Cross-modal interactions between auditory attention and oculomotor control

Sijia Zhao[1], Claudia Contadini-Wright[2], Maria Chait[2]

1 Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, United Kingdom
2 Ear Institute, University College London, London, WC1X 8EE, United Kingdom

Abbreviated title: Auditory attention modulates microsaccadic inhibition

Corresponding Authors:
Sijia Zhao
sijia.zhao@psy.ox.ac.uk
Department of Experimental Psychology, New Radcliffe House, Walton Street, Oxford OX2 6AE, United Kingdom

Maria Chait
m.chait@ucl.ac.uk
Ear Institute, University College London, 332 Gray's Inn Road, London WC1X 8EE, UK

Number of figures: 4
Abstract length: 249
Significance Statement length: 120
Introduction length: 647
Discussion length: 1497

Conflicts of interest: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgements: We thank Mert Huviyetli for contributing to data collection. This work was supported by the NIHR UCLH BRC Deafness and Hearing Problems Theme. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Data sharing: The data reported in this manuscript alongside related information will be available on GitHub upon publication.

ORCID iD:
Sijia Zhao https://orcid.org/0000-0002-6246-0702
Maria Chait https://orcid.org/0000-0002-7808-3593

Keywords: attention; microsaccades; pupil dilation; pupil constriction

Abbreviations:
MS - Microsaccades
MSI - Microsaccadic inhibition
FEF - Frontal eye fields

SC - Superior colliculus
LC - Locus coeruleus
NE - Norepinephrine
PDR - Pupillary dilation response

Abstract

Microsaccades are small, involuntary eye movements that occur during fixation and play a crucial role in maintaining visual stability and attentional focus. Microsaccade inhibition (MSI) refers to the abrupt suppression of microsaccades, typically evoked within 0.1 second after new stimulus onset. The functional significance and neural underpinnings of MSI are subjects of ongoing research. It has been hypothesised that MSI is a component of the brain's attentional re-orienting network which facilitates the allocation of attention to new environmental occurrences by reducing disruptions or shifts in gaze that could interfere with processing.

The extent to which MSI is reflexive or influenced by top-down mechanisms remains debated. We developed a task that specifically examines the impact of auditory top-down attention on MSI, allowing us to disentangle ocular dynamics from visual sensory processing. Participants (N=24 and 27; both sexes) listened to two simultaneous streams of tones and were instructed to attend to one stream while detecting specific task "targets." We quantified MSI in response to occasional task-irrelevant events presented in both the attended and unattended streams (frequency steps in Experiment 1, omissions in Experiment 2).

The results show that initial stages of MSI are not affected by auditory attention. However, later stages (~0.25s post-event onset), affecting the extent and duration of the inhibition, are enhanced for sounds in the attended stream compared to the unattended stream. These findings provide converging evidence for the reflexive nature of early MSI stages and robustly demonstrate the involvement of auditory attention in modulating the later stages.

Significance Statement

Microsaccades are rapid eye movements occurring during fixation. Their precise role is not known but a major hypothesis is that they reflect automatic sampling of the environment. A feature of microsaccades is that they exhibit abrupt suppression (MSI) after the presentation of new stimuli. This is thought to be part of attentional re-orienting. To understand the neural circuit that controls MSI, and by extension, the brain's response to novel events in the environment, it is essential to determine which factors affect MSI. We show, for the first time, that auditory attention affects the latter (but not initial) stages of MSI. Thus, the early stages of MSI are automatic, but subsequent phases are affected by the perceptual state of the individual.

Introduction

Microsaccades (MS) are small involuntary fixational eye movements occurring at a rate of approximately 2-Hz. Initially believed to play a role in preventing visual fading during fixation (Martinez-Conde et al., 2004, 2006), evidence now suggests a more complex role in the unconscious continuous exploration of the environment (Otero-Millan et al., 2008; Benedetto et al., 2011; McCamy et al., 2014). Thus, understanding the perceptual processes influencing MS generation is vital for unravelling the brain mechanisms that underlie automatic scene analysis.

Gradual changes in sustained MS incidence have been consistently associated with the attentional load experienced by individuals (Pastukhov and Braun, 2010; Benedetto et al., 2011; Siegenthaler et al., 2014; Widmann et al., 2014; Dalmaso et al., 2017; Lange et al., 2017; Yablonski et al., 2017; Contadini-Wright et al., 2023). MS dynamics also exhibit fast changes, evoked by sudden stimuli. The presentation of a new visual stimulus triggers a rapid transient inhibition in MS activity (microsaccadic inhibition; MSI) around 0.10-0.15 sec post onset (Engbert and Kliegl, 2003; Rolfs et al., 2008; Valsecchi and Turatto, 2009; Wang et al., 2017; Xue et al., 2017). An abrupt sound can also elicit a MSI, although usually smaller in magnitude (Valsecchi et al., 2007; Rolfs et al., 2008; Wang et al., 2017; Wang and Munoz, 2021). The underpinnings of this response and its functional role remain relatively poorly understood. One hypothesis suggests that MSI represents a primitive attention re-orienting mechanism that interrupts ongoing processing, including eye movements, to enable the organism to quickly assess the behavioural relevance of a novel environmental stimulus and choose the best course of action. Supporting this idea, research has shown that the magnitude of MSI is influenced by the salience of the stimulus (Bonneh et al., 2015; Kadosh and Bonneh, 2022). Understanding the factors that affect this early response can provide critical insight into the intricate processes that govern the brain's response in fight or flight situations.

The neural circuits controlling MSI involve a network comprising the frontal eye fields (FEF), Superior colliculus (SC) and visual cortex (Hafed et al., 2009; Otero-Millan et al., 2011;

Martinez-Conde et al., 2013; Peel et al., 2016; Veale et al., 2017; Hsu et al., 2021). Although the specific contributions of different components of the network to MSI remain unresolved, recent evidence suggests that V1 lesions and inactivation of SC and FEF do not influence MSI (Hafed et al., 2021). Rather, the initial inhibition may be mediated by omnipause neurons (OPN) in a low-level circuit downstream from the SC (Hafed and Ignashchenkova, 2013; Hafed et al., 2021). Investigating the impact of top-down attention on MSI can provide insight into the reflexive nature of this circuit and its susceptibility to broader top-down information flow. However, the majority of existing research has primarily focused on the visual modality, leaving a critical gap in our understanding of how top-down auditory attention influences MSI, despite its potential to decouple the influences of visual processing from ocular dynamics.

To address this gap, we developed a paradigm (Figure 1) to examine the modulation of MSI by auditory attention. If MSI is predominantly driven by a low-level visual circuit in an autonomous and reflex-like manner (e.g.,Hafed et al., 2021), one would expect minimal effects of top-down attention in a non-visual modality. Alternatively, if MSI can be modulated by attention in the auditory domain, it would suggest that the circuits responsible for MSI generation receive inputs from higher-level brain systems. Exploring the temporal characteristics of any observed effects, whether occurring early or late, will help further pinpoint the nature of this interaction.

