialized auditory distractors may have. These could manifest in a variety of ways, including new recommended hiring procedures (e.g., Biggs et al., 2017) or multi-sensory devices to refocus drivers' attention (e.g., Ho et al., 2005), among others. The results of this experiment underscore the importance of understanding the complex interplay between auditory and visual stimuli and the distractibility effects they might have in daily life.

Experiment 2

Extensive research has explored how spatial cues potentially influence visual attention, though fewer have focused on spatial auditory cues. Our first experiment contributed to that less-explored sector of the cross-modal literature; informed by existing studies (e.g., Koelewijn et al., 2009a; Lunn et al., 2019; Mazza et al., 2007), we hypothesized (and later observed) that visual search abilities would be compromised by spatialized auditory distractors. However, we also had the exploratory hypothesis that the interrupting spatialized auditory task could unintentionally cue participants to resume the visual search task by looking in the direction from which the last interrupting tone originated. Unlike other studies in this area though (e.g., Klein, 1977; Spence & Driver, 1997b), we did not observe this spatial cueing effect. Questions, then, remain as to whether the negative effects on visual search we observed in Experiment 1 were due to the spatial aspect of the auditory interruption task, or perhaps for other reasons. As introduced in Experiment 1, it is possible that modality-switching or task-switching effects, rather than the spatial aspect of the interrupting task, could explain our findings. Experiment 2 was conducted as a means of verifying whether the spatial nature of the tones was responsible for the impaired visual search performance we observed in Experiment 1.

Experiment 1 demonstrated the potential for cross-modal distractors — stimuli in one modality that may draw attention away from a task in another modality — to negatively impact performance in a primary cognitive task; specifically, we examined the role *spatialized auditory* distractors played during *visual* search and how switching between two modalities impacted performance in the primary visual task. While we hypothesized that the spatial nature of our auditory stimuli was the reason for impaired performance in the primary task, guided by the literature discussed above, we also acknowledged that modality-switching could compromise primary task performance on its own (e.g., the auditory recall task during a primary visual task employed by Richard et al., 2002). Experiment 2 was pursued to determine if Experiment 1's findings can be attributed to the *spatial* tones we used rather than the fact that we used *any* variety of auditory stimuli during a primary visual task.

The impaired performance experienced by participants of Experiment 1 could also be explained by the “distraction hangover” — or resumption lag — effect; returning to a primary task following an interrupting secondary task could produce costs (e.g., longer response times) associated with task-switching. During the interrupting secondary task of Experiment 1, participants were asked to identify *where* they heard each of the three interrupting tones originate before quickly returning to identify *what* letter they observed in visual target. The rapid switch between localization and identification tasks asked of participants could have produced resumption costs. This explanation is bolstered by the notion that secondary tasks with spatial elements may induce greater performance costs than non-spatial secondary tasks (Ratwani & Trafton, 2008). However, even without spatial elements, the increased cognitive load taken on during task-switching can also disrupt primary task performance (Borowsky et al., 2016; Snider et al., 2023). By varying whether our secondary task uses spatialized auditory stimuli or non-spatialized auditory stimuli (Experiment 1 vs Experiment 2), we aimed to provide additional evidence that spatial distractors were the reason for significant visual search performance impairments in our first study.

The individual suppositions regarding modality and task-switching as potential alternative explanations for our Experiment 1 findings are supported by the literature on task-switching across modalities. The literature on this topic reveals that performance impairments in the form of longer response times and higher error rates may be produced when experimental tasks vary between the auditory and visual modalities (Hunt & Kingstone, 2004; Kreutzfeldt et al., 2015; Murray et al., 2009). Mirroring the findings of Experiment 1, Kreutzfeldt and colleagues (2015) observed performance costs when participants experienced sequential trials that varied in modality (i.e., sequential trials that, for example, went from the visual to the auditory modality led to longer response times than sequential trials that stayed within the same modality). They specifically noted that these costs were exacerbated in trials that went from the auditory to the visual modality, similar to our findings for our interrupted trials. Reaction time costs when switching modality were also found by Hunt and Kingstone (2004), and, importantly, this study highlighted that the cost of both task *and* modality switching was greater than when either switch type was observed on its own.

Contrasting these findings, Murray, DeSantis, Thut, and Wylie (2009) saw less prominent performance impairments when sequential trial modalities switched (i.e., sequential trials that switched from the visual to the auditory modality or vice versa) than when they repeated (i.e., back-to-back trials of either visual or auditory stimuli), and even less costs when both task and modality switched. This result was attributed to the functionally separate neural pathways required for cross-modal switches reducing the extent to which a task performed in one modality affected the activity in the brain areas associated with a different task in the other modality. It is argued that the reduction of competition between the networks underlying the performance of the two tasks decreased the costs associated with switching between task and modality. However, this study, like others in this research space (e.g., Kreutzfeldt et al., 2015) employed cueing to indicate to participants which modality or which task to allocate attention to. Without task or modality cues, though, responses may be more reactive than prepared (i.e., responding directly to the stimulus, rather than having advanced notions of the variety of stimulus and expecting it), ultimately contributing to task performance impairments (Sandhu & Dyson, 2013).

Unfortunately, it is the latter scenario that is more realistic to what individuals experience in their daily or professional lives. Radiologists, for example, may not get advanced warning that their visual scan of a mammogram will be suddenly interrupted by a phone ringing that they must respond to. It is with this in mind that it was necessary to verify that our Experiment 1 findings were due to the spatial aspect of the auditory stimuli presented, rather than task and modality switching costs, with the present study.

Additionally, semantically-relevant (“characteristic”) auditory stimuli have been studied in relation to how they impact visual attention (i.e., sounds that conceptually relate to the visual stimuli they are presented alongside). Like spatialized auditory cues, these have been noted to enhance visual search capabilities (Chen & Spence, 2011; Iordanescu et al., 2008; Knoeferle et al., 2016), but little has been studied on the effects of non-semantically-relevant (“uncharacteristic”), non-spatial sounds might have on this same process. This is of particular interest as people may encounter any manner of auditory distraction (i.e., spatialized, non-spatialized, characteristic, or uncharacteristic) in daily life, and understanding how individuals respond to these interruptions during visual search is essential. We explored this in Experiment 2 by changing the auditory stimuli used from spatialized tones to naturalistic, inanimate sounds. Altering these stimuli then allowed us to identify if the visual search impairments seen in Experiment 1 were due to the spatialized nature of the tones or simply due to the distracting nature of any auditory stimulus during visual search.

Given our existing findings and the majority consensus in the literature on cross-modal task-switching, we hypothesized that visual search performance would remain impaired as a result of modality switching, but perhaps not as significantly as it did in Experiment 1. This was predicted as both the visual search and interrupting auditory tasks in Experiment 2 ask participants to respond to the *identity* of the stimuli, rather than the *identity* of the visual target and then the *location* of the auditory (similar to the conclusions of Hunt & Kingstone's 2004 paper). Removing the task-switching element of varying between "what" and "where," we believed, would reduce the likelihood of a "distraction hangover" and thereby reduce general impairment during the visual search task. This is further supported by the literature on spatialized cross-modal distraction, which argues that non-spatial auditory distractors would not impair visual search performance as significantly as spatialized auditory distractors would.

Our exploratory hypotheses are tied to our Experiment 1 findings. The performance advantages experienced by attention-focusers over attention-shifters in Experiment 1 during distraction conditions were not specifically tied to the spatialized nature of the tones used. Consequently, we hypothesized that attention-focusers would continue to adapt better to distractor conditions than their attention-shifting peers in Experiment 2.

It was also predicted that we would observe stronger visual distractor effects for this experiment than in Experiment 1. Not observing an effect of a discontinuous color distractor, like we did in Experiment 1, has been well-documented in the literature on top-down and bottom-up influences in visual attention, particularly in the context of distraction. However, we also felt it important to take into consideration the experience of our participants during the study when examining this result. During the Experiment 1 debriefing phase, participants were asked how they felt the visual distractor, in the form of the red circle, and auditory distractor, in the form of the interrupting task, impacted their performance on the visual search task, if at all. While some participants responded that they did not feel that these distractors impacted their performance at all, 44.1% of participants (N = 15) indicated that the red circle specifically was able to be "tuned out" as trials progressed, was less distracting than the auditory distractor, or was not a distractor at all. Comparatively, 17.6% of participants (N = 6) indicated that the visual distractor was consistently distracting throughout the trials. One comment noted that the red of the visual distractor was more similar to the black background of the task than the green of the target singleton and remaining array, prompting an investigation into the perceptual distance between the colors used for this task.

