Visual-cortical enhancement by acoustic distractors: The effects of endogenous spatial attention and visual working memory load

Shari Cavicchi , Andrea De Cesarei , Matteo Valsecchi , Maurizio Codispoti [*]

Department of Psychology, University of Bologna, Italy

A B S T R A C T

Past work has shown that when a peripheral sound captures our attention, it activates the contralateral visual cortex as revealed by an event-related potential component labelled the auditory-evoked contralateral occipital positivity (ACOP). This cross-modal activation of the visual cortex has been observed even when the sounds were not relevant to the ongoing task (visual or auditory), suggesting that peripheral sounds automatically activate the visual cortex. However, it is unclear whether top-down factors such as visual working memory (VWM) load and endogenous attention, which modulate the impact of task-irrelevant information, may modulate this spatially-specific component. Here, we asked participants to perform a lateralized VWM task (change detection), whose performance is supported by both endogenous spatial attention and VWM storage. A peripheral sound that was unrelated to the ongoing task was delivered during the retention interval. The amplitude of sound-elicited ACOP was analyzed as a function of the spatial correspondence with the cued hemifield, and of the memory array set-size. The typical ACOP modulation was observed over parieto-occipital sites in the 280–500 ms time window after sound onset. Its amplitude was not affected by VWM load but was modulated when the location of the sound did not correspond to the hemifield (right or left) that was cued for the change detection task. Our results suggest that sound-elicited activation of visual cortices, as reflected in the ACOP modulation, is unaffected by visual working memory load. However, endogenous spatial attention affects the ACOP, challenging the hypothesis that it reflects an automatic process.

1. Introduction

In everyday life, selection of relevant information from the world around us is controlled by cognitive factors such as current goals, expectations, and past learning experiences (Egeth & Yantis, 1997; Luck, Gaspelin, Folk, Remington & Theeuwes, 2021). Although attention can be deployed voluntarily, salient environmental events can also capture attention, even if irrelevant for the ongoing task (Codispoti, De Cesarei, Biondi, & Ferrari, 2016; Failing & Theeuwes, 2018; Lleras, Buetti, & Mordkoff, 2013). The processing of these task-irrelevant stimuli may interrupt ongoing goal-directed activity (i.e. distraction), but it is also adaptive in informing us about environmental changes.

Previous findings suggest that the ability to ignore irrelevant distractors is modulated by the type and level of processing load involved in the current task (Lavie, 2010; Lavie, Hirst, De Fockert & Viding, 2004). Accordingly, recent studies indicate that high visual working memory (VWM) load leads to reduced distractor processing, in the same way as when perceptual resources are exhausted (Konstantinou, Beal, King &

Lavie, 2014; Roper & Vecera, 2014). The visual cortex plays a key role in supporting VWM, for instance during the maintenance of visual objects (sensory-recruitment hypothesis; e.g., Harrison & Tong, 2009). While behavioral studies have suggested that both perceptual and visual working memory load affect distractor interference in flanker tasks (Konstantinou et al., 2014), less is known regarding the impact of visual working memory load on auditory distractor processing.

Sudden sounds represent an example of a potential source of distraction and are especially effective at capturing attention. Since the seminal study by Spence and Driver (1997) a large body of evidence has accumulated indicating that a lateralized sound can act as an exogenous cue and capture visual attention crossmodally (see Hillyard, Störmer, Feng, Martinez & McDonald, 2016; Störmer, 2019 for reviews), even overriding the voluntary orienting of attention promoted by arrow cues (Van der Lubbe & Postma, 2005; Van der Lubbe, Havik, Bekker & Postma, 2006; McDonald, Teder-Sälejärvi, Russo & Hillyard, 2003; McDonald & Ward, 2000). Recent research revealed that sudden sounds activate the visual cortex, even when they are not relevant for the

ongoing task (Hillyard et al., 2016; McDonald, Störmer, Martinez, Feng & Hillyard, 2013). Specifically, it has been demonstrated that peripheral sounds elicited a positive ERP over the contralateral occipital hemisphere; this lateralized ERP component was labelled "Auditory-Evoked Contralateral Occipital Positivity" (ACOP; Hillyard et al., 2016). The ACOP is a slow positive ERP (200–500 ms from sound onset), arising from the visual cortex, that is larger in the hemisphere contralateral to the stimulus location and is associated with an involuntary orienting of attention to salient sounds (Feng, Störmer, Martinez, McDonald & Hillyard, 2014; Keefe & Störmer, 2021; McDonald et al., 2013). It is also well-documented that spatially non-predictive sound cues enhance the processing of colocalized visual targets, and a more pronounced ACOP amplitude is predictive of enhanced perceptual processing (Hillyard et al., 2016). Moreover, the ACOP can also be triggered by an irrelevant sound presented in a randomized sequence with auditory targets, which lead McDonald and colleagues to suggest that the orienting to sounds indexed by the ACOP is automatic (Hillyard, et al., 2016; McDonald, et al., 2013). However, a recent study presented task-relevant sounds and asked participants to categorize one of several stimulus dimensions; under these conditions, an ACOP was only observed when the spatial location of the sound was relevant, leading the authors to suggest that the ACOP may be sensitive to task set (Retsa, Matusz, Schnupp & Murray, 2020).

The present study examined whether acoustic distractor processing, indexed by the ACOP, can be modulated by endogenous spatial attention and visual working memory. Since it has been shown that VWM recruits sensory processing areas to maintain visual information online (Harrison & Tong, 2009; Gayet et al., 2017a; 2017b; Teng & Postle, 2021), in the present study, we investigated whether visual working memory load affects the visual cortex response (ACOP) to distracting sounds. In addition, we examined if the ACOP reflects a reflexive orienting to peripheral sounds or can be modulated by endogenous spatial attention.

We asked participants to perform a lateralized version of a change detection task (Luck & Vogel, 1997; Phillips, 1974) with variable set-sizes (low vs. high load; Luck & Vogel, 1997; Vogel & Machizawa, 2004). During the retention interval of each trial, a task-irrelevant sound was delivered unilaterally. We analyzed changes in the ACOP modulation as a function of set size and correspondence with the location of the memoranda. We also measured Pashler's K score and Contralateral Delay Activity (CDA), in order to use them as indexes of VWM load. The CDA is an occipital negativity contralateral to the cued hemifield, and it is observed during the delay interval in change detection tasks, starting from 300 ms after the onset of the memory array. Previous studies have shown that CDA amplitude increases along with the number of items to be remembered, until capacity limit is reached (typically three or four items; Luria, Balaban, Awh & Vogel, 2016).