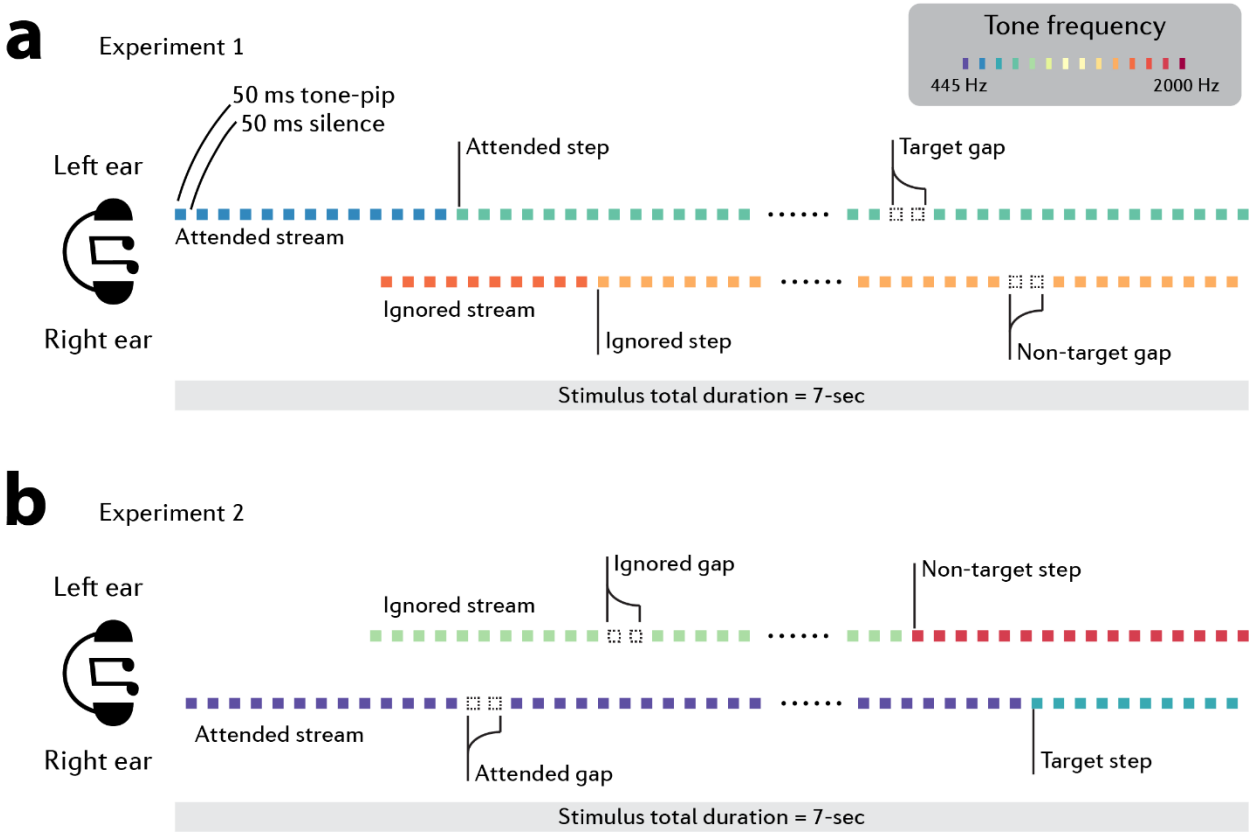


Figure 1. Experimental paradigm. *We designed a tightly controlled auditory paradigm to manipulate top-down attention. Two rapid streams of tones, differing in pitch, were presented simultaneously, one to each ear. The listener is required to selectively attend to one side and monitor for "target" events. The to-be-attended stream was indicated with an earlier onset time [a] in Experiment 1 listeners were instructed to detect brief silent gaps in the attended stream. Gaps were present, with equal probability, in both the attended and ignored stream, but only those in the attended stream were targets. Both attended and ignored streams additionally contained frequency steps - a salient change in tone-pip frequency. These events were always task irrelevant. Microsaccadic inhibition (MSI) evoked by step events in the attended stream ("attended step") were compared to those evoked by step events in the ignored stream ("ignored step") [b] in Experiment 2 stimuli were identical to those in Experiment 1 except that the 'target' events were now the frequency steps. The benefit of this paradigm is that the events of interest ("step" in Experiment 1 and "gap" in Experiment 2) are embedded in both attended and ignored streams rather than being explicitly attended. The events are physically identical, and all trials were blended to guarantee that there were no artefacts brought on by the different baselines in ocular responses.*

Materials and Methods

Ethics

The research was approved by the Research Ethics Committee of University College London. Participants were provided written informed consent and were paid for their participation.

Participants

We aimed for a sample size of ~20 participants, based on previous data indicating that this number is sufficient to obtain a stable measure of mean MSI (Zhao et al., 2019b). To allow for attrition due to noise/performance, we recruited 30 participants per experiment. In Experiment 1, two participants did not complete the session due to technical issues; one participant was excluded following the pre-processing stage due to a very low baseline incidence of microsaccades (fewer than 0.25 per second); two additional participants were rejected due to high false alarm rate (see below *Data Analysis – Behaviour*). Data from N=24 participants are reported (age 24.7, SD 5.2, min 18, max 39, 13 females and 11 males). In Experiment 2, three participants were excluded in the pre-processing stage due to low baseline incidence of microsaccades. Data from N=27 participants are reported (age 23.9, SD 4.3, min 19, max 34, 23 females and 4 males). All participants were naïve to the aims of the experiment, reported normal hearing and no history of neurological disorders.

Stimuli and Procedure

Experiment 1

The stimuli (Figure 1a; 7-second long) consisted of two, simultaneously presented (one to each ear) streams of 0.05-s (50ms) tone pips, separated by 0.05-s (50ms) inter-tone-intervals (ITI). Tone-pips were ramped on and off with a 0.005s (5ms) raised cosine ramp and their frequencies were selected from a pool of 14 fixed, logarithmically spaced values between 445 and 2000 Hz (12% steps).

Since temporal coherence is a strong binding cue (Elhilali et al., 2009; Shamma et al., 2011; Krishnan et al., 2014), to support the segregated ("two concurrent streams") percept, left and right tones were temporally interleaved such that a tone in one ear coincided with the ITI in the other ear. Furthermore, a frequency separation of at least 8 frequency pool steps between ears was maintained at all times. Taking a trial as an example, if the tone in the left ear was chosen to be 445 Hz, the tone in the right ear had to be higher than 1122 Hz. The same constraint applied to the trials with a step change (see below).

The tone-pips in each ear were arranged according to one of two frequency patterns, generated anew for each participant and each trial (Figure 1a). CONT sequences consisted of a single repeating tone, chosen by randomly selecting a frequency from the pool. STEP sequence consisted of a STEP transition from one repeating tone to another repeating tone of a different frequency; both frequencies were randomly drawn on each trial. The STEP could occur anywhere between 2 and 5 seconds after sequence onset. Therefore, a given trial could either contain two concurrent CONT sequences, a CONT and STEP sequence, or two concurrent STEP sequences (with the constraint that the steps in the right and left ears occurred at least 2 seconds apart).

Participants were naïve to these conditions and instead instructed to monitor one of the ears ("to-be-attended" stream) for brief silent gaps and indicate detection with a button press. The to-be-attended stream (determined quasi-randomly on each trial) started 1s before the ignored stream. Gaps were 0.15s long (two omitted tones plus 0.05s inter-tone interval), occurred equiprobably in both streams and could appear anywhere between 2 and 5s after stimulus onset. Therefore, to succeed in the task, participants had to focus attention on the to-be-attended stream and resist distraction from the other ("ignored") stream. Frequency step events were always task irrelevant. An example sound is available to download at https://shorturl.at/bcnU8 (listen with headphones).

The experiment started with a practice block, which consisted of 4 trials with a target gap (i.e., a gap in the to-be-attended stream), 4 trials with a distractor gap (i.e., a gap in the ignored stream), and 8 trials with no gap. All participants performed well in the practice and progressed to the main experiment.