This color difference is typically identified using Euclidean distances. We employed the International Commission on Illumination’s CIEDE2000 formula to calculate the ΔE values (i.e., the color difference values) between each of the colors seen in the additional singleton task arrays (See Table 1.5) (Luo et al., 2001). As indicated by the ΔE values for these colors, the observation that the shade of red used was closer to the black of the background than the green was correct. The original literature on the additional singleton task specifically notes that the salience of a distractor is an essential factor in producing the effect we had intended in using this task (Theeuwes, 1991). As such, discovering that there could be a difference in color salience between the distractor and target colors from the background is noteworthy when considering our initial findings. Experiment 2 was adjusted to use a grey background, as outlined in Table 2.1, to reduce the influence color difference may have on attention during the AST.

To further bolster the salience of the visual distractors in Experiment 2, the AST array objects were color counterbalanced. This change emerged in response to our Experiment 1 findings, where we did not observe the presence of our visual distractors having a main effect on error rates or response time (i.e., we did not observe the visual attention capture typical of the AST). Unlike Experiment 1, which exclusively had green targets and red distractors, Experiment 2 also had arrays with red targets and green distractors. When attempting to capture visual attention, it is *crucial* to maintain the saliency (i.e., uniqueness) of the distractor object (Adam et al., 2021; Adam & Serences, 2021); counterbalancing the colors of targets and distractors in the visual search arrays assists in subverting the learned expectation that one particular color (red, in the case of Experiment 1) should be ignored (Cosman & Vecera, 2013; Turatto & Galfano, 2001; Zehetleitner et al., 2012). This change to the stimuli should, then, preserve the distractibility of the distractor objects and contribute to greater visual attention capture.

Table 2.1

ΔE values between additional singleton task stimuli colors

Color	Feature	*ΔE* from red	*ΔE* from green	*ΔE* from black
Red (#ff0000)	Distractor	0.0	87.18	50.41
Green (#00ff00)	Target	87.18	0.0	87.88
Black (#000000)	Background	50.41	87.88	0.0
Grey (#999999)†	-	*34.4*	*34.89*	*49.86*

Note. † Background color for follow-up studies.

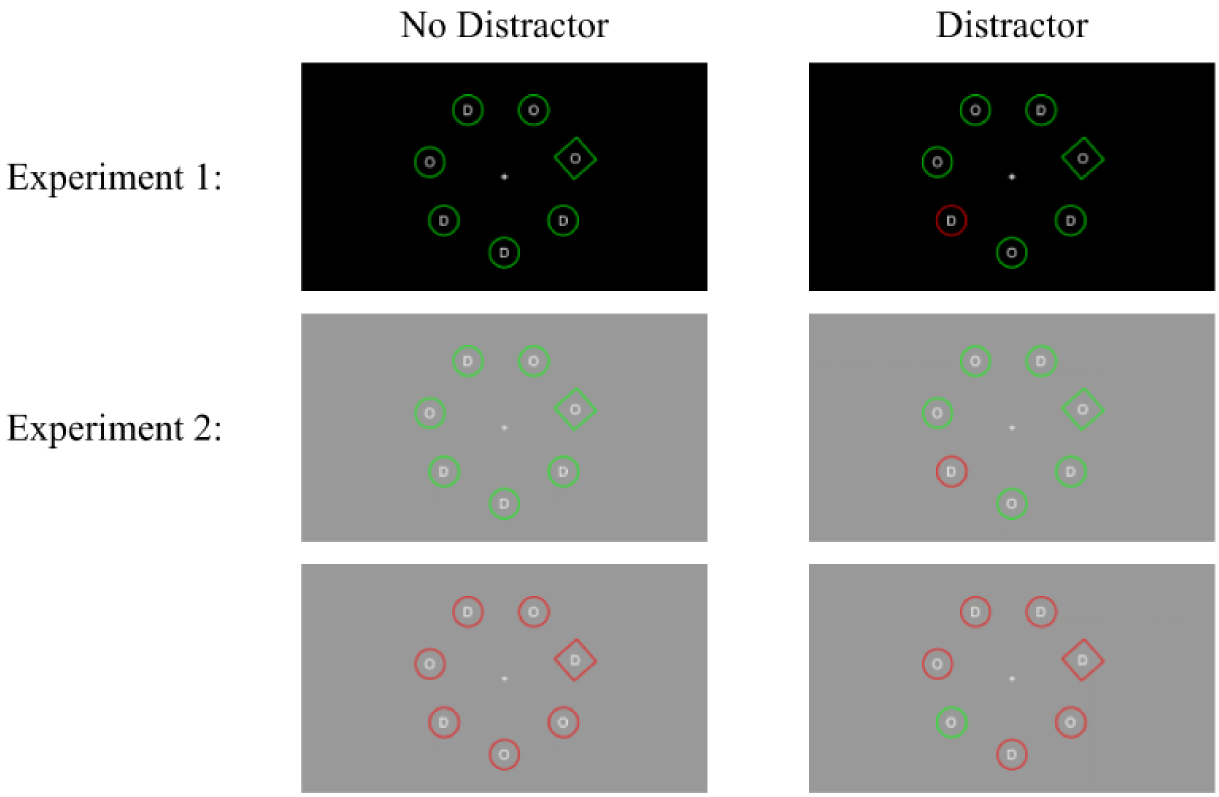
Experiment 2 Methods

Design

This study continued to utilize the 2 x 2 within-subjects design used in Experiment 1 to assess visual search performance during interrupting auditory distractions. Participants were asked to complete 240 trials of the novel additional singleton task — 25% without any distractors, 25% with only a visual distractor in the search array, 25% with only an interrupting auditory localization task, and 25% with both a visual distractor and an interrupting auditory localization task (trial timing of each condition may be viewed in the figures in Appendix B). As such, visual and auditory distractors randomly occurred in 50% of all trials.

Figure 2.1

Visual search arrays in Experiment 1 compared to Experiment 2.



Materials and Procedure

Visual Search Performance

The array background color was changed from black to grey in this experiment to ensure both red and green stimuli stood out against the background equally (See Figure 2.1). The stimuli were also counterbalanced such that 50% of visual search arrays had *green targets* with *red distractors*, and the other 50% of search arrays had *red targets* with *green distractors*. This was done to preserve the distractibility of the additional singleton in distractor-present trials, as participants could not learn to specifically "tune out" one color over the other.

Auditory Identification Performance

While the procedural aspects of the interrupting auditory task were consistent across both experiments, the nature of Experiment 2 dictates that new sounds were used. Rather than playing three spatialized tones, this experiment required participants to identify three naturalistic sounds: a whistle blow, a car horn, or a ringing phone. These sounds each have an alerting quality to them, ensuring one sound did not draw more attention than the others. It should be noted that these audio clips were not exactly 440 Hz and 85.08 dB, as the tones in Experiment 1 were, but were similar in pitch and intensity to the tones and to each other so that perception should not be impacted in a noticeable way. Each audio clip was 1 second in length; additional details about these sounds may be seen in Table 2.2. These sounds played with equal strength from both headphone channels to mitigate any spatial properties of the clips.

As with Experiment 1, participants of Experiment 2 were asked to identify three sequentially-presented sounds during each auditory interruption trial using the computer keyboard arrow keys. Stickers with visual cues pertaining to each of the three potential sounds were placed on the arrow keys (i.e., rather than being cued to use the left arrow key when encountering a “left” tone with the visual of the left arrow on that key, participants were cued to use the left arrow key when encountering the whistle sound with the visual of a whistle on that key; See Figure 2.2). These stickers were applied to the arrow keys to reduce the memory load of participants attempting to remember which key aligned with which sound.

Table 2.2

Average pitch and intensity of Experiment 2 sounds

Sound	Pitch (Hz)	Intensity (dB)
Whistle	476.87	84.09
Car Horn	432.05	84.25
Phone Ring	457.84	81.51

Note. Values derived via analysis in Praat (v6.4.13; Boersma & Weenink, 2024)

Figure 2.2

Response keys and corresponding visual cues for Experiment 2 sounds

Attentional Control

The attentional control methodology remained consistent across both experiments.