We can formulate two distinct predictions based on the hypotheses that VWM and/or endogenous spatial attention affect the processing of distractor sounds. Firstly, if VWM load reduces distractor processing, then ACOP amplitude should be attenuated under high compared to low VWM load. Secondly, if the ACOP reflects an automatic process, then endogenous spatial attention should not modulate it. Otherwise, a mismatch between the spatial correspondence between the cued hemifield and sound location should attenuate or even eliminate the ACOP.

2. Methods

2.1. Participants

A total of 18 participants (7 women) took part in the study. The sample size was determined based on previous studies (Lakens, 2022). We decided to collect a similar number of participants as in previous ACOP studies (e.g., Feng, Störmer, Martinez, McDonald & Hillyard, 2014; Matusz et al., 2016; McDonald et al., 2013; Störmer, Feng, Martinez, McDonald & Hillyard, 2016). Age ranged from 19 to 36 years (M = 24.11 years, SD = 4.14 years). All participants had normal or corrected-to-normal vision and hearing, and none of them reported current or past neurological or psychopathological problems. All participants provided written informed consent in accordance with the Bioethics Committee protocols of the University of Bologna, and the experiment was conducted in accordance with the tenets of the Declaration of Helsinki. Participants had no previous experience with this task, and volunteered to participate in the study without receiving any monetary compensation or course credit.

2.2. Apparatus and stimuli

Participants were seated with a chin rest in a comfortable chair in an electrically shielded room. Visual stimuli were presented on an LCD computer screen (ViewSonic XG2530; 100 Hz refresh rate; 1280 × 768 pixels) placed at a distance of 60 cm from participants. Auditory stimuli were delivered through a pair of external loudspeakers (Trust Arys, 4 Ohm, 28 W) positioned to the left and right sides of the monitor. The two speakers were arranged symmetrically in the frontal azimuth plane, to the immediate right and left of the monitor. The distance (center-to-center) of one speaker from the other was 63 cm, the azimuthal angles were approximately ± 27.7°. Visual and auditory stimuli were presented using the E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). All stimuli appeared on a gray background (RGB = 128 128 128). A small black cross (0.67° × 0.67°) marking the center of the display was present during the whole trial. The cue consisted of a white arrow (1.4° × 0.3°) positioned 1.4° above the fixation cross. The memory array consisted of one or four squares per hemifield (each square measuring 1.1° × 1.1° of visual angle, distanced from the next square by at least 2°) whose color was selected randomly from a set of 7 possible colors (Red = RGB 255 0 0; Dark Green = 0 130 60; Blue= 0 0 255; Yellow = 255 255 0; Brown = 139 69 19; Cyan = 0 255 255; Green = 0 255 0). The squares could appear within an imaginary rectangular box, centered 3° from the screen center (to the left or right side), and measuring 4° horizontally and 10° vertically. The task-irrelevant sound was a pink noise burst (500 –15,000 Hz, 78 dB SPL, 83 ms duration with 5 ms rise and fall ramps) delivered unilaterally from one of the two speakers.

2.3. Procedure

A schematic illustration of a sample trial is depicted in Fig. 1. The observers performed a lateralized change detection task (e.g., Vogel & Machizawa, 2004). Each trial began with the cue arrow (200 ms), indicating the hemifield of the array that the observers had to remember (50 % probability of left cues and 50 % of right cues). Then a blank display remained on screen for a randomly jittered interval of 400–500 ms.[1] After this interval, two arrays appeared (100 ms), one in each hemifield, and participants were instructed to memorize the squares from the cued side and to ignore the other side entirely. The number of squares was always the same on both sides of the screen. The same hemifield could never be cued for more than 5 consecutive trials. After 900 ms or 1200 ms (50 % probability of 900 ms interval and 50 % of 1200 ms interval) from the offset of memory array, the task-irrelevant sound was delivered randomly from the speaker placed beside the cued side of the screen, or from the one placed beside the uncued side (50 % probability each; 50 % left, 50 % right). The interval after which the task-irrelevant sound was delivered was varied to reduce the temporal anticipation of the distractor. Participants were clearly informed that the sounds were task-irrelevant. Sounds were delivered from the same speaker for a maximum of 4 consecutive trials, to avoid the learning of contextual regularities. To keep the retention interval duration at 1800 ms every time, trials in which the sound occurred 900

---

[1] Because of a technical error in the program, for the first three participants the interval duration was set to 400 ms.
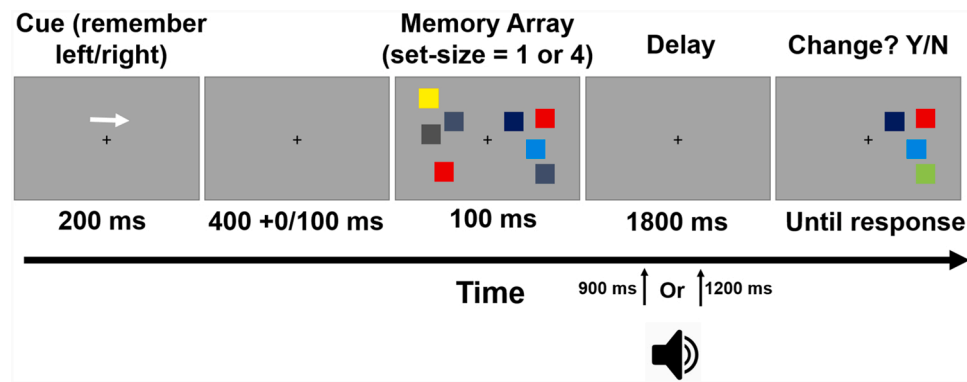
Fig. 1. Schematic illustration of the trial procedure. In the example a change trial from the high load condition can be seen. The auditory distractor could be delivered after 900 or 1200 ms from the memory array offset, randomly. Visually, only a gray background and central fixation cross were present during the whole retention interval.

ms after the memory array had another interval of 900 ms; for the remaining trials the second interval lasted 600 ms. At the end of the retention interval the probe display appeared. It consisted of squares located only on the side that had been cued in the original array and stayed on screen until a response was given. In half of the trials the colors were identical to the memory display ('no change' trials), in the remaining trials one of the colors was different ('change' trials). Participants responded by pressing with the index finger of their dominant hand one of two keys ("b" or "h") on a standard Italian keyboard that was placed in front of them. One of the keys associated with a 'change' response, the other with a 'no change' response, and the key-response association was counterbalanced across participants. Participants were instructed to prioritize accuracy over speed when responding. No feedback was given after the response. The same trial type ('change' or 'no change') was repeated for a maximum of 5 consecutive trials. After the response, the trial ended and the next trial started after a blank intertrial interval of 1000 ms.