The main experiment consisted of four blocks (8 minutes each). There were 32 trials per block for a total of 128 trials. The inter-trial interval was at least 6.5s, including 1.5s during which the visual feedback for each trial response was displayed (see below) and a minimum of 5s waiting time before playing the next trial.

In each block, 8 trials (25%) contained a gap. In four of those trials, the gap appeared in the cued ("to-be-attended") stream ("target"). In the others, it appeared in the ignored stream ("non-target"). Participants were instructed to press a keyboard button as soon as they detected the target gaps. Button presses that occurred within 2s after the target gap were considered a hit. Other button presses were considered false alarms. In particular, if a button press occurred within 2s after the non-target gap it was considered as a distractor-related false alarm. Most of the

participants achieved ceiling performance (see below *Data Analysis - Behaviour*). All trials which contained a gap or any response, were excluded from the eye movement/pupillometry analysis.

Of the remaining (no-gap) trials, 24 trials contained a STEP sequence in the to-be-attended stream and a CONT sequence in the ignored stream; 24 trials contained a CONT sequence in the to-be-attended stream and a STEP sequence in the ignored stream; 24 trials contained a STEP sequence in both streams (with steps occurring at least 2s apart); and 24 trials contained CONT sequences in both streams. All stimuli were presented in a random order, such that on each trial the specific condition was unpredictable.

Participants were instructed to fixate at a black centre cross "+" on a grey background throughout the experiment. At the end of each trial, visual feedback was given for the response of that trial; a blue circle "O" above the fixation cross indicated that the response was correct (correct rejection or hit), while a red cross "X" indicated an incorrect response (miss or false alarm). The visual feedback lasted 1.5s and was followed by an additional 5-second-long inter-trial interval. During the inter-trial interval, no sound or visual cue was presented, and the participants were instructed to rest and continue fixating. Further feedback was given at the end of each block, indicating the total number of correct responses, false alarms, and average response time. The experimental session—including introduction, practice and the main experiment—lasted one hour. A short break of a few minutes was imposed between blocks to reduce the effects of fatigue.

Participants sat with their head fixed on a chinrest in front of a monitor (24-inch BENQ XL2420T with a resolution of 1920x1080 pixels and a refresh rate of 60 Hz), in a dimly lit and acoustically shielded room (IAC triple-walled sound-attenuating booth). The distance between the chinrest and the screen was 62 cm. Sounds were delivered diotically to the participants' ears with Sennheiser HD558 headphones (Sennheiser, Germany) via a Roland DUO-CAPTURE EX USB Audio Interface (Roland Ltd, UK), at a comfortable listening level (self-adjusted by each participant). Stimulus presentation and response recording were controlled with Psychtoolbox (Psychophysics Toolbox Version 3; Brainard, 1997) on MATLAB (The MathWorks, Inc.).

Experiment 2

We further replicated the effect of attention on MSI with the same paradigm but looking at MSI evoked by omissions—silent gaps. These stimuli are interesting because the neural responses evoked by omissions, as measured through electroencephalogram (EEG) and direct neural recordings, often exhibit distinct characteristics compared to responses elicited by deviant tones (such as step events in our previous experiment) (Heilbron and Chait, 2018; Braga and

Schönwiesner, 2022). As a result, they may be associated with different patterns of MSI and potentially influenced differently by top-down attention. The stimuli and procedures for Experiment 2 (Figure 1b) were identical to those in Experiment 1, except the task significance of step and gap events was switched. The task now involved monitoring for frequency step events whilst MSI to gaps in the attended and non-attended ears were measured. The proportions of the two event-types (step and gap) were adjusted to mirror those in Experiment 1 such that each block contained 8 STEP trials (25%). In four of those trials, the step appeared in the to-be-attended ear. In the others, it appeared in the ignored stream. Participants were instructed to press a keyboard button as soon as they detected the step in the to-be-attended ear ("target"). All trials which contained a step, or any response were excluded from the eye movement/pupillometry analysis.

Of the remaining (non-step) trials, 24 trials contained a gap sequence in the to-be-attended ear and a CONT sequence in the ignored ear; 24 trials contained a CONT sequence in the to-be-attended ear and a gap sequence in the ignored ear; 24 trials contained a gap sequence in both ears (with gaps occurring at least 2 seconds apart); and 24 trials contained CONT sequences in both ears. All stimuli were presented in a random order, such that on each trial the specific condition was unpredictable.

Pupil recording

An infrared eye-tracking camera (Eyelink 1000 Desktop Mount, SR Research Ltd.) positioned just below the monitor continuously tracked gaze position and recorded pupil diameter, focusing binocularly with a sampling rate of 1000 Hz. The standard five-point calibration procedure for the Eyelink system was conducted prior to each experimental block. Participants were instructed to blink naturally. They were also encouraged to rest their eyes briefly during inter-trial intervals. Prior to each trial, the eye-tracker automatically checked that the participants' eyes were open and fixated appropriately; trials would not start unless this was confirmed.

Data Analysis

Behaviour

In Experiment 1, any response within a 2-second time window after the onset of a "target" gap was considered to be a hit or false positive for a "non-target". It was critical that participants perform well on the task (high hit rate, low false positive rate) as this was taken to indicate that attention was correctly allocated to the "to-be-attended" ear. Participants who did not achieve a hit rate of at least 70% and a false alarm rate below 30% were excluded from further analysis. This resulted in the exclusion of 2 participants from Experiment 1.

Preprocessing of pupil data

Where possible the right eye was analyzed. To measure the pupil dilation response (PDR) associated with the key events in the auditory streams (step events in Experiment 1 or gaps in Experiment 2), the pupil data from each trial were epoched from 0.2 second prior to step/gap onset to 2 second post the step/gap.

As mentioned previously, in both experiments, we had 24 trials containing an event in the to-be-attended ear, 24 trials containing an event in the ignored ear, 24 trials containing no event in either to-be-attended or ignored ears and 24 trials containing an event in both ears. After epoching, we then had 48 epochs of events in the attended stream and 48 epochs of events in the ignored stream. Matched no-event conditions were processed in a similar manner around dummy event times set to match those in the event conditions. Thus, each condition had 48 epochs per subject.

Intervals where the participant gazed away from fixation (> radius of 100 pixels around the centre of the fixation cross) or where full or partial eye closure was detected (e.g., during blinks) were automatically treated as missing data. Epochs with excessive missing data (>50%) were excluded from further analysis. For the rest, missing data were recovered with shape preserving piecewise cubic interpolation.

On average, approximately two trials per condition per participant were rejected. There was no effect of condition in the number of valid trials in Experiment 1 (Repeated Measures ANOVA, $F(1.2, 28.4)=0.82$, $p=0.40$, $\eta 2=0.034$; mean number of valid trials: control=45.1±1.6, attended event=44.4±1.5, ignored event=44.4±1.6) or Experiment 2 (Repeated Measures ANOVA, $F(1.1, 28.8) = 0.92$, $p=0.36$, $\eta 2=0.034$; mean number of valid trials: no gap=46.0±0.7, attended gap=45.5±0.6, ignored gap=45.7±0.5).

Microsaccade analysis

Microsaccade detection was based on the algorithm proposed by Engbert and Kliegl (2003). In short, microsaccades were extracted from the continuous horizontal eye-movement data based on the following criteria: (a) a velocity threshold of $\lambda = 6$ times the median-based standard deviation within each block; (b) above-threshold velocity lasting for longer than 0.005s but less than 0.1s; (c) the events are binocular (detected in both eyes) with onset disparity less than 0.01s; and (d) the interval between successive microsaccades is longer than 0.05s.