Procedure

Experiment 2 was approved by the Tufts University Institutional Review Board and the experimental sessions were nearly identical in procedure to that of Experiment 1. They differed only in the instructions given to participants regarding the auditory response keys. For this study, the left arrow key was now covered with a whistle sticker, the up arrow key with a steering wheel sticker, and the right arrow key with a phone sticker (See Figure 2.2). The object on each sticker corresponded to the correct responding key for each of the three sounds, and participants were instructed to place their right pointer finger on the whistle, their right middle finger on the steering wheel, and their right ring finger on the phone. They then previewed each sound, completed practice trials, and were given the opportunity to ask additional questions before proceeding to the main task. Each experiment session took approximately 15 minutes in total.

Participants

Participants (N = 36) were again recruited through the SONA website, using the same sample size determination as Experiment 1. Each experimental session lasted approximately 15-30 minutes, and participants were given the option to participate for course credit or for monetary compensation at a rate of either 0.5 course credit or $10 per half-hour. Eligibility requirements for this study were consistent with those in Experiment 1.

Data exclusion criteria also remained consistent across experiments. Initially, no participants were removed from the data set due to error rates or not completing the Attentional Control Scale, and only three trial-level response times met our criteria. However, upon further inspection of response time and accuracy data by trial type (i.e., no distractors, visual distractor only, auditory distractor only, and both distractors), data from 2 participants were removed as outliers. Our final sample size was 34 (M_{Age} = 18.41, *SD* = 0.66 years. See additional demographic information in Appendix A).

Table 2.3

Experiment 2 Visual search GLMMs

Model Formula	*AIC*	*Marginal R^2/ Conditional R^2*	*RMSE*
Error Rate			
Visual Search Error Rate ~ 1 + (1|Participant)	1800.41	0.00 / 0.16	0.152
Visual Search Error Rate ~ 1 + (1|Trial Image)	1859.45	0.00 / 0.015	0.153
Visual Search Error Rate ~ Visual Distraction * Auditory Distraction + (1|Participant) + (1|Trial Image)	1799.70	0.01 / 0.18	0.152
Response Time			
Visual Search Response Time ~ 1 + (1|Participant)	3824.47	0.00 / 0.034	0.408
Visual Search Response Time ~ 1 + (1|Trial Image)	4782.47	0.00 / 0.006	0.433
Visual Search Response Time ~ Visual Distraction*Auditory Distraction + (1|Participant) + (1|Trial Image)	2133.21	0.044 / 0.094	0.374
Attention Tendency			
Visual Search [Error Rate or Response Time] ~ Visual Distraction * Attention Tendency Score + (1|Participant) + (1|Trial Image)			
Stimuli Color			
Visual Search [Error Rate or Response Time] ~			

Visual Distraction * Auditory Distraction *
Primary Color of Array Stimuli + (1|Participant)
+ (1|Trial Image)

Trial Number

Visual Search [Error Rate or Response Time] ~
Primary Color of Array Stimuli * Trial Number
+ (1|Participant) + (1|Trial Image)

Data Preparation and Analysis

Like Experiment 1, generalized mixed linear models estimated with the R package glmmTMB (Brooks et al., 2017) were used to assess the impact of the non-spatialized interrupting auditory distractions on visual search performance in Experiment 2. Participant and stimuli were used as random effects for the within-subject analyses. Models used to evaluate error rates employed a binomial distribution family and a logit link function while response time data better aligned with a lognormal family model and a log link function. Model fit was determined by a convergence of AIC values, marginal and conditional R^2 values, and RMSE values (See Table 2.3). ICC values of the random-intercept-only models revealed that 21% of the variance in the data for Experiment 2 error rates can be attributed to random effects. Exploratory analyses were run to evaluate the potential role individual differences in attention may play during visual search, as were analyses assessing the role that primary stimulus color could have.

Experiment 2 Results

Table 2.4

Means and standard errors by trial type and variable

	Response Time (sec)		*Error Rate (%)*	
Trial Type	*Avg*	*SE*	*Avg*	*SE*
Additional Singleton Task				
VD Absent + AD Absent	0.828	0.007	1.86	0.299
VD Present + AD Absent	0.912	0.008	2.06	0.315
VD Absent + AD Present	1.102	0.009	2.94	0.374
VD Present + AD Present	1.188	0.011	2.8	0.365
Auditory Identification Task				
Whistle	0.42	0.006	1.70	0.206
Car Horn	0.445	0.007	0.81	0.14
Phone	0.464	0.007	1.88	0.215

Note: VD = Visual Distractor, AD = Auditory Distractor.

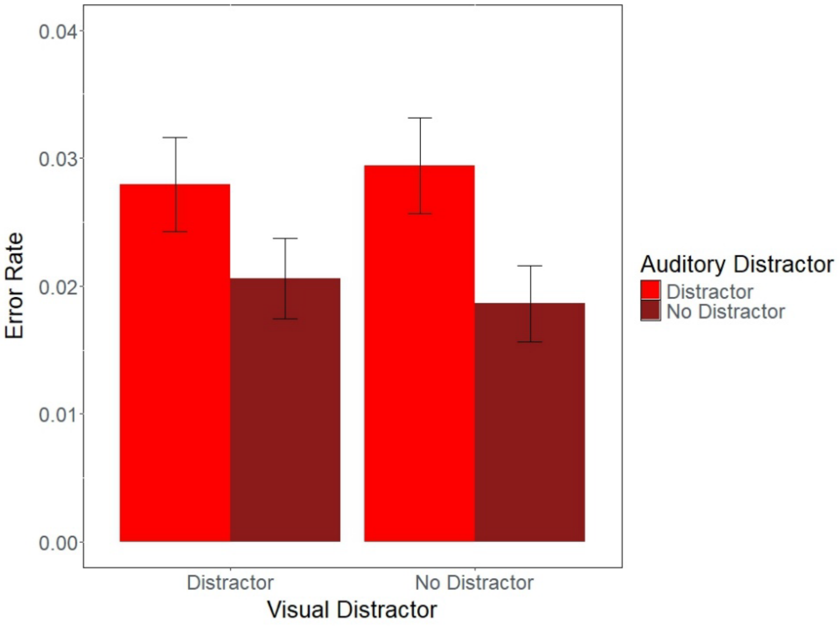
Primary Analyses

Error Rate.

We observed a main effect of auditory distraction on visual search error indicating that participants were 1.22 times more likely to incorrectly identify the AST target letter when their search resumed following an auditory interruption (β = 1.22, SE = 0.09, z(8158) = 2.69, *p* = 0.007) (See Figure 2.3). No statistically significant effects were seen for visual distractors (*p* = 0.87) or for the interaction between visual and auditory distractors (*p* = 0.59).

Figure 2.3

Visual search error rate



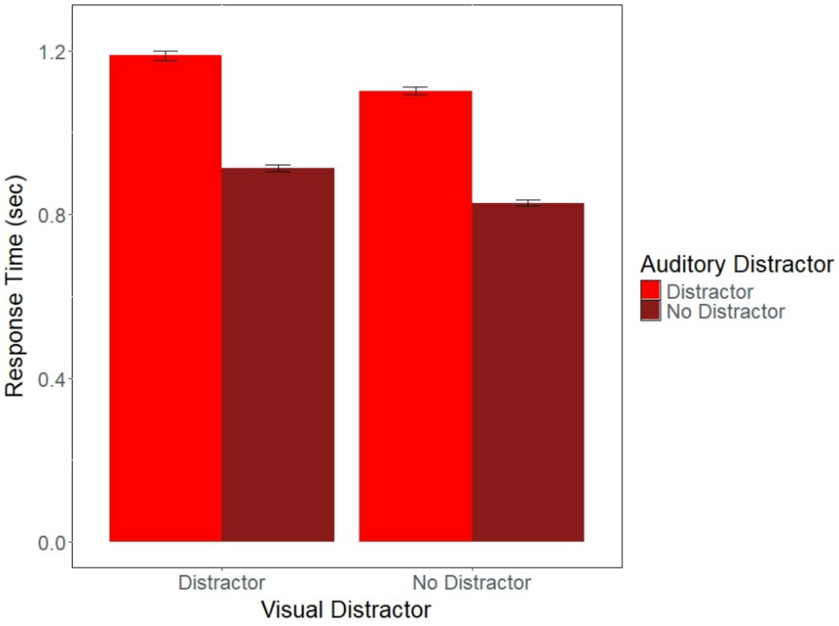
Note. Error bars are +/− SE.