The experiment was divided into two blocks, each one with a different set size. For consistency with the prior WM load literature (e.g., Gao & Theeuwes, 2020; Konstantinou et al., 2014; Tavares, Logie, & Mitchell, 2016; see also Lavie et al., 2004, note 5), the load condition was manipulated block-wise. In the low load block only one item per hemifield was presented in each trial, whereas in the high load block four items per hemifield were presented in each trial. Each block consisted of 336 trials, divided into 4 sub-blocks separated by brief pauses, the duration of which was decided by participants. An additional break was provided at the end of the first block. The total number of trials was 672, always preceded by a practice phase consisting of 20 trials (10 per load condition). The order of blocks was counterbalanced across participants. The whole experiment lasted about 1 h, including breaks.

2.4. EEG recording and processing

Continuous EEG was recorded at a sampling rate of 512 Hz from 64 active sites using an ActiveTwo Biosemi system. Electrodes were mounted in an elastic cap positioned according to the International 10–10 system. Two additional sensors were placed near the outer canthi of the left and right eyes to monitor horizontal eye-movements (horizontal EOG), one below the participant's left eye to monitor vertical eye-movements (vertical EOG), and two were placed near the left and right mastoid process. An additional reference electrode located near Cz served as the reference during data acquisition. A hardware fifth-order low-pass filter with a − 3 dB attenuation factor at 50 Hz was applied online. Off-line analysis was performed using Emegs 3.1 (Peyk, De Cesarei, & Junghöfer, 2011), a toolbox operating in the MATLAB (Mathworks, Natick, MA) environment (MATLAB 2015 version). Data were filtered using a low pass 23rd order IIR Butterworth

forward-backward filter with a passband ripple of 3 dB and a stopband attenuation of 45 dB (passband: 0–40 Hz, transition band: 40–50 Hz, stopband: 50 Hz and above), and an high pass 4th order IIR Butterworth forward-backward filter with a passband ripple of 1 dB and 18 dB attenuation in the stopband (stopband: 0–0.05 Hz, transition band: 0.05–0.1 Hz, passband: 0.1 and above). Then, blink, horizontal, and vertical eye-movement corrections were carried out by means of an automated regressive method (Schlögl, Keinrath, Zimmermann, Scherer, Leeb & Pfurtscheller, 2007). The EEG data were then segmented into epochs spanning 500 ms pre-stimulus to 1000 ms post-stimulus. Two separate analyses were carried out for CDA and ACOP analysis, the first based on the onset of the memory array, the latter on the onset of the sound (900–1200 ms after the offset of the memory array, see Fig. 1). Artifact detection was performed, separately for the CDA and ACOP epochs, through a semi-automatic procedure (Junghöfer, Elbert, Tucker & Rockstroh, 2000). For each sensor, a distribution of values of maximum absolute amplitude, standard deviation of amplitude, and absolute maximum temporal gradient is created. Based on these distributions, for each sensor Emegs automatically discards trials that exceed the typical values. Then, a single index of data quality is generated for each individual trial, based on the number of to-be-retained sensors and on their topographical distribution. Finally, depending on a threshold which can be adjusted by the user, bad trials are excluded from further analysis. For the remaining trials, data from excluded sensors are interpolated based on the nearest retained sensors using a spherical spline function. The mean percentage of rejected electrodes corresponded to 7.9 % (SD = 2.6 %). The mean percentage of rejected epochs corresponded to 12.6 % (SD = 3.6 %). Finally, data were re-referenced to the average of the right and left mastoid electrodes.

2.5. Statistical analyses

All the statistical analyses were performed using IBM SPSS Statistics 20 (IBM Corp., 2011), with the exception of the quantification of the Bayes factor (BF), that was performed using Bayesian hypothesis testing in JASP 0.16.4 (JASP Team, 2022). For each p value resulting from ANOVAs, we report the corresponding BF_{excl} for matched models, i.e. the BF that compares models that contain the effect to equivalent models stripped of the effect and excluding higher-order interactions.

2.5.1. Behavioral performance

Performance in the change detection task was examined in terms of error rate, reaction times (RTs), and an estimate of participants' visual working memory capacity. All the indexes of behavioral performance were first analyzed in a repeated measures univariate ANOVA with factors Correspondence (same/different side relative to sound location) and Load (low/high). Trials with incorrect responses or responses that

were faster or slower than 2 SD from each observer's mean were excluded from reaction times analysis. This resulted in the rejection of 15.6 % of trials (SD = 4.39). To obtain an average estimate of participants' visual working memory capacity, K values were computed for each load condition. We chose to use Pashler's formula (*K= N[(hit rate – false alarms) / (1 – false alarms)];* Pashler, 1988) to obtain K, since it has been suggested as more appropriate for whole-display recognition versions of the change detection task (Rouder, Morey, Morey & Cowan, 2011).

2.5.2. Electrophysiological data

To collapse ERP data for visualization and statistical analysis, the regions and time intervals of interest were separately selected for the ERP components of interest, namely the CDA and the ACOP, as described below. Data from trials with incorrect responses were excluded from all the ERP analyses.

2.5.3. CDA analysis

The CDA was calculated as the difference in waveform amplitude recorded over the contralateral and the ipsilateral hemispheres using cued side as a reference, and time-locked to the onset of the memory array. For this purpose, ERP waveforms were collapsed across cued location (left/right) and hemisphere of recording (left/right) to obtain ERPs recorded over the contralateral hemisphere and over the ipsilateral hemisphere in relation to where the cue arrow pointed. The CDA is normally sampled in a set of parieto-occipital sites, but the specific electrodes vary across studies (e.g., Adam, Robison, Vogel, & 2018; Allon & Luria, 2019; Brady, Störmer, & Alvarez, 2016; Fukuda, Kang, & Woodman, 2016; Hakim, Adam, Gunseli, Awh & Vogel, 2019; Vogel & Machizawa, 2004). In the present study, the electrodes included in the CDA analysis were selected within the scalp area that was used in previous literature based on the visual inspection of the grand average waveforms across conditions. Here, the CDA was measured as the mean amplitude difference over 4 pairs of parieto-occipital electrode sites (PO7/PO8, PO3/PO4, O1/O2, P3/P4). Based on previous CDA literature (Hakim, Feldmann-Wüstefeld, Awh & Vogel, 2020), a 200 ms interval preceding the onset of the memory array served as a baseline, and statistical analyses were carried out over the 300–900 ms time window locked to memory array onset. This time window was chosen a priori to include the part of the retention interval that preceded the acoustic interruption. The CDA amplitudes were first analyzed in a repeated measures univariate ANOVA with factors Hemisphere (contralateral/ipsilateral to the cued side) and Load (low/high). Significant interactions were followed by post hoc tests. The total number of trials that were retained after artefact rejection in the low load condition is M= 275, SD = 24.1, while in the high load condition is M = 221, SD = 28.3.