For deriving the microsaccade rate time series, a causal smoothing kernel was applied to each epoch with a decay parameter of $\alpha = 1/50$ ms (Dayan et al., 2005; Rolfs et al., 2008;

Widmann et al., 2014), paralleling a similar technique for computing neural firing rates from neuronal spike trains (Dayan et al., 2005; Rolfs et al., 2008; Joshi et al., 2016). The obtained time series was shifted by 0.05s (the peak of the convoluted curve) and baseline corrected by subtracting the mean microsaccade rate over 0.2s pre-event interval. Mean microsaccade rate time series, obtained by averaging across epochs for each participant and then averaging across participants, are reported below.

Pupil diameter analysis

To allow for comparison across conditions and subjects, data for each subject in each block were normalized. To do this, the mean and standard deviation across all data points in that block were calculated and used to z-score normalize all data points in the block. A baseline correction was then applied by subtracting the mean pupil size over the pre-onset period; subsequently, data were smoothed with a 0.15s Hanning window. For each participant, pupil diameter was time-domain averaged across all epochs to produce a single time series per condition.

Pupil dilation and constriction incidence rate analysis

Pupil event rate analysis compared the incidence of pupil dilation or constriction events. Following Joshi et al. (2016), events were defined as local minima (dilations; PD) or local maxima (constrictions; PC) with the constraint that continuous dilation or constriction is maintained for at least 0.1s. The pupil events were extracted from the continuous data smoothed with a 0.15s Hanning window. The rate was estimated for each epoch by convolving with an impulse function (Rolfs et al., 2008; Joshi et al., 2016; Zhao et al., 2019a; Milne et al., 2021) in the same way that microsaccade rate was computed (see above).

Time-series statistical analysis

To identify time intervals where a given pair of conditions exhibited differences in microsaccade rate/pupil diameter/pupil dilation rate/pupil constriction rate, a non-parametric bootstrap-based statistical analysis was used (Efron and Tibshirani, 1994). Using the average pupil diameter at each time point, the difference time series between the conditions was computed for each participant and these time series were subjected to bootstrap re-sampling (1000 iterations: with replacement). At each time point, differences were deemed significant if the proportion of bootstrap iterations that fell above or below zero was more than 95% (i.e., $p < 0.05$). Any significant differences in the pre-onset interval would be attributable to noise, therefore the largest number of consecutive significant samples pre-onset was used as the threshold for the statistical analysis for the entire epoch.

Results

Frequency step-induced microsaccadic inhibition is modulated by attention (Experiment 1)

To confirm that the listeners successfully directed their attention to the to-be-attended stream we first examined their behavioral performance. Figure 2a confirms that participants followed the instructions: listeners accurately and quickly detected the target gap in the to-be-attended stream (hit rate = 93.8±1.6%, reaction time = 0.86±0.05s) and successfully ignored the distractor gap in the ignored stream (false alarm rate to gap in Experiment 1 = 5.6±1.5%). Most importantly, no participant made any false alarms to step.

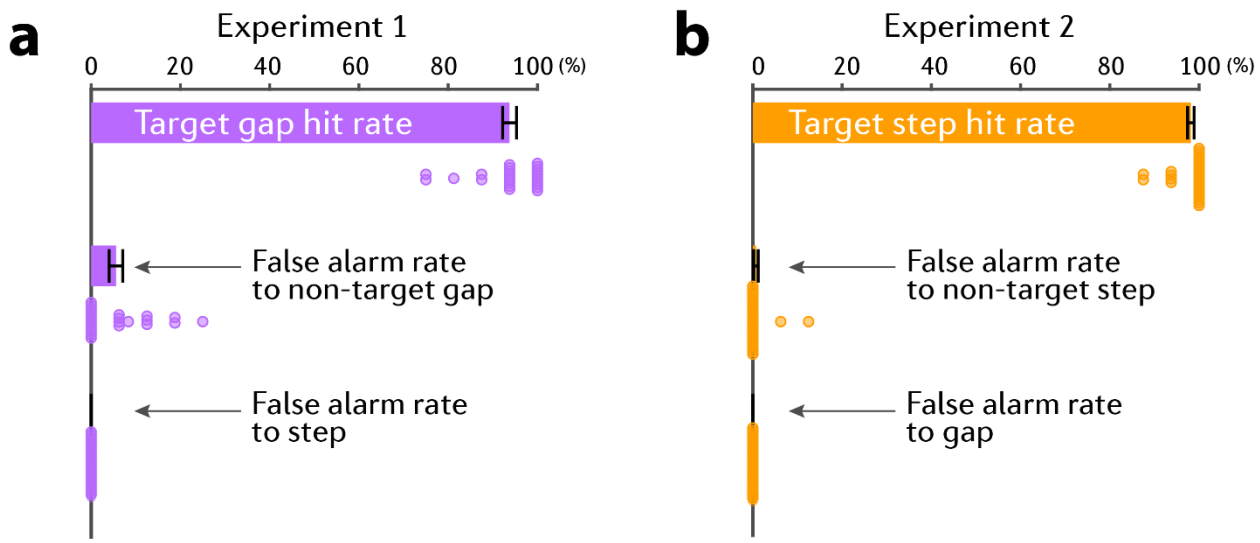


Figure 2. Behavioral performance. *Error bars are ±1 SEM. Coloured circles represent individual participant data. There was no statistical difference between experiments in any of the measures.*

While listening to the dichotic tone-pip streams, the baseline microsaccadic rate was around 1.27±0.12 Hz. The baseline-corrected time course of the microsaccadic rate is shown in Figure 3a. Three conditions are plotted: MSI to step events in the to-be-attended stream, MSI to STEP events in the ignored stream and the control condition (CONT). Responses to step in the attended and ignored streams exhibited a rapid, simultaneous drop in the microsaccades rate (MSI) which started around 0.07s and became significant (as assessed by comparing statistically against the control condition) from about 0.15s post-step. The MSI in the attended stream reached its minimum at 0.27s after the onset of the step event. The MSI to step in the ignored stream reached a much shallower through (-0.52 events per second) than the step in the attended stream

(-0.68 events per second). This distinction is especially clear when comparing the time courses of the two conditions: a significant difference emerged at 0.26s and persisted until 0.4s after the step (black horizontal line in Figure 3a). This provides direct evidence that abrupt frequency step event evoked MSI is modulated by attention.

We analyzed three aspects of pupil responsivity to compare to MSI effects. (a) Pupil size: is a widely used measure of LC-NE mediated arousal. Joshi et al. (2016) demonstrated that the trial-by-trial average pupil size correlated with the mean spiking activity in the LC. This suggests that pupil size can serve as an indicator of LC activity. (b) Pupil dilation rate: quantifies the incidence of pupil dilation events. Joshi et al (2016) highlighted the direct link between spiking activity in the LC, SC and inferior colliculus and pupil dilation, with spikes in all three areas correlated with subsequent pupil dilation events. Incidentally SC spikes triggered the fastest pupil dilation. Consequently, if attention-related MSI effects are mediated by the SC, we might expect to observe a similar pattern in the pupil dilation rate. (c) While less researched, pupil constriction rate may reflect activity in the parasympathetic nervous system (PNS), which is responsible for controlling pupil constriction. The PNS, along with the prefrontal cortex, plays a role in top-down regulation of attention. It aids in inhibiting irrelevant or distracting stimuli and maintaining focus on relevant tasks. PNS activity, potentially revealed through pupil constriction, is associated with increased cognitive control and reduced susceptibility to distraction.