Response Time.

A main effect of auditory distraction was identified such that search trials interrupted by the secondary auditory task were more likely to contribute to longer response times than uninterrupted trials (β = 1.12, SE = 0.00, z(8158) = 40.76, $p < 0.001$; See Figure 2.4). Similarly, we observed a main effect of visual distractor presence, indicating that response times were longer when a visual distractor was present in the search array (β = 1.03, SE = 0.01, z(8158) = 4.17, $p < 0.001$; See Figure 2.4). The interaction between visual and auditory distractors on response time was not significant ($p = 0.67$).

Figure 2.4

Visual search response time



Note. Error bars are +/− SE.

Exploratory Analysis

Attention Tendency.

Attention tendency did not significantly impact visual search error rate (*p* = 0.076) or response times (*p* = 0.070). No interactions between attention tendency and auditory or visual distractors reached significance for error rates. However, for response times, we observed a significant interaction between attention tendency and auditory distractors (β = 0.91, SE = 0.03, z(8158) = -2.59, *p* = 0.010), such that individuals with stronger attention-focusing abilities were more impacted by the varying auditory distractor conditions than their attention-shifting counterparts. This is further confirmed by the differences in response time between auditory distractor conditions for attention-focusers and attention-shifters. At the most extreme attention-focusing score (-0.213), the difference in predicted response time for auditory distractor versus no auditory distractor conditions equates to approximately 0.289 seconds, while the difference at the most extreme attention-shifting score (0.151) equates to approximately 0.203 seconds.

Stimuli Color.

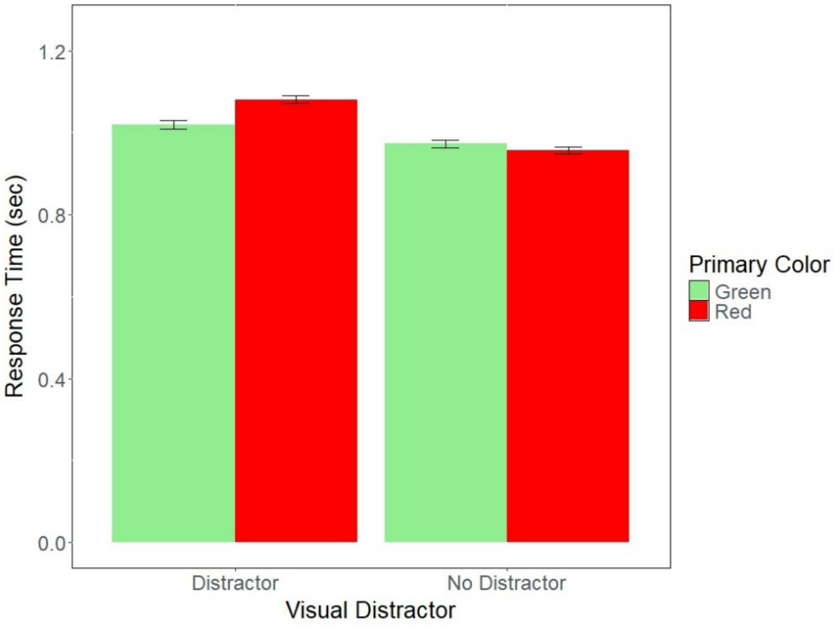
The colors of the visual search array objects were counterbalanced in this experiment to maintain the attention-capture effect of salient distractors; we expected no effect of stimuli color. This hypothesis held true in regards to visual search accuracy, as we observed no main effect or interactions involving primary stimulus color reached significance. However, a significant main effect for primary stimulus color and response time *did* emerge ($\beta = 1.02$, SE = 0.01, z(8158) = 2.01, $p = 0.045$). This main effect is qualified by the significant interaction between visual distractor condition and primary stimulus color ($\beta = 1.04$, SE = 0.01, z(8158) = 2.95, $p = 0.003$; see Figure 2.5) such that visual distractor-present trials with primarily red stimuli (i.e., red target and array objects with a green distractor present), yielded longer response times compared to trials either without a present visual distractor or to trials with primarily green stimuli.

Trial Number.

Observing the interaction between primary color and visual distractor on response time prompted additional analyses investigating the potential effect of trial number (see Table 2.3 for model). This was done to assess for potential contextual learning effects specific to the colors used in the AST stimuli (i.e., the potential for participants to learn which colors to suppress over the course of the experiment, ultimately reducing the distractibility of the additional singleton). For response time, we observed main effects of primary color ($\beta = 1.06$, SE = 0.02, z(8158) = 3.85, $p < 0.001$) and trial number ($\beta = 1.00$, SE = 0.00, z(8158) = -10.78, $p < 0.001$) such that red stimuli contributed to longer response times but participants ultimately got faster over time. However, the interaction between primary color and trial number was not significant ($p = 0.199$), indicating that the additional singleton remained distracting across trials. No main effects or interactions emerged in regards to error rate.

Figure 2.5

Interaction between primary stimulus color and visual distractor condition on response time



Note. Error bars are +/− SE.

Experiment 2 Discussion

Experiment 2 explored the impact of non-spatialized auditory interruptions on visual search performance by asking participants to identify what letter was held in an AST target and what an interrupting sound represents. Results showed that non-spatialized auditory interruptions negatively impacted both metrics — participants made more errors and took longer to respond on AST trials when an auditory distractor was present. Our findings, where impaired performance emerges from trials where participants must complete the auditory task before resuming visual search, are typical of modality-switching experimental conditions (Hunt & Kingstone, 2004;

Kreutzfeldt et al., 2015). This is especially the case when the switch in modality is uncued (Sandhu & Dyson, 2013), as in the present experiment. Response times were *also* lengthened by the presence of a visual distractor, consistent with the visual attention capture intended by the AST (Theeuwes, 1991, 1992). However, contrary to our hypothesis, interactions between visual and auditory distractor conditions for neither error rate nor response time were observed, meaning that trials with both distractor types were *not* more impaired than those with either or no distractors.

We had also hypothesized that the visual search performance of participants who reported more attention-focusing tendencies would not be as impacted by the presence of a distractor. This was based on our findings in Experiment 1 and the literature on individual differences influencing visual search (Adamo et al., 2017; Biggs et al., 2017). The results of Experiment 2 supported this hypothesis, indicating an interaction where the search performance of individuals with attention-focusing tendencies was less disrupted by the presence of an auditory distractor. It is possible that attention *control strategies*, rather than or in addition to attention *tendencies*, play a role in this continued observation. Goal-directed attention control has the potential to substantially assist in visual search, and while this strategy is not inherently tied to attention-focusing tendencies, it may be the case that individuals who are more attention-focusing in nature are better equipped to maintain this goal across visual search trials regardless of a secondary task (Irons & Leber, 2016, 2018, 2020). This is, however, merely speculation, and will be explored further in the future.

Exploratory analyses on the potential effects of color counterbalancing the AST stimuli also yielded interesting results. We expected this change to impact the distractibility of the visual distractors by preventing color-specific intentional “tune-outs.” This hypothesis was supported by longer response times in distractor-present AST trials and no effect of color-specific learning over the course of the experiment. However, we did not expect the primary color of the stimuli to have its own effect on search performance. The interaction observed between visual distractor condition and primary stimulus color indicates that when there is no distractor, response times for green- and red-dominant stimuli are similar; in distractor-present trial though, response times for the red-dominant stimuli are significantly longer than for the green-dominant stimuli. This would imply that, as a distractor, green is the more salient of the two.

Without an effect of learning, which could indicate the implementation of color-specific attentional control strategies, additional exploratory evaluation of the stimuli themselves was conducted. The grey background used in this experiment was employed to ensure that both red and green objects could equally stand out against the background, but this was only in terms of color distance. This interaction between visual distractor and primary color prompted an additional investigation into the *luminance* of both colors used in our AST arrays. Using the Web Content Accessibility Guidelines (2024) luminance formula, it was determined that the shade of green used has a relative luminance of 0.715 while the red used has a relative luminance of 0.213. This may explain the heightened distractibility of green distractors during red-dominant AST trials, and future iterations of this study will explore equally luminant color options.