2.5.4. ACOP analysis

The ACOP was calculated as the difference in waveform amplitude recorded over the contralateral and the ipsilateral hemispheres, using sound laterality as a reference. For ACOP preprocessing, epochs were time-locked to the sound onset. ERP waveforms were collapsed across auditory stimulus location (left/right) and hemisphere of recording (left/right) to obtain ERPs recorded on the contralateral hemisphere and on the ipsilateral hemisphere. The ACOP component was measured as the mean amplitude difference over 5 pairs of posterior electrode sites (P7/P8, PO3/PO4, P1/P2, P5/P6, P3/P4), and waveforms were corrected relative to a 100 ms pre-sound baseline. The set of sites included in the analysis was chosen from the parieto-occipital region, based on previous ACOP studies (e.g., Feng et al., 2014). The baseline period, similarly to the time window of interest, was selected to be consistent with previous literature (Feng et al., 2014; McDonald et al., 2013). The exact time interval for the ACOP analysis was then chosen based on when the ipsilateral and contralateral waveforms started to differentiate (based on the grand average waveforms across conditions), and it

corresponded to 280–500 ms. As our results did not differ between trials in which the auditory stimulus was delivered from the left or right speaker, we report the results collapsed between the two conditions. The averaged amplitudes were first analyzed in a repeated measures univariate ANOVA with factors Hemisphere (contralateral/ipsilateral to side of sound), Load (low/high), and Correspondence with cued side (same/different). Significant interactions were followed by post hoc tests. For data visualization, averages for each condition are accompanied by within-participant SEMs (Loftus & Masson, 1994), calculated following the procedure suggested by O'Brien and Cousineau (2014). The total number of trials per condition that were retained after artifact rejection was: low load, same hemifield M= 139, SD= 10.9; low load, different hemifield M= 135, SD= 12.7; high load, same hemifield M= 111, SD= 13.1; high load, different hemifield M= 111, SD= 14.1.

3. Results

3.1. Behavioral performance

Behavioral results are reported in Fig. 2. VWM task error rates were significantly higher in the high load (M = 21 %, SD = 7 %) than in the low load (M = 3 %, SD = 2 %) condition, Load $F(1, 17) = 176.86$, $p < .001$, $\eta^2_p = .912$, BF$_{excl} < .001$. However, we found no evidence that the acoustic distractor had an impact on error rates in the VWM task, as indicated by the lack of significant effect of Correspondence, or interaction between Correspondence and Load (ps >.455, BF$_{excl}$ = 2.664).

RTs were significantly slower in the high load (M = 978.46 ms, SD = 226.04) than in the low load (M = 748 ms, SD =194.91) condition, Load $F(1, 17) = 21.26$, $p < .001$, $\eta^2_p = .556$, BF$_{excl} = .027$. As for accuracy results, we found no evidence that the acoustic distractor had an impact on RTs in the VWM task. The effect of Correspondence, and the interaction between Correspondence and Load were not significant (ps >.53, BF$_{excl}$ = 2.994).

The estimated amount of information maintained in VWM using Pashler's K (Pashler, 1988) increased significantly from the low (K = .99 SD = .02) to the high (K = 2.84, SD = .53) VWM load condition (Fig. 3), Load $F(1, 17) = 166.18$, $p < .001$, $\eta^2_p = .907$, BF$_{excl} < .001$. The presence of an auditory distractor did not disrupt the maintenance of items in VWM, as suggested by the fact that we observed neither a significant effect of Correspondence, nor a significant interaction between Correspondence and Load (ps > .66, BF$_{excl}$ = 3.263).

3.2. CDA

As shown in Fig. 3, the ERP waveforms were more negative over the contralateral to the cued location compared to the ipsilateral
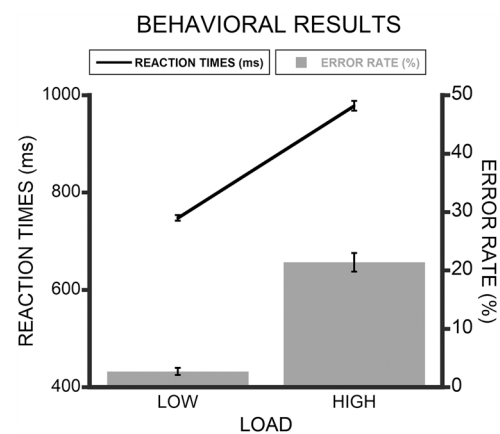