The pupil dilation response (PDR, Figure 3b) started around 0.3s and became significant from around 0.5s after step onset. This response to step in the attended and ignored streams diverged shortly thereafter, with the PDR to step in the to-be-attended stream eliciting a larger pupil dilation response. The difference between the two attention conditions was significant between 0.53s and 0.87s post onset. Note that the MSI response had completely subsided by that stage.

To further examine pupil response dynamics, we separately analyzed pupil dilation incidence rate and pupil constriction incidence rate. The likelihood of pupil dilation is known to be closely connected with the firing of noradrenergic neurons in the LC and SC (Joshi et al., 2016; Megemont et al., 2022). Therefore, it is conceivable that events in the attended stream might evoke a stronger arousal/reorienting response which might be revealed in increased pupil dilation rate. As shown in Figure 3c, the pupil dilation incidence started to increase after 0.16s and reached its peak 0.36 second after the step event onset; This roughly corresponds to the timing of the MSI dynamics discussed above. From 0.6 to 1.4 seconds after the step event, the pupil dilation incidence rate dropped below baseline with a small rebound thereafter. Remarkably, there

was no difference between the attention conditions. This suggests that though attention affected pupil diameter between 0.5-0.9s after step onset, this effect was not underpinned by the incidence rate of pupil dilation events.

A different pattern was observed for the analysis of the pupil constriction rate (Figure 3d). Pupil size depends on the interplay between antagonistic sympathetic impulses (norepinephrine, NE) acting on the pupil dilator muscle and parasympathetic impulses (acetylcholine, ACh) acting on sphincter muscles which causes pupil constriction. ACh has been hypothesized to play a role in focused attention and distractor suppression (Sarter et al., 2001). As expected, overall constriction dynamics showed a complimentary profile to that seen for the pupil dilation incidence rate (Figure 3c) – exhibiting throughs that are temporally coincident with the pupil dilation rate peaks observed in Figure 3c. Notably, a clear significant difference between the attention conditions is seen around 0.69s; with step in the to-be-attended stream being approximately 30% more likely to evoke a pupil constriction event than that in the ignored stream. The constriction effect is also visible in the pupil diameter data (Figure 3b), manifested as a sharper drop in pupil dimeter in the attended condition.

Overall, the pattern of results suggests that frequency step events presented in the attended stream were associated with increased MSI (between 0.25~0.4s following event onset), increased pupil dilation (between 0.2~0.5s following event onset) and increased pupil constriction rate (between 0.5~0.9ms). We discuss this pattern of results below.

Figure 3. Microsaccade and pupillometry data of Experiment 1 (n = 24). *Solid lines represent the average microsaccade rate (a), change in pupil diameter evoked by step onset (b), change in pupil dilation rate evoked by step onset (c), and change in pupil constriction rate evoked by step onset (d). All baseline corrected to the pre-step interval. In all plots of this figure, the shaded area shows ±1 SEM. Colour-coded horizontal lines indicate the time interval where bootstrap resampling confirmed a significant difference between each step condition (dark or light purple) and the no-step control (grey). The black horizontal line indicates when the response to attended step (dark purple) significantly differs from ignored step (light purple). The green shaded area marks the microsaccadic inhibition (MSI) interval, spanning from 0.1 to 0.4 seconds post step onset.*

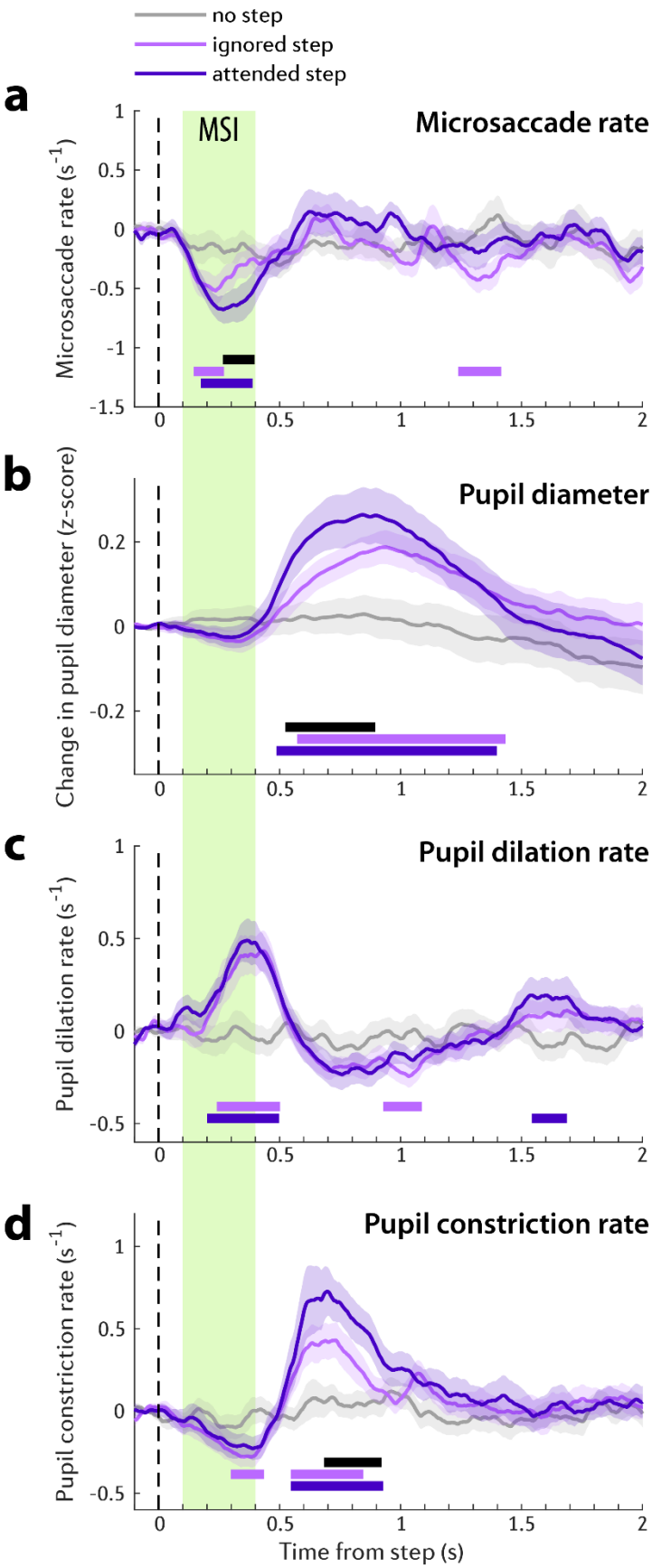
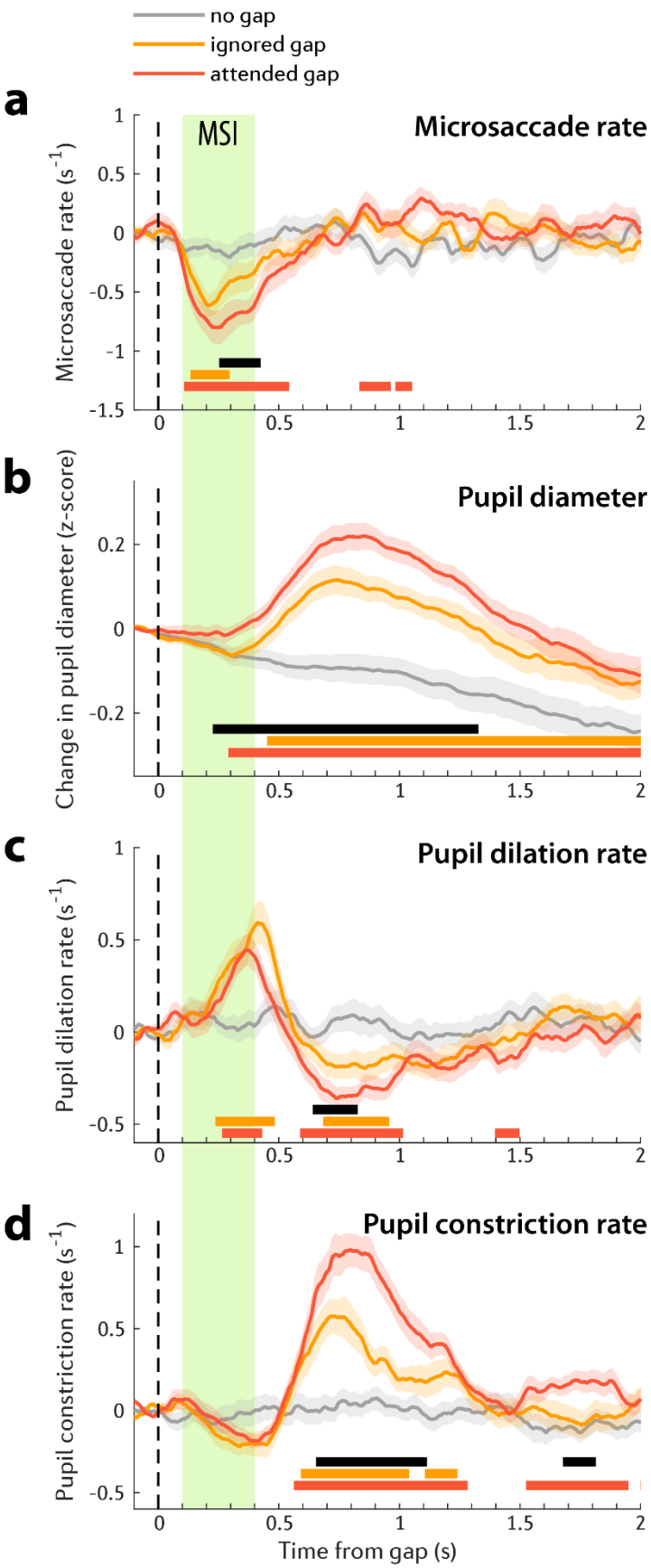