Cross-Experiment Analyses

Data Preparation and Analysis

Cross-experiment analyses were conducted to explore whether the differing task stimuli in Experiments 1 and 2 differentially affected visual search performance. As discussed, there were differences in both the auditory and visual stimuli between the two experiments (i.e., spatialized tones and a black AST background in Experiment 1 vs. identifiable sounds and a grey background with counterbalanced colors in Experiment 2). These additional analyses allowed us to assess if any post-interruption impairments to visual search (i.e., more errors or longer response times) were due simply to the modality/task-switching nature of the interrupting auditory trials or due to the spatial localization aspect of the tones from Experiment 1. We also assessed if the color alterations to the visual search stimuli enhanced salience — and consequently, the distractibility — of our visual distractors. The data sets from both experiments were combined prior to assessing visual search error rates and response times with generalized mixed linear models. Nearly identical models to those used for the primary analyses of Experiment 2 were again run, using “Experiment” as an additional factor.

Results

Primary Analyses.

Error Rate.

A main effect of auditory distractor condition was observed ($\beta = 1.21$, SE = 0.06, z(16318) = 3.79, $p < 0.001$) such that, regardless of experiment or visual distractor condition, trials with an auditory distractor yielded higher visual search error rates (See Figure 2.6). Neither visual distractors ($p = 0.594$) nor experiment ($p = 0.259$) significantly impacted error rates, and no interactions between variables were significant.

Response Time.

Main effects were observed for visual distractors ($\beta = 1.02$, SE = 0.01, z(16318) = 3.27, $p = 0.001$) and auditory distractors ($\beta = 1.12$, SE = 0.00, z(16318) = 59.73, $p < 0.001$), regardless of experiment ($\beta = 1.03$, SE = 0.02, z(16318) = 1.93, $p = 0.053$; See Figure 2.7). No interactions reached significance, meaning that the nature of the auditory distractor stimuli (spatial tones versus non-spatial sounds) or of the additional singleton (red distractors only versus red or green distractors) did not impact response times.

Exploratory Analyses.

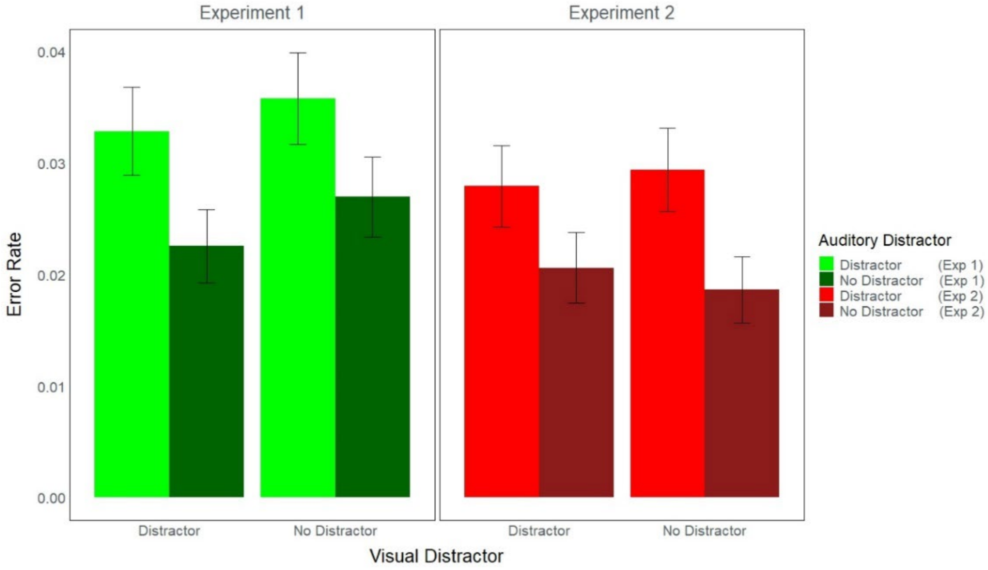
Stimuli Color.

We changed the background color of the additional singleton task stimuli from black to grey in Experiment 2 to increase the salience of the AST distractor and hypothesized that this change would yield greater impairment in visual search performance. To account for differences in the interrupting auditory task, exploratory analyses compared visual search performance of trials *not* interrupted by an auditory distraction. All uninterrupted visual search trials from Experiment 1 were compared to the uninterrupted, green-target trials of Experiment 2. Trials from Experiment 2 with red targets were excluded for comparability. See Table 2.5 for model details.

No main effects or significant interactions were observed when analyzing both visual search error rates and response times across experiments. Evidently, the change in stimulus background color did not achieve its intended effect of inducing greater visual search impairment in Experiment 2 compared to Experiment 1.

Figure 2.6

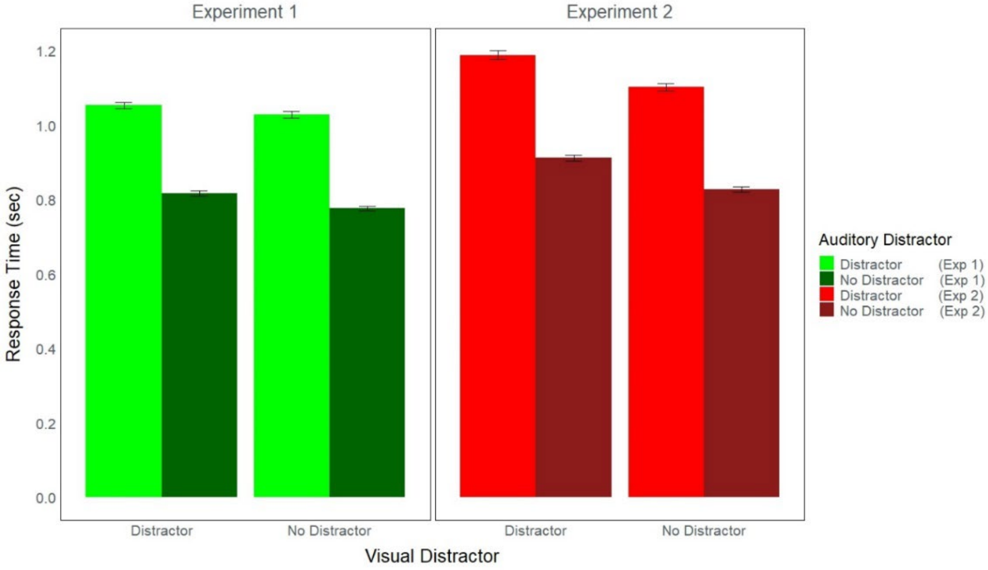
Visual search error rate by experiment and distractor condition



Note. Error bars are +/− SE.

Figure 2.7

Visual search response time by experiment and distractor condition



Note. Error bars are +/− SE.

Table 2.5

Cross-experiment color analysis GLMMs

Model Formula	*AIC*	*Marginal R^2/ Conditional R^2*	*RMSE*
Error Rate			
Visual Search Error Rate ~ Visual Distraction * Experiment + (1|Participant) + (1|Trial Image)	2948.79	0.01 / 0.21	0.160
Response Time			
Visual Search Response Time ~ Visual Distraction * Experiment + (1|Participant) + (1|Trial Image)	3146.30	0.00 / 0.05	0.363

General Discussion

The experiments included in this study sought to clarify the effect spatialized auditory distractions may have on visual search performance. This was evaluated by varying whether participants identified where an interrupting sound originates (Experiment 1) or what it represents (Experiment 2) while engaging with the Additional Singleton Task. In Experiment 1, we found that visual search error rates were negatively impacted by the presence of spatialized auditory distractors, regardless of the presence of visual distractors, and response times were longer when either an auditory distractor or both distractors were present. Experiment 2 was conducted to determine whether the visual search results from Experiment 1 were influenced by the spatialized nature of the interrupting auditory tones or if an alternative explanation, such as task- or modality-switching) better accounted for the findings.