Fig. 2. Behavioral results for error rate (bars) and RTs (line), showing the effects of Load (Low vs. High) on performance. In each plot, error bars represent the SEM for within-participant designs.
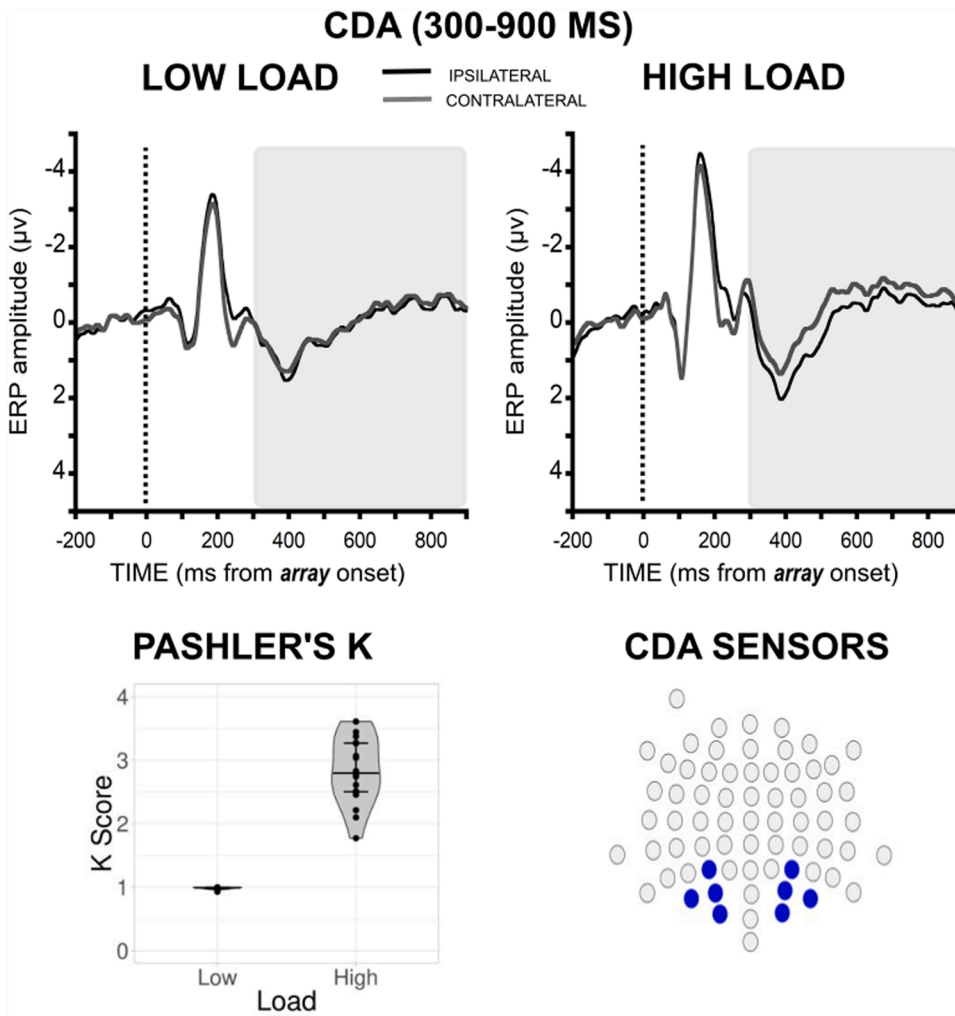
Fig. 3. CDA and behavioral results. The top row represents grand-averaged ERP waveforms for low and high load conditions analyzed in the CDA interval (300–900 ms) over 4 pairs of electrodes (colored in blue on the top-view electrode layout in the bottom right panel) placed ipsilaterally and contralaterally to the cued side of the display. The bottom left panel represents Pashler's K scores shown as a violin plot reflecting the data distribution. Results from each participant are shown as dots. A vertical bar indicates the 95 % confidence interval determined by bootstrapping for each median.

hemisphere during the CDA time interval (300 – 900 ms), as described by the main effect of Hemisphere $F(1, 17) = 7.94$, $p = .012$, $\eta^2_p = .32$, $BF_{excl} = .293$. A significant interaction was also observed between Hemisphere and Load $F(1, 17) = 8.44$, $p = .01$, $\eta^2_p = .33$, $BF_{excl} = .181$, with a more pronounced CDA (contralateral – ipsilateral difference) in the high load condition. Specifically, the contralateral waveform was significantly more negative compared to that of the ipsilateral hemisphere in the high load condition, Hemisphere $F(1,17) = 10.13$, $p = .005$, $\eta^2_p = .37$, $BF_{excl} = .143$, but not in the low load condition, Hemisphere $F(1,17) = 1.17$, $p = .296$, $\eta^2_p = .06$, $BF_{excl} = 2.045$. When comparing the contralateral waveforms in the low vs. high load conditions; as well as the ipsilateral in the low vs. high load conditions we observed no significant differences (all Load $ps > .071$, $BF_{excl} = .836$).

3.3. ACOP [2]

As shown in Figs. 4 and 5, the ERP waveforms were more positive over the contralateral than the ipsilateral hemisphere in terms of sound

location during the ACOP time interval (280 – 500 ms). The repeated measures univariate ANOVA we performed showed a main effect of Hemisphere, $F(1, 17) = 38.24$, $p < .001$, $\eta^2_p = .69$, $BF_{excl} = .001$, with a more pronounced positivity in the contralateral hemisphere compared to the ipsilateral one. The magnitude of this difference was modulated by the correspondence factor,[3] Hemisphere x Correspondence $F(1, 17) = 10.5$, $p = .005$, $\eta^2_p = .38$, $BF_{excl} = .415$. The amplitude of the contralateral vs. ipsilateral difference was more pronounced when the sound occurred in the same location as VWM contents, even though the ACOP was significant both when the sound occurred in same location as VWM contents, Hemisphere $F(1, 17) = 42.9$, $p < .001$, $\eta^2_p = .72$, $BF_{excl} < .001$, and when it occurred in the differing location, Hemisphere $F(1, 17) = 19.73$, $p < .001$, $\eta^2_p = .54$, $BF_{excl} = .014$. When comparing the amplitude of the ipsilateral waveform in the same vs. different conditions, as well as the contralateral waveforms in the same vs. different conditions, we observed no significant differences (all Correspondence $ps > .17$, $BF_{excl} = 1.891$). No significant interaction between Hemisphere and Load, $F(1, 17) = .3$, $p = .592$, $\eta^2_p = .02$, $BF_{excl} = 2.914$ was observed. The Bayes factor indicates that, albeit weak, the evidence is against the inclusion of the interaction in the model. The interaction between

---

[2] Further analyses were conducted to evaluate the N1 component. The ERP waveforms were more negative over the hemisphere contralateral vs. ipsilateral with respect to the location of the sound during the 80–150 ms time interval. The repeated measures univariate ANOVA we performed confirmed that the main effect of hemisphere was statistically significant, Hemisphere $F(1,17) = 15.165$, $p = .001$ $\eta^2_p = .471$, $BF_{excl} = .121$, but no significant interaction with Load or Cued side factors was observed (all $ps > .27$, $BF_{excl} = 16.382$).