Figure 4. Microsaccade and pupillometry data of Experiment 2 (n = 27). *Solid lines represent the average microsaccade rate (a), change in pupil diameter evoked by gap onset (b), change in pupil dilation rate evoked by gap onset (c), and change in pupil constriction rate evoked by gap onset (d). All baseline corrected to the pre-gap interval. In all plots of this figure, the shaded area shows ±1 SEM. Colour-coded horizontal lines indicate the time interval where bootstrap resampling confirmed a significant difference between each gap condition (dark or light orange) and the no-gap control (grey). The black horizontal line indicates when the response to attended gap (dark orange) significantly differs from ignored gap (light orange). The green shaded area marks the microsaccadic inhibition (MSI) interval, spanning from 0.1 to 0.4 seconds post gap onset.*



The attentional modulation on MSI is not limited to the physical presence of a stimulus (Experiment 2)

Experiment 2 sought to replicate the results of Experiment 1 but asking whether similar response dynamics would be evoked by silent gaps as opposed to frequency steps. The same stimuli and paradigm were used except frequency steps were now behavioral targets. We analysed microsaccade and pupil dilation dynamics in response to silent gaps occurring in the to-be-attended vs ignored streams. Omission events provide a compelling avenue of investigation in this context due to their distinct characteristics compared to deviants (i.e., step events in Experiment 1). Specifically, EEG and neural responses to omitted stimuli often exhibit later peak latencies and lower amplitudes compared to responses evoked by deviant sounds (Heilbron and Chait, 2018; Braga and Schönwiesner, 2022). These differences suggest the possibility of observing a distinct MSI signature in relation to auditory omissions.

It is also worth noting that auditory omission responses are primarily observed in the cortex (Auksztulewicz et al., 2023; Lao-Rodríguez et al., 2023) with limited evidence of such responses in the IC or brain stem (Lehmann et al., 2016). This implies that any potential influence of attention on omission-evoked MSI would need to be conveyed through cortical circuits.

Figure 2b shows the behavioral performance (frequency step detection). Listeners detected the target step accurately (hit rate=98.1±0.7%) and quickly (reaction time = 0.77±0.05s), successfully ignored the non-target step in the ignored stream (false alarm rate to distractor=0.7±0.5%) and did not respond to any gaps (false alarm rate to gap = 0).

The Microsaccade/Pupillometry result pattern observed in Experiment 1 was fully replicated in Experiment 2. Figure 4a depicts the microsaccadic rate in attended and ignored streams following the commencement of a 0.15s gap. Responses to gap in the attended and ignored streams exhibited a significant MSI at about 0.11s post-gap-onset. The MSI in the attended stream reached its minimum at 0.25s after gap onset. The MSI to gaps in the ignored stream reached a much shallower through (-0.61 events per second relative to the baseline) than the gaps in the attended stream (-0.80 events per second). This distinction is especially clear when comparing the time courses of the two conditions: a significant difference emerged at 0.26s and persisted until 0.43s after the gap (black horizontal line in Figure 4a).

PDR emerged from roughly 0.3 seconds after gap onset (Figure 4b). PDRs to gaps in the attended and ignored streams diverged shortly thereafter, with the PDR to gap in the to-be-

attended stream eliciting a larger pupil dilation response. The difference between two attention conditions lasted till 1.32s post-gaps.

Similar to Experiment 1, there was no difference in pupil dilation rate between attention conditions during the major pupil dilation rate peak (between 0.2-0.5 s post gap onset, Figure 4c). We did observe a later difference between conditions (between 0.65-0.82 s) where pupil dilation rate evoked by the gap in the to-be-attended stream was smaller, than that for gap in the ignored stream. This effect appears to be linked to the increased pupil constriction rate (Figure 4d). Indeed, as in Experiment 1, a robust difference between attention conditions is seen in the pupil constriction rate data, with gaps in the to-be-attended stream associated with substantially larger pupil constriction incidence in the interval between 0.66 and 1.11s post gap onset.

Discussion

MSI is affected by auditory attention

Microsaccadic inhibition (MSI) refers to a rapid and brief decrease in the occurrence of microsaccades following sudden sensory stimuli (Rolfs et al., 2008; Rolfs, 2009). Traditionally, MSI has been thought to originate from a primitive sensory circuit, serving as an interrupt process that halts ongoing activities to facilitate attentional shifts towards abrupt sensory events. However, recent evidence suggests that MSI is not automatic but influenced by visual salience (Bonneh et al., 2015; Wang et al., 2017; Roberts et al., 2019), and conscious awareness (White and Rolfs, 2016). Despite the significance of microsaccades in indexing visual attentional sampling, our understanding of how the auditory system interfaces with the MSI-generating network remains limited.