We had hypothesized that visual search performance in Experiment 2 would continue to be impaired by interrupting auditory distractions, but less so than when those distractions were spatialized (i.e., Experiment 1's visual search results were negatively impacted by the *spatial* tones, rather than the auditory disruption in general). However, the lack of significant interactions between experiments and distractor conditions of either modality for our performance measures indicates that this hypothesis was not supported (per our cross-experiment analyses). Regardless of the content of the interrupting auditory task, whether the stimuli were spatialized or not, visual search trials subject to auditory distractors produced more errors and longer response times on the AST; similarly, visual distractors in either experiment increased response times on the AST. Though participants may have begun either experiment intent on locating the target object in the array and reporting the letter contained within, bottom-up visual attention, susceptible to salient distractors, led to response delays consistent with the AST (Theeuwes, 1991, 1992).

These findings imply that modality switching, rather than task-switching between spatial and non-spatial identification, has a powerful negative effect on visual search performance. Though our hypothesis was not supported, our results fit well within the literature specific to isolated modality switching (e.g., Richard et al., 2002) and modality switching with concurrent task-switching (e.g., Hunt & Kingstone, 2004; Kreutzfeldt et al., 2015) where the cost associated with switching modalities has been well-established. Additionally, we had not used modality-switch cues, hoping to better mirror the unexpected nature of cross-modal distractions that occur in daily life; our findings further align with the costs of reactive, rather than prepared, modality-switch responses (Sandhu & Dyson, 2013).

Limitations

Informed by Experiment 1, several changes were made to the AST stimuli in Experiment 2 to promote the salience and distractibility of the additional singleton. These included adjusting the background of the AST arrays to be grey and counterbalancing the colors for targets and distractors. However, these changes may have prevented adequate, comparable analyses by introducing confounds across experiments. For instance, our exploratory analysis regarding the background color of our AST arrays, in order to successfully compare between experiments, required that we use only a quarter of the data from Experiment 2 (as opposed to half of the AST data from Experiment 1). While that analysis did not render any significance of Experiment (and consequently, the background color of the stimuli), we acknowledge the amount of data included in that analysis could impact those findings.

It is also important to note that the Attentional Control Scale used in both experiments is a subjective, self-report measure. There are indications that it may be better at capturing a participant's *belief* in their attention tendencies than the ability itself (Khodami et al., 2024; Quigley et al., 2017). As a result, the findings we report in regards to attention tendency must be appreciated in the context of the validity of the Attentional Control Scale.

Summary and Implications

Taking into account both experiments, this study identified that unexpected auditory distractions, regardless of their spatial properties, are detrimental to visual search performance. Many high professional environments (e.g., radiology, airport security) regularly use visual search, with dire consequences emerging from errors made or delays in identifying a target. Individuals in these professions also regularly encounter (and cannot avoid) auditory distractions both spatialized and non-spatialized in nature, which our findings indicate may be detrimental to efficient and accurate visual search. To minimize the effects of such cross-modal distractors on professional search, noise-reduction solutions (such as noise-cancelling headphones with transparency modes or noise-reduction earplugs) should be explored. Suggestions like this could contribute to error reduction or efficiency boosts in organizations like the TSA or by hospital radiology departments, as fewer auditory distractions would be able to pull attention away from the search process.

Not only professional searchers could benefit from noise reduction, though. This study began with the discussion of how visual search is an inherent part of near all individuals' daily lives, for example, while driving. Unexpected auditory distractors may not be completely avoidable, such as emergency sirens or those caused by passengers in the vehicle (Koppel et al., 2011), but taking steps to reduce auditory distractions to visual search like silencing mobile devices may reduce the effect of these stimuli on visual attention.

Future Directions

The experiments discussed in this paper are intended to be the first two of four studies to parametrically explore the relationship between the audiovisual ventral and dorsal pathways by varying whether participants will identify what or where visual or auditory stimuli are or represent. Experiment 1 had participants identify *what* letter is in a visual singleton shape and *where* an interrupting sound originates, while Experiment 2 had participants identify *what* letter is in the singleton shape and *what* the interrupting sound represents. Before proceeding with the remaining two experiments, however, we intend to rerun Experiment 1 with counterbalanced AST arrays and with the same background color as used in Experiment 2. Minimizing the visual stimuli differences between experiments, we believe, will allow for a more faithful assessment of how visual search performance differs when interrupted by spatialized versus non-spatialized auditory distractions. We will also be evaluating color pairs for equal salience and luminance to be used in the AST, rather than the current shades of red and green, to ensure color counterbalanced distractor objects are equally capable of drawing visual attention.

Building on these adjustments, Experiment 3 will have participants identify *where* the singleton shape is in the array and *where* the interrupting sound originates. Experiment 4, finally, will have participants identify *where* the singleton shape is in the array and *what* the interrupting sound represents. In completing this series, we hope to clarify the extent of cross-modal attention abilities when facing distractions that are encountered in daily life.

References

Adam, K. C. S., Patel, T., Rangan, N., & Serences, J. T. (2021). Classic Visual Search Effects in an Additional Singleton Task: An Open Dataset. *Journal of Cognition*, *4*(1), 34. https://doi.org/10.5334/joc.182

Adam, K. C. S., & Serences, J. T. (2021). History Modulates Early Sensory Processing of Salient Distractors. *Journal of Neuroscience*, *41*(38), 8007–8022. https://doi.org/10.1523/JNEUROSCI.3099-20.2021

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2017). An individual differences approach to multiple-target visual search errors: How search errors relate to different characteristics of attention. *Vision Research*, *141*, 258–265. https://doi.org/10.1016/J.VISRES.2016.10.010

Balint, B. J., Steenburg, S. D., Lin, H., Shen, C., Steele, J. L., & Gunderman, R. B. (2014). Do Telephone Call Interruptions Have an Impact on Radiology Resident Diagnostic Accuracy? *Academic Radiology*, *21*(12), 1623–1628. https://doi.org/10.1016/J.ACRA.2014.08.001

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809–823. https://doi.org/10.1016/S0896-6273(04)00070-4

Biggs, A. T., Clark, K., & Mitroff, S. R. (2017). Who should be searching? Differences in personality can affect visual search accuracy. *Personality and Individual Differences*, *116*, 353–358. https://doi.org/10.1016/J.PAID.2017.04.045

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.4.13) [R].

Borowsky, A., Horrey, W. J., Liang, Y., Garabet, A., Simmons, L., & Fisher, D. L. (2016). The effects of brief visual interruption tasks on drivers' ability to resume their visual search for a pre-cued hazard. *Accident Analysis & Prevention*, *93*, 207–216. https://doi.org/10.1016/J.AAP.2016.04.028

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., & Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience 1999 2:8*, *2*(8), 759–766. https://doi.org/10.1038/11239

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649–657. https://doi.org/10.1016/S0960-9822(00)00513-3

Campbell, A., Kirkpatrick, A., Connor, J., & Cooper, M. (Eds.). (2024). *Web Content Accessibility Guidelines (WCAG) 2.1* [Technical Report]. Accessibility Guidelines Working Group. https://www.w3.org/TR/WCAG21/

Chapman, P. R., & Underwood, G. (1998). Visual Search of Driving Situations: Danger and Experience. *Perception*, *27*(8), 951–964. https://doi.org/10.1068/P270951

Chen, Y., Fang, W., Guo, B., & Bao, H. (2022). The moderation effects of task attributes and mental fatigue on post-interruption task performance in a concurrent multitasking environment. *Applied Ergonomics*, *102*, 103764. https://doi.org/10.1016/J.APERGO.2022.103764

Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1554–1568. https://doi.org/10.1037/a0024329

Cosman, J. D., & Vecera, S. P. (2013). Context-dependent control over attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 836–848. https://doi.org/10.1037/a0030027

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*(2), 225–236. https://doi.org/10.1037/0021-843X.111.2.225

Folk, C. L., & Annett, S. (1994). Do locally defined feature discontinuities capture attention? *Perception & Psychophysics*, *56*(3), 277–287.