---

[3] A sensitivity power analysis conducted using MorePower 6.0.4 (Campbell & Thompson, 2012), given an alpha value equal to .05, a power value of .8 and our sample size of 18, yielded a minimum detectable effect size of $\eta^2_p = .342$. This is lower than our reported effect size ($\eta^2_p = .382$), suggesting that the study had sufficient power.
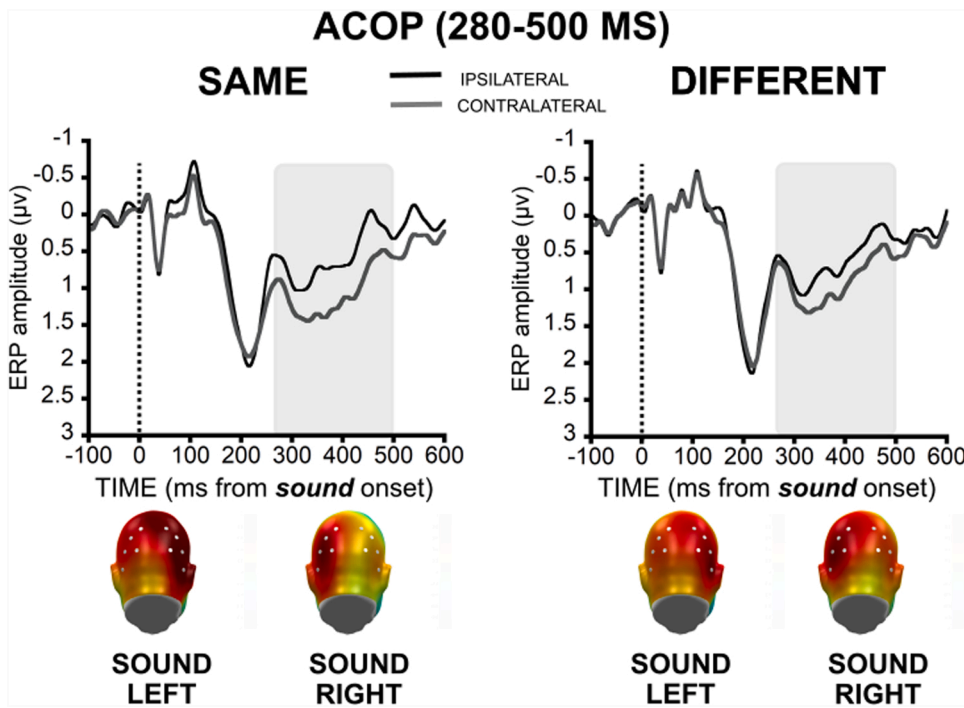
Fig. 4. Sound-elicited ACOP waveforms and scalp topographies. Grand-averaged ERP waveforms were averaged over 5 pairs of electrodes positioned ipsilaterally and contralaterally to the sound, marked in white over each topography. The ACOP interval (280–500 ms) is colored in gray. In the top left panel are waveforms that were averaged across trials in which the sound and cued locations corresponded (same). Below, scalp topographies of the ACOP interval are plotted separately for trials with left and right sounds. In the right panel, waveforms (top) and topographies (bottom), averaged across trials, where sound laterality and cued location did not correspond, are shown (different).
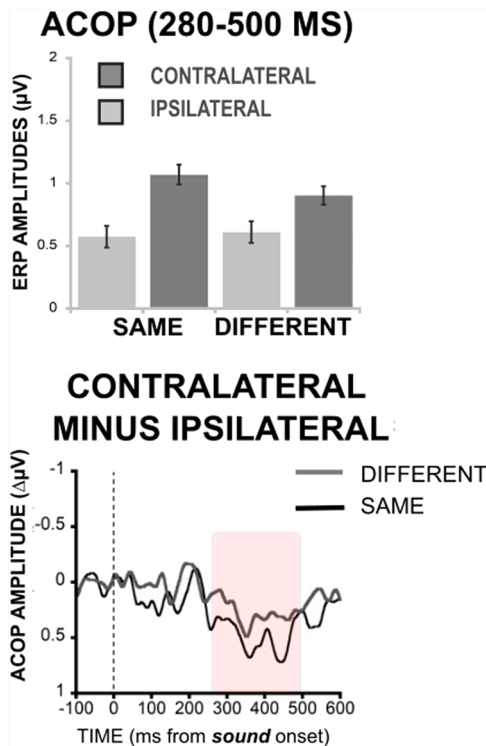

Fig. 5. Bar plot and the ACOP calculated as the contralateral minus ipsilateral positivity. The bar graph (top) shows ERP amplitude averages and within-participant SEM for ipsilateral and contralateral sites, in the same (left) and different (right) conditions. In the bottom panel, the ACOP modulation (contralateral-minus-ipsilateral) is represented and compared between the same and different conditions in the time window of interest (280–500 ms).

Hemisphere, Correspondence, and Load, $F(1, 17) = .07$, $p = .802$, $\eta^2_p = .004$, $BF_{excl} = 8.174$, was also not significant, with the Bayes factor providing evidence against the inclusion of the interaction in the model.

4. Discussion

In the present study, we investigated whether visual working memory load affects the visual cortex response to distracting sounds. Additionally, we examined if the auditory-evoked contralateral occipital positivity reflects a reflexive orienting to peripheral sounds or can be modulated by endogenous spatial attention. Peripheral sounds are especially effective at capturing attention, and recent research revealed that they could elicit activations of the contralateral occipital cortex, even when sounds are not relevant to the ongoing task (Hillyard, et al., 2016; Roye, Schroger, Jacobsen & Gruber, 2010; Spence & Driver, 1994). Our findings indicate that endogenous spatial attention affects the ACOP elicited by task-irrelevant acoustic distractors. At the same time, our results do not provide evidence for an effect of visual working memory load on the sound-elicited activation of visual cortices.

Previous studies reported that the ACOP could be elicited by lateralized sounds even when they were task-irrelevant, and participants were performing tasks on visual stimuli, or when they were presented in randomized sequence with central target tones (or bilateral target tones; McDonald et al., 2013). These findings led McDonald, Hillyard, and colleagues to suggest that task-irrelevant sounds automatically attract spatial attention to their location, and that spatial attention is inherently linked to visual processing (Hillyard et al., 2016; McDonald et al., 2013). However, as Hillyard and colleagues remarked, this hypothesis was at odd with the latency of the ACOP, which is unusually late (about 400 ms) to index an automatic process (Hillyard et al., 2016). It is well documented that changes in cortical activity with a latency larger than around 100 ms are attributed to attentive and post-attentive stages, which are usually influenced by contextual factors (Hillyard et al., 2016). Consistent with this observation, our findings showed that the ACOP is modulated by endogenous spatial attention. In the present study, sound-elicited activation of visual cortices was attenuated when participants were remembering visual items that were placed in the opposite location relative to sounds. Even though ACOP was not fully eliminated here, we report that goal-driven processes play a role (i.e., endogenous attention) in modulating its amplitude, making further research necessary to determine other goal-driven processes that could modulate cross-modal distraction elicited by acoustic events. Our

findings are consistent with other evidence that questioned the automaticity of sound-elicited activation of visual cortices, such as the finding that statistical regularities in the location of the sounds eliminated the ACOP modulation (Matusz, Retsa, & Murray, 2016). Specifically, task-irrelevant sounds affect visual processing and activate the visual cortex only if their location is unpredictable (Matusz et al., 2016). Moreover, in a recent study which used sounds as task-relevant targets, cross-modal activation of the visual cortex was not observed unless the top-down attention of participants was focused on the spatial location of the sounds (Retsa et al., 2020).