Previous studies have established that auditory salience can modulate MSI (Zhao et al., 2019b), but it is unclear whether audio-evoked MSI modulation arises from a local, bottom-up, circuit specifically wired for sensitivity to low-level acoustic features that convey salience (e.g. roughness, Zhao et al., 2019b) or if it is also subject to top-down control. Some evidence for the latter is available through demonstrations that auditory oddball stimuli - a deviant sound in a sequence of standard sounds - can trigger enhanced MSI (Valsecchi and Turatto, 2009; Widmann et al., 2014; Kadosh and Bonneh, 2022). To explore the relationship between auditory-evoked MSI and attention, we targeted top-down attention by designing a task (Figure 1a) where participants listened to concurrent "attended" and "ignored" streams. We then analysed MSI evoked by non-behaviourally relevant events in both streams. This allowed us to control overall vigilance levels and isolate the effects of top-down attention. Notably, we used two types of MSI

triggering events: frequency steps (Experiment 1) and silent gaps (Experiment 2). Whilst MSI to frequency steps may arise from a relatively low-level IC-SC circuit, omission responses are not usually observed in the IC (Auksztulewicz et al., 2023; Lao-Rodríguez et al., 2023), therefore MSI to silent gaps must likely involve a cortical contribution.

We reveal a robust attentional effect on MSI, across both trigger types, with larger and more prolonged MSI responses observed for events within attended streams compared to those in ignored streams. A consistent attentional modulation was seen for step- and gap- evoked MSI, both in terms of timing and extent. This attentional effect emerged approximately 0.25s after event onset, in line with the timing observed in the oddball studies (Valsecchi and Turatto, 2009; Kadosh and Bonneh, 2022). Therefore, auditory attention, whether driven by bottom-up (as in the oddball work) or top-down mechanisms (in the present study), influences the later stage of MSI rather than the initial inhibition phase.

This observation is consistent with recent developments in the understanding of the network that supports microsaccade generation. SC has long been believed to be the primary regulator of MSI (Hafed et al., 2009). However, recent data indicate that the SC (and FEF) may not be causally involved in MSI (Hafed and Ignashchenkova, 2013; Hafed et al., 2021). Instead, emerging evidence suggests that the earliest MSI effects (initial inhibition) originate downstream of the SC and that subsequent processes, within a broader network likely encompassing the SC and FEF, determine the degree and duration of MSI. The present results add to this evolving understanding by providing further evidence for the reflexive nature of the early stage of MSI.

While the SC is traditionally known for its contribution to visual processing, it also receives auditory inputs from the IC and is involved in auditory processing, including sound localization, orienting responses to auditory stimuli, and integration of auditory information with visual and spatial cues (Meredith and Stein, 1986; Meredith et al., 1987; King et al., 1996; King, 2004; Ito and Feldheim, 2018; Hu and Dan, 2022). It is plausible that the attentional effects observed here are mediated, to some extent, through the connection between the SC and the auditory system. However, we note that the generation of omission-evoked MSI, as mentioned earlier, likely involves additional contributions from cortical regions. A plausible route is via the FEF. Traditionally known for their strong projections to the SC and their role in oculomotor control and visual attention, there is emerging evidence suggesting that the FEF may also play a role in auditory attention. Studies have shown that attention to auditory stimuli can modulate FEF activity and functional connectivity (Lee et al., 2013; Braga et al., 2016). Additionally, as part of the dorsal attention network (Corbetta et al., 2008), the FEF is activated during the maintenance of top-down

attention to spatial locations regardless of modality (Kastner et al., 1999; Braga et al., 2016). Whilst physiological recordings of microsaccade-related neural firing in the FEF are lacking, it has been shown that cooling the FEF affected MSI properties (Peel et al., 2016; but see Hsu et al., 2021). Investigating, the FEF and SC's role in the context of auditory influences on microsaccades can provide valuable insights into the neural mechanisms underlying the integration of auditory and visual attention.

MSI evoked by sound omission

In Experiment 2, we specifically examined MSI in response to silent gaps. To the best of our knowledge, this is the first report of MSI that is elicited by omission (but see Watanabe et al, 2014).

We introduced silent intervals between tones, ensuring that the MSI was triggered by the absence of an anticipated sound onset rather than the presence of a sound offset. Notably, the omission-induced MSI observed in Experiment 2 was not different from the step-event-induced MSI observed in Experiment 1, indicating that the presence of a new stimulus is not necessary for MSI. Instead, MSI, as an orienting response, is sensitive to salient changes in the sensory environment including the non-arrival of expected events.

Ocular dynamics reveal a sequence of attentional processes

Pupil responsivity is a primary measure of attentional engagement, much more researched than MS (Aston-Jones and Cohen, 2005; Bradley et al., 2008; Nassar et al., 2012; Dercksen et al., 2023). It has been used as a proxy of instantaneous arousal, e.g., as triggered by surprising events, and hypothesized to reflect activity in the locus coeruleus-norepinephrine (LC-NE) system, the brain’s main arousal regulator.

LC is interconnected with the FEF and SC (Matsumoto et al., 2018), together forming part of a network involved in attention, sensory processing, and arousal regulation (Joshi et al., 2016; Wang and Munoz, 2021). However, the precise functional roles and mechanisms of interaction between these structures are still an active area of research and further investigation is needed to fully understand their interplay.

The results of our analysis showed interesting patterns in the pupil dilation response and pupil constriction rate in relation to attentional processes. Firstly, events in the attended stream elicited a greater pupil dilation response, indicating increased arousal compared to events in the unattended stream. Notably, the attention effect on MSI (0.25~0.4s after the event) emerged significantly earlier than in pupil size (after 0.5s after event). However, this delay is difficult to

interpret. It may have functional significance or could be attributed to the slower pathway from the LC to the pupil musculature.

Next, we explored whether the attention effect on pupil dilation response was driven by an overall increase in absolute pupil size or an increase in the number of pupil dilation events. Interestingly, we did not observe any evidence of attentional modulation on the pupil dilation rate, suggesting that attention did not influence the occurrence of pupil dilation events but rather resulted in larger pupil sizes. Despite previous demonstrations that IC and SC control pupil dilation rate, the fact that no modulation was seen here may be interpreted to suggest that the attention effects observed for the MSI data are driven by a circuit that bypasses the SC (e.g. directly via FEF).

Furthermore, pupil constriction rate increased significantly in the attended stream around 0.65s after the event began. This observation is noteworthy because it has rarely been reported, predominantly due to the fact that pupil constriction rates are rarely studied. The attentional effect on pupil constriction may reflect resistance to non-target distraction in the attended stream. When individuals are successfully resisting distraction, their pupils respond less to irrelevant stimuli (Laeng et al., 2011; Hsu et al., 2020). In other words, the pupil shows a smaller dilation in response to distractors when attention is effectively maintained on the task at hand.

Taken together, the data reveal a sequence of attentional events, starting from approximately 0.2s after event onset in the context of MSI reflecting increased attentional capture by auditory events in attended relative to ignored stream. This is followed by increased arousal, starting around 0.5 seconds and subsequently increased rate of pupil constriction in the attended stream, likely reflecting active resistance of distraction by the task irrelevant events.

Overall, the findings presented here offer compelling evidence for the role of auditory attention in modulating microsaccade inhibition (MSI). These results contribute to our growing understanding of the neural network involved in microsaccade generation and shed light on the intricate interplay between attentional capture, as reflected by MSI, and the modulation of arousal, as indexed by pupil size.

References

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450.

Auksztulewicz R, Rajendran VG, Peng F, Schnupp JWH, Harper NS (2023) Omission responses in local field potentials in rat auditory cortex. BMC Biology 21:130.

Benedetto S, Pedrotti M, Bridgeman B (2011) Microsaccades and Exploratory Saccades in a Naturalistic Environment. Journal of Eye Movement Research 4 Available at: https://bop.unibe.ch/JEMR/article/view/2318 [Accessed May 23, 2018].

Bonneh YS, Adini Y, Polat U (2015) Contrast sensitivity revealed by microsaccades. Journal of Vision 15:11.