Folk, C. L., & Remington, R. (1998). Selectivity in Distraction by Irrelevant Featural Singletons: Evidence for Two Forms of Attentional Capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 847–858.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary Covert Orienting Is Contingent on Attentional Control Settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Gandomkar, Z., & Mello-Thoms, C. (2019). Visual search in breast imaging. *British Journal of Radiology*, *92*(1102). https://doi.org/10.1259/BJR.20190057/7449310

Gaspelin, N., & Luck, S. J. (2018). Distinguishing among potential mechanisms of singleton suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(4), 626–644. https://doi.org/10.1037/xhp0000484

Ho, C., Tan, H. Z., & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(6), 397–412. https://doi.org/10.1016/j.trf.2005.05.002

Hunt, A. R., & Kingstone, A. (2004). Multisensory executive functioning. *Brain and Cognition*, *55*(2), 325–327. https://doi.org/10.1016/j.bandc.2004.02.072

Ingle, D., Goodale, M. A., & Mansfield, R. J. W. (1982). *Analysis of visual behavior*. Cambridge, Massachusetts: MIT Press.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*(3), 548–554. https://doi.org/10.3758/PBR.15.3.548

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, *78*(7), 2031–2048. https://doi.org/10.3758/s13414-016-1125-4

Irons, J. L., & Leber, A. B. (2018). Characterizing individual variation in the strategic use of attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(10), 1637–1654. https://doi.org/10.1037/xhp0000560

Irons, J. L., & Leber, A. B. (2020). Developing an Individual Profile of Attentional Control Strategy. *Current Directions in Psychological Science*. https://doi.org/10.1177/0963721420924018

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346–354.

Khodami, M. A., Battaglini, L., Jansarvatan, M., Kireeva, S., & Bagheri, S. (2024). Comparing Self-Report vs. Performance Measures of Attentional Control and Efficiency. *NeuroSci*, *5*(2), 114–127. https://doi.org/10.3390/NEUROSCI5020008

Klein, R. M. (1977). Attention and Visual Dominance: A Chronometric Analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(3), 365–378.

Knoeferle, K. M., Knoeferle, P., Velasco, C., & Spence, C. (2016). Multisensory brand search: How the meaning of sounds guides consumers' visual attention. *Journal of Experimental Psychology: Applied*, *22*(2), 196–210. https://doi.org/10.1037/xap0000084

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2009a). Auditory and Visual Capture During Focused Visual Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1303–1315. https://doi.org/10.1037/a0013901

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2009b). Competition between auditory and visual spatial cues during visual task performance. *Experimental Brain Research*, *195*(4), 593–602. https://doi.org/10.1007/s00221-009-1829-y

Koppel, S., Charlton, J., Kopinathan, C., & Taranto, D. (2011). Are child occupants a significant source of driving distraction? *Accident Analysis & Prevention*, *43*(3), 1236–1244. https://doi.org/10.1016/j.aap.2011.01.005

Kreutzfeldt, M., Stephan, D. N., Sturm, W., Willmes, K., & Koch, I. (2015). The role of crossmodal competition and dimensional overlap in crossmodal attention switching. *Acta Psychologica*, *155*, 67–76. https://doi.org/10.1016/j.actpsy.2014.12.006

Lewald, J., Dörrscheidt, G. J., & Ehrenstein, W. H. (2000). Sound localization with eccentric head position. *Behavioural Brain Research*, *108*(2), 105–125. https://doi.org/10.1016/S0166-4328(99)00141-2

Lüdecke, D. (2023). *sjPlot: Data visualization for statistics in social science* (Version 2.8.15) [Computer software]. https://CRAN.R-project.org/package=sjPlot

Lunn, J., Sjoblom, A., Ward, J., Soto-Faraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition*, *187*, 38–49. https://doi.org/10.1016/J.COGNITION.2019.02.008

Luo, M. R., Cui, G., & Rigg, B. (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, *26*(5), 340–350. https://doi.org/10.1002/COL.1049

Ma, Y., Qi, S., Zhang, Y., Lian, G., Lu, W., & Chan, C. Y. (2020). Drivers' Visual Attention Characteristics under Different Cognitive Workloads: An On-Road Driving Behavior Study. *International Journal of Environmental Research and Public Health*, *17*(15), 5366. https://doi.org/10.3390/IJERPH17155366

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, *39*(12), 1304–1316. https://doi.org/10.1016/S0028-3932(01)00119-1

Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(5482), 1206–1208. https://doi.org/10.1126/science.289.5482.1206

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024). The Surprising Robustness of Visual Search Against Concurrent Auditory Distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(1), 99–118. https://doi.org/10.1037/xhp0001168

Mandal, A., Röer, J. P., & Liesefeld, H. R. (2024). Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant. *Visual Cognition*, 1–17. https://doi.org/10.1080/13506285.2024.2397825

Mazza, V., Turatto, M., Rossi, M., & Umiltà, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia*, *45*(3), 514–522. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2006.02.010

Milner, D., & Goodale, M. (2006). *The Visual Brain in Action* (2nd ed.). Oxford University Press.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Mitroff, S. R., Ericson, J. M., & Sharpe, B. (2018). Predicting Airport Screening Officers' Visual Search Competency With a Rapid Assessment. *Human Factors*, *60*(2), 201–211. https://doi.org/10.1177/0018720817743886

Moher, J. (2020). Distracting Objects Induce Early Quitting in Visual Search. *Psychological Science*, *31*(1), 31–42. https://doi.org/10.1177/0956797619886809

Murray, M. M., Santis, L. D., Thut, G., & Wylie, G. R. (2009). The costs of crossing paths and switching tasks between audition and vision. *Brain and Cognition*, *69*(1), 47–55. https://doi.org/10.1016/j.bandc.2008.05.004

NCSA. (2023). *Distracted Driving in 2021* (DOT HS 813 443). National Highway Traffic Safety Administration.

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

Quigley, L., Wright, C. A., Dobson, K. S., & Sears, C. R. (2017). Measuring Attentional Control Ability or Beliefs? Evaluation of the Factor Structure and Convergent Validity of the Attentional Control Scale. *Journal of Psychopathology and Behavioral Assessment*, *39*(4), 742–754. https://doi.org/10.1007/s10862-017-9617-7

R Core Team. (2024). *R: A Language and Environment for Statistical Computing* (Version 4.3.3) [Computer software].

Ratwani, R. M., & Trafton, J. G. (2008). Spatial memory guides task resumption. *Visual Cognition*, *16*(8), 1001–1010. https://doi.org/10.1080/13506280802025791

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800–11806. https://doi.org/10.1073/PNAS.97.22.11800

Richard, C. M., Wright, R. D., Ee, C., Prime, S. L., Shimizu, Y., & Vavrik, J. (2002). Effect of a Concurrent Auditory Task on Visual Search Performance in a Driving-Related Image-Flicker Task. *Human Factors*, *44*(1), 108–119. https://doi.org/10.1518/0018720024494874

Sandhu, R., & Dyson, B. J. (2013). Modality and task switching interactions using bi-modal and bivalent stimuli. *Brain and Cognition*, *82*(1), 90–99. https://doi.org/10.1016/j.bandc.2013.02.011

Santangelo, V., Belardinelli, M. O., & Spence, C. (2007). The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 137–148. https://doi.org/10.1037/0096-1523.33.1.137

Santangelo, V., & Spence, C. (2007a). Assessing the automaticity of the exogenous orienting of tactile attention. *Perception*, *36*(10), 1497–1505. https://doi.org/10.1068/p5848

Santangelo, V., & Spence, C. (2007b). Multisensory Cues Capture Spatial Attention Regardless of Perceptual Load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1311–1321. https://doi.org/10.1037/0096-1523.33.6.1311

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception and Psychophysics*, *70*(2), 208–218. https://doi.org/10.3758/APP.72.3.672

Sestieri, C., Di Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., Olivetti Belardinelli, M., & Romani, G. L. (2006). “What” versus “Where” in the audiovisual domain: An fMRI study. *NeuroImage*, *33*(2), 672–680. https://doi.org/10.1016/j.neuroimage.2006.06.045

Snider, J., Spence, R. J., Engler, A.-M., Moran, R., Hacker, S., Chukoskie, L., Townsend, J., & Hill, L. (2023). Distraction “Hangover”: Characterization of the Delayed Return to Baseline Driving Risk After Distracting Behaviors. *Human Factors*, *65*(2), 306–320. https://doi.org/10.1177/00187208211012218

Spence, C., & Driver, J. (1997a). Audiovisual links in exogenous covert spatial orienting. *Perception and Psychophysics*, *59*(1), 1–22. https://doi.org/10.3758/BF03206843

Spence, C., & Driver, J. (1997b). On measuring selective attention to an expected sensory modality. *Perception and Psychophysics*, *59*(3), 389–403. https://doi.org/10.3758/BF03211906

Steinbach, R., & Tefft, B. C. (2023). *American Driving Survey: 2022*. Washington, D.C.: AAA Foundation for Traffic Safety.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science*, *292*(5515), 290–293. https://doi.org/10.1126/science.1058911

Treisman, A. M., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, *12*, 97–136.