Differently from endogenous attention, we found no evidence that VWM load had an impact on ACOP amplitude. Even though our manipulation of set size was consistent with previous studies, and behavioral and CDA data indicate that VWM load was significantly modulated by set size, our results are rather consistent with load having no appreciable effect on acoustic distractor processing as reflected by the ACOP. Previous research has suggested that distractor processing is reduced when the VWM load is high (Konstantinou et al., 2014; Kim, Kim, & Chun, 2005; SanMiguel, Corral, & Escera, 2008), although a few recent studies failed to replicate this effect (Gil-Gómez de Liaño, Stablum, & Umiltà, 2016; Guo et al., 2019; Yao, Guo, Liu, Shen, & Gao, 2020). The observation of an effect of VWM on distraction may depend both on the modality of stimuli to be used for the memory and the distraction tasks, and on the type of behavioral task used to assess distraction. Concerning stimulus modality, most of the previous research was based on unimodal paradigms; here, stimuli were in different modalities, with visual memory stimuli and auditory distractors. One possibility is that visual working memory load might not tap on the resources that are involved in filtering auditory distractors. Concerning the type of task, studies differ in the way distraction is assessed, with most studies using a flanker task to infer distraction from interference. In flanker tasks, the magnitude of interference (the congruency effect) reflects the ability to reduce the impact of stimulus-response associations elicited by the flankers. As suggested by Buetti and colleagues, flankers are "candidate targets" because visually identical to one of the potential targets in the task (often only the location is different; Buetti, Lleras, & Moore, 2014; Lleras et al., 2013), and therefore the "flanker interference effect" (congruency effect) "should not be used to infer how stimuli that are entirely unrelated to a participant's main task would be processed" (Buetti et al., 2014). Consequently, studies examining distractor processing using flanker tasks have examined congruency effects (or "identity compatibility effects"; Folk, 2013) rather than genuine distraction by task-irrelevant stimuli. Differently, in the present study, we examined the impact of VWM load on the processing of task-irrelevant sounds. The reason why the reduction of the available visual WM resources failed to produce a modulation of the ACOP might be that processing the location where a sudden sound occurs, regardless of its task compatibility, is an efficient and prioritized process. Indeed, another efficient process, i.e. the extraction of spatial regularities is also not affected by VWM load (Gao & Theeuwes, 2020).

In summary, we observed that endogenous spatial attention modulates the activation of visual cortices to task-irrelevant sounds, challenging the hypothesis that the ACOP reflects an automatic process. It should be noted that the ACOP is attenuated but not eliminated when there is a mismatch in the spatial correspondence between the cued hemifield and sound location, suggesting that even in this condition, spatial attention is triggered by the distracting sound. One direction for future research might be extending these results to other distractor modalities, e.g. visual (Störmer, 2019), to understand whether they reflect a general characteristic of the visual system response, or a specific characteristic of acoustic-visual interaction. Here, we observed that the system remains receptive to sudden sounds even when the amount of available VWM resources is reduced and when endogenous spatial attention is directed to the opposite spatial position.

CRediT authorship contribution statement

Shari Cavicchi: Conceptualization, Investigation, Methodology, Formal analysis, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Andrea De Cesarei:** Conceptualization, Formal analysis, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Matteo Valsecch:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Maurizio Codispoti:** Conceptualization, Methodology, Data curation, Formal analysis, Funding acquisition, Supervision, Writing – original draft; Writing – review & editing.

Conflict of interest

Authors report no conflict of interest.

Data Availability

Experimental Data are available through the Zenodo repository (10.5281/zenodo.7327510)

References

Adam, K. C., Robison, M. K., & Vogel, E. K. (2018). Contralateral delay activity tracks fluctuations in working memory performance. *Journal of Cognitive Neuroscience, 30* (9), 1229–1240. https://doi.org/10.1162/jocn_a_01233

Allon, A. S., & Luria, R. (2019). Filtering performance in visual working memory is improved by reducing early spatial attention to the distractors. *Psychophysiology, 56* (5), Article e13323. https://doi.org/10.1111/psyp.13323

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences, 113*(27), 7459–7464. https://doi.org/10.1073/pnas.1520027113

Buetti, S., Lleras, A., & Moore, C. M. (2014). The flanker effect does not reflect the processing of "task-irrelevant" stimuli: Evidence from inat- tentional blindness. *Psychonomic Bulletin & Review, 21*, 1231–1237. https://doi.org/10.3758/s13423-014-0602-9

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavioral Research, 44*, 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Codispoti, M., De Cesarei, A., Biondi, S., & Ferrari, V. (2016). The fate of unattended stimuli and emotional habituation: Behavioral interference and cortical changes. *Cognitive, Affective and Behavioral Neuroscience, 16*, 1063–1073. https://doi.org/10.3758/s13415-016-0453-0

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*, 267–297. https://doi.org/10.1146/annurev.psych.48.1.269

Failing, M., & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin & Review, 25*(2), 514–538. https://doi.org/10.3758/s13423-017-1380-y

Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2014). Sounds activate visual cortex and improve visual discrimination. *Journal of Neuroscience, 34* (29), 9817–9824. https://doi.org/10.1523/JNEUROSCI.4869-13.2014

Folk, C. L. (2013). Dissociating compatibility effects and distractor costs in the additional singleton paradigm. *Frontiers in Psychology, 4*, 434.

Fukuda, K., Kang, M. S., & Woodman, G. F. (2016). Distinct neural mechanisms for spatially lateralized and spatially global visual working memory representations. *Journal of Neurophysiology, 116*(4), 1715–1727. https://doi.org/10.1152/jn.00991.2015

Gao, Y., & Theeuwes, J. (2020). Learning to suppress a distractor is not affected by working memory load. *Psychonomic Bulletin & Review, 27*(1), 96–104. https://doi.org/10.3758/s13423-019-01679-6

Gayet, S., Guggenmos, M., Christophel, T. B., Haynes, J. D., Paffen, C. L., Van der Stigchel, S., & Sterzer, P. (2017a). Visual working memory enhances the neural response to matching visual input. *Journal of Neuroscience, 37*(28), 6638–6647. https://doi.org/10.1523/JNEUROSCI.3418-16.2017

Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2017b). Visual working memory storage recruits sensory processing areas. *Trends in Cognitive Sciences, 22*, 189–190.