Trick, L. M., Enns, J. T., Mills, J., & Vavrik, J. (2004). Paying attention behind the wheel: A framework for studying the role of attention in driving. *Theoretical Issues in Ergonomics Science*, *5*(5), 385–424. https://doi.org/10.1080/1463922041233129893

Turatto, M., & Galfano, G. (2001). Attentional capture by color without any relevant attentional set. *Perception & Psychophysics*, *63*(2), 286–297. https://doi.org/10.3758/BF03194469

Walker, H. E. K., & Trick, L. M. (2018). Mind-wandering while driving: The impact of fatigue, task length, and sustained attention abilities. *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 81–97. https://doi.org/10.1016/J.TRF.2018.08.009

Williams, L. H., & Drew, T. (2017). Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? *Cognitive Research: Principles and Implications*, *2*(1), 1–11. https://doi.org/10.1186/s41235-017-0050-y

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour 2017 1:3*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Zehetleitner, M., Goschy, H., & Müller, H. J. (2012). Top-down control of attention: It's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 941–957. https://doi.org/10.1037/a0027629

Appendix

Appendix A

Demographic characteristics of participants

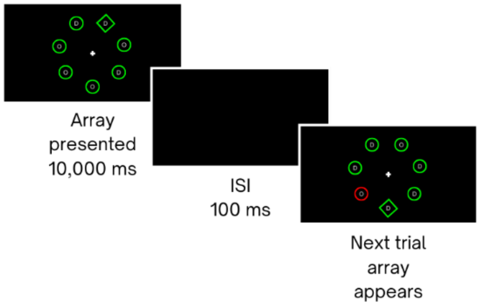
Characteristic	*Experiment 1*		*Experiment 2*	
	n	*%*	*n*	*%*
Gender				
Female	22	64.7	21	61.8
Male	12	35.3	12	35.3
Non-Binary	0	0.0	1	2.9
Race				
Asian	8	23.5	8	23.5
Black/African American	7	20.6	0	0.0
Hispanic/Latino	1	2.9	0	0.0
Indigenous American or Alaskan Native	0	0.0	0	0.0
Middle Eastern or North African	0	0.0	1	2.9
Mixed Race	2	5.9	5	14.7
Pacific Islander/Native Hawaiian	0	0.0	0	0.0
White	15	44.1	20	58.8
Prefer Not to Respond	1	2.9	0	0.0
Handedness				
Ambidextrous	1	2.9	0	0.0
Left-Handed	1	2.9	4	11.8
Right-Handed	32	94.1	30	88.2
ADHD Diagnosis				
No	31	91.2	31	91.2
Yes	3	8.8	3	8.8

Note. $N_{Exp1} = 34$ ($M_{age} = 19.91$, $SD = 2.69$ years); $N_{Exp2} = 34$ ($M_{age} = 18.41$, $SD = 0.66$ years).

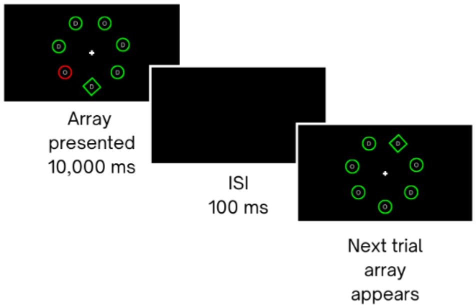
Appendix B

Trial timing by condition

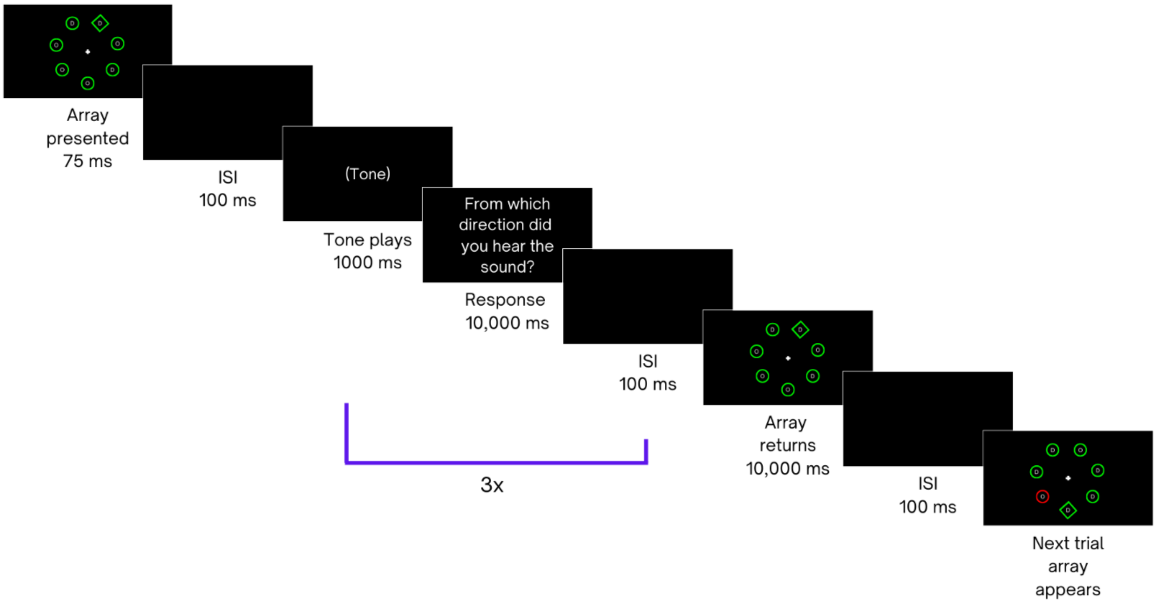
A. No Visual Distractor/No Auditory Distractor



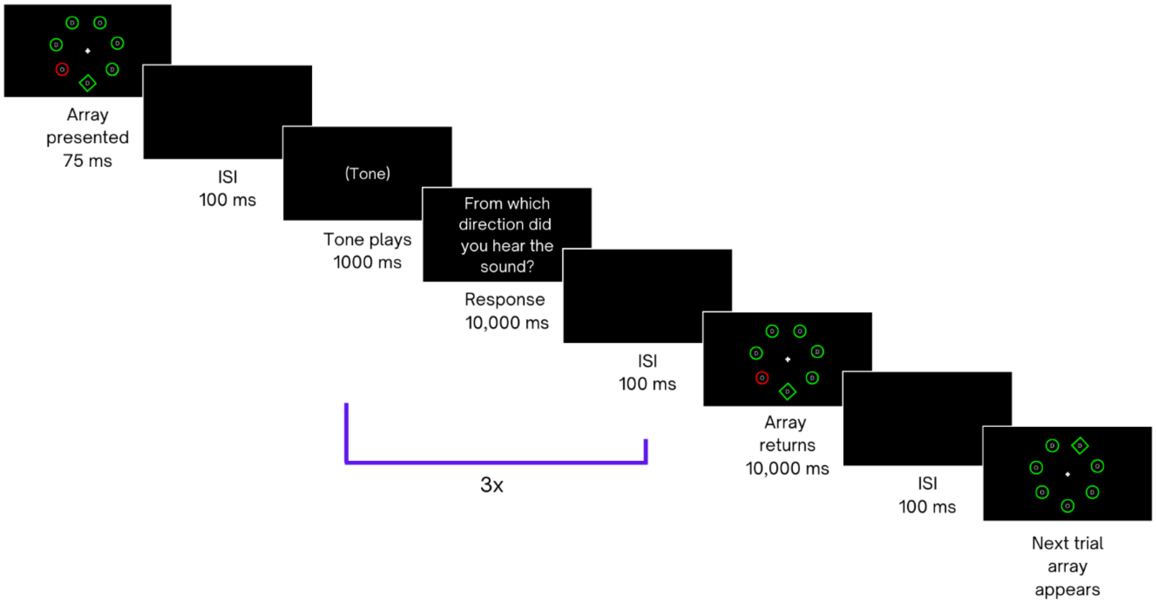
B. Visual Distractor/No Auditory Distractor



C. No Visual Distractor/Auditory Distractor



D. Visual Distractor/Auditory Distractor



Note. Jittering was not employed in the timing of trials.