Gil-Gómez de Liaño, B., Stablum, F., & Umiltà, C. (2016). Can concurrent memory load reduce distraction? A replication study and beyond. *Journal of Experimental psychology: General, 145*(1), e1–e12. https://doi.org/10.1037/xge0000131

Guo, Y., Yao, N., Liu, Y., Gao, Z., Shen, M., & Shui, R. (2019). Visual working memory capacity load does not modulate distractor processing. *Journal of Vision, 19*(10), 103. https://doi.org/10.1167/19.10.103

Hakim, N., Adam, K. C., Gunseli, E., Awh, E., & Vogel, E. K. (2019). Dissecting the neural focus of attention reveals distinct processes for spatial attention and object-based storage in visual working memory. *Psychological Science, 30*(4), 526–540. https://doi.org/10.1177/0956797619830384

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing neural representations of working memory with task-irrelevant interruption. *Journal of Cognitive Neuroscience, 32*(3), 558–569. https://doi.org/10.1162/jocn_a_01481

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632–635. https://doi.org/10.1038/nature07832

Hillyard, S. A., Störmer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia, 83*, 170–178. https://doi.org/10.1016/j.neuropsychologia.2015.06.003

IBM CorpReleased (2011). *IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp.

JASP Team (2022). *JASP (Version 0.16.3)*.

Junghöfer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology, 37*(4), 523–532. https://doi.org/10.1111/1469-8986.3740523

Keefe, J. M., & Störmer, V. S. (2021). Lateralized alpha activity and slow potential shifts over visual cortex track the time course of both endogenous and exogenous orienting of attention. *Neuroimage, 225*, Article 117495. https://doi.org/10.1016/j.neuroimage.2020.117495

Kim, S. Y., Kim, M. S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America, 102*(45), 16524–16529. https://doi.org/10.1073/pnas.0505454102

Konstantinou, N., Beal, E., King, J. R., & Lavie, N. (2014). Working memory load and distraction: Dissociable effects of visual maintenance and cognitive control. *Attention, Perception, & Psychophysics, 76*(7), 1985–1997. https://doi.org/10.3758/s13414-014-0742-z

Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), 33267. https://doi.org/10.1525/collabra.33267

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental psychology: General, 133*(3), 339. https://doi.org/10.1037/0096-3445.133.3.339

Lleras, A., Buetti, S., & Mordkoff, J. T. (2013). When do the effects of distractors provide a measure of distractibility?. In *Psychology of Learning and Motivation* (pp. 261–315). Academic Press. https://doi.org/10.1016/B978-0-12-407187-2.00007-1.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490. https://doi.org/10.3758/BF03210951

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition, 29*(1), 1–21. https://doi.org/10.1080/13506285.2020.1848949

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews, 62*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003

MATLAB. (2015b). *The MathWorks, Inc*. Massachusetts, United States: Natick.

Matusz, P. J., Retsa, C., & Murray, M. M. (2016). The context-contingent nature of cross-modal activations of the visual cortex. *Neuroimage, 125*, 996–1004. https://doi.org/10.1016/j.neuroimage.2015.11.016

McDonald, J. J., Störmer, V. S., Martinez, A., Feng, W., & Hillyard, S. A. (2013). Salient sounds activate human visual cortex automatically. *Journal of Neuroscience, 33*(21), 9194–9201. https://doi.org/10.1523/JNEUROSCI.5902-12.2013

McDonald, J. J., Teder-Sälejärvi, W. A., Russo, F. D., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience, 15*(1), 10–19. https://doi.org/10.1162/089892903321107783

McDonald, J. J., & Ward, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science, 11*(2), 167–171. https://doi.org/10.1111/1467-9280.00233

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The Quantitative Methods for Psychology, 10*(1), 56–67.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*(4), 369–378. https://doi.org/10.3758/BF03210419

Peyk, P., De Cesarei, A., & Junghöfer, M. (2011). ElectroMagnetoEncephalography software: Overview and integration with other EEG/MEG toolboxes. *Computational Intelligence and Neuroscience*, 2011. https://doi.org/10.1155/2011/861705

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283–290. https://doi.org/10.3758/BF03203943

Retsa, C., Matusz, P. J., Schnupp, J. W., & Murray, M. M. (2020). Selective attention to sound features mediates cross-modal activation of visual cortices. *Neuropsychologia, 144*, Article 107498.

Roper, Z. J., & Vecera, S. P. (2014). Visual short-term memory load strengthens selective attention. *Psychonomic Bulletin & Review, 21*(2), 549–556. https://doi.org/10.3758/s13423-013-0503-3

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review, 18*(2), 324–330. https://doi.org/10.3758/s13423-011-0055-3

Roye, A., Schroger, E., Jacobsen, T., & Gruber, T. (2010). Is my mobile ringing? Evidence for rapid processing of a personally significant sound in humans. *Journal of Neuroscience, 30*, 7310–7313.

SanMiguel, I., Corral, M. J., & Escera, C. (2008). When loading working memory reduces distraction: Behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. *Journal of Cognitive Neuroscience, 20*(7), 1131–1145. https://doi.org/10.1162/jocn.2008.20078

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology, 118*(1), 98–104. https://doi.org/10.1016/j.clinph.2006.09.003

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

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 555–574.

Störmer, V. S. (2019). Orienting spatial attention to sounds enhances visual processing. *Current Opinion in Psychology, 29*, 193–198. https://doi.org/10.1016/j.copsyc.2019.03.010

Störmer, V. S., Feng, W., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2016). Salient, irrelevant sounds reflexively induce alpha rhythm desynchronization in parallel with slow potential shifts in visual cortex. *Journal of Cognitive Neuroscience, 28*(3), 433–445. https://doi.org/10.1162/jocn_a_00915

Tavares, T. P., Logie, K., & Mitchell, D. G. (2016). Opposing effects of perceptual versus working memory load on emotional distraction. *Experimental Brain Research, 234* (10), 2945–2956. https://doi.org/10.1007/s00221-016-4697-2

Teng, C., & Postle, B. R. (2021). Spatial specificity of feature-based interaction between working memory and visual processing. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000899

Van der Lubbe, R. H., & Postma, A. (2005). Interruption from irrelevant auditory and visual onsets even when attention is in a focused state. *Experimental Brain Research, 164*(4), 464–471. https://doi.org/10.1007/s00221-005-2267-0

Van der Lubbe, R. J., Havik, M. M., Bekker, E. M., & Postma, A. (2006). Task-dependent exogenous cuing effects depend on cue modality. *Psychophysiology, 43*(2), 145–160. https://doi.org/10.1111/j.1469-8986.2006.00393.x

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748–751. https://doi.org/10.1038/nature02447

Yao, N., Guo, Y., Liu, Y., Shen, M., & Gao, Z. (2020). Visual working-memory capacity load does not modulate distractor processing. *Attention, Perception, & Psychophysics, 82*(7), 3291–3313. https://doi.org/10.3758/s13414-020-01991-7