Bradley MM, Miccoli L, Escrig MA, Lang PJ (2008) The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology 45:602–607.

Braga A, Schönwiesner M (2022) Neural Substrates and Models of Omission Responses and Predictive Processes. Frontiers in Neural Circuits 16 Available at: https://www.frontiersin.org/articles/10.3389/fncir.2022.799581 [Accessed July 5, 2023].

Braga RM, Fu RZ, Seemungal BM, Wise RJS, Leech R (2016) Eye Movements during Auditory Attention Predict Individual Differences in Dorsal Attention Network Activity. Frontiers in Human Neuroscience 10 Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2016.00164 [Accessed June 29, 2023].

Contadini-Wright C, Magami K, Mehta N, Chait M (2023) Pupil dilation and microsaccades provide complementary insights into the dynamics of arousal and instantaneous attention during effortful listening. J Neurosci Available at: https://www.jneurosci.org/content/early/2023/04/27/JNEUROSCI.0242-23.2023 [Accessed June 24, 2023].

Corbetta M, Patel G, Shulman GL (2008) The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron 58:306–324.

Dalmaso M, Castelli L, Scatturin P, Galfano G (2017) Working memory load modulates microsaccadic rate. Journal of Vision 17:6–6.

Dayan P, Sejnowski TJ, Poggio TA (2005) Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems, Revised ed. edition. Cambridge, Mass.: MIT Press.

Dercksen TT, Widmann A, Wetzel N (2023) Salient omissions-pupil dilation in response to unexpected omissions of sound and touch. Front Psychiatry 14:1143931.

Efron B, Tibshirani RJ (1994) An Introduction to the Bootstrap. New York: Chapman and Hall/CRC.

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61:317–329.

Engbert R, Kliegl R (2003) Microsaccades uncover the orientation of covert attention. Vision Research 43:1035–1045.

Hafed ZM, Goffart L, Krauzlis RJ (2009) A Neural Mechanism for Microsaccade Generation in the Primate Superior Colliculus. Science 323:940–943.

Hafed ZM, Ignashchenkova A (2013) On the Dissociation between Microsaccade Rate and Direction after Peripheral Cues: Microsaccadic Inhibition Revisited. J Neurosci 33:16220–16235.

Hafed ZM, Yoshida M, Tian X, Buonocore A, Malevich T (2021) Dissociable Cortical and Subcortical Mechanisms for Mediating the Influences of Visual Cues on Microsaccadic Eye Movements. Frontiers in Neural Circuits 15 Available at: https://www.frontiersin.org/articles/10.3389/fncir.2021.638429 [Accessed March 17, 2023].

Heilbron M, Chait M (2018) Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? Neuroscience 389:54–73.

Hsu T-Y, Chen J-T, Tseng P, Wang C-A (2021) Role of the frontal eye field in human microsaccade responses: A TMS study. Biological Psychology 165:108202.

Hsu Y-F, Baird T, Wang C-A (2020) Investigating cognitive load modulation of distractor processing using pupillary luminance responses in the anti-saccade paradigm. Eur J Neurosci 52:3061–3073.

Hu F, Dan Y (2022) An inferior-superior colliculus circuit controls auditory cue-directed visual spatial attention. Neuron 110:109-119.e3.

Ito S, Feldheim DA (2018) The Mouse Superior Colliculus: An Emerging Model for Studying Circuit Formation and Function. Frontiers in Neural Circuits 12 Available at: https://www.frontiersin.org/articles/10.3389/fncir.2018.00010 [Accessed July 5, 2023].

Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron 89:221–234.

Kadosh O, Bonneh YS (2022) Involuntary oculomotor inhibition markers of saliency and deviance in response to auditory sequences. Journal of Vision 22:8.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. Neuron 22:751–761.

King AJ (2004) The superior colliculus. Current Biology 14:R335–R338.

King AJ, Schnupp JW, Carlile S, Smith AL, Thompson ID (1996) The development of topographically-aligned maps of visual and auditory space in the superior colliculus. Prog Brain Res 112:335–350.

Krishnan L, Elhilali M, Shamma S (2014) Segregating complex sound sources through temporal coherence. PLoS Comput Biol 10:e1003985.

Laeng B, Ørbo M, Holmlund T, Miozzo M (2011) Pupillary Stroop effects. Cogn Process 12:13–21.

Lange EB, Zweck F, Sinn P (2017) Microsaccade-rate indicates absorption by music listening. Consciousness and Cognition 55:59–78.

Lao-Rodríguez AB, Przewrocki K, Pérez-González D, Alishbayli A, Yilmaz E, Malmierca MS, Englitz B (2023) Neuronal responses to omitted tones in the auditory brain: A neuronal correlate for predictive coding. Sci Adv 9:eabq8657.

Lee A, Rajaram S, Xia J, Bharadwaj H, Larson E, Hämäläinen M, Shinn-Cunningham B (2013) Auditory Selective Attention Reveals Preparatory Activity in Different Cortical Regions for Selection Based on Source Location and Source Pitch. Frontiers in Neuroscience 6 Available at: https://www.frontiersin.org/articles/10.3389/fnins.2012.00190 [Accessed June 29, 2023].

Lehmann A, Arias DJ, Schönwiesner M (2016) Tracing the neural basis of auditory entrainment. Neuroscience 337:306–314.

Martinez-Conde S, Macknik SL, Hubel DH (2004) The role of fixational eye movements in visual perception. Nat Rev Neurosci 5:229–240.

Martinez-Conde S, Macknik SL, Troncoso XG, Dyar TA (2006) Microsaccades counteract visual fading during fixation. Neuron 49:297–305.

Martinez-Conde S, Otero-Millan J, Macknik SL (2013) The impact of microsaccades on vision: towards a unified theory of saccadic function. Nature Reviews Neuroscience 14:83–96.

Matsumoto M, Inoue K, Takada M (2018) Causal Role of Neural Signals Transmitted From the Frontal Eye Field to the Superior Colliculus in Saccade Generation. Frontiers in Neural Circuits 12 Available at: https://www.frontiersin.org/articles/10.3389/fncir.2018.00069 [Accessed June 29, 2023].

McCamy MB, Otero-Millan J, Di Stasi LL, Macknik SL, Martinez-Conde S (2014) Highly Informative Natural Scene Regions Increase Microsaccade Production during Visual Scanning. J Neurosci 34:2956–2966.

Megemont M, McBurney-Lin J, Yang H (2022) Pupil diameter is not an accurate real-time readout of locus coeruleus activity Reimer J, Gold JI, eds. eLife 11:e70510.

Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci 7:3215–3229.

Meredith MA, Stein BE (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology 56:640–662.

Milne A, Zhao S, Tampakaki C, Bury G, Chait M (2021) Sustained pupil responses are modulated by predictability of auditory sequences. J Neurosci Available at: https://www.jneurosci.org/content/early/2021/06/01/JNEUROSCI.2879-20.2021 [Accessed June 10, 2021].

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046.

Otero-Millan J, Macknik SL, Serra A, Leigh RJ, Martinez-Conde S (2011) Triggering mechanisms in microsaccade and saccade generation: a novel proposal. Annals of the New York Academy of Sciences 1233:107–116.

Otero-Millan J, Troncoso XG, Macknik SL, Serrano-Pedraza I, Martinez-Conde S (2008) Saccades and microsaccades during visual fixation, exploration, and search: Foundations for a common saccadic generator. Journal of Vision 8:21–21.

Pastukhov A, Braun J (2010) Rare but precious: Microsaccades are highly informative about attentional allocation. Vision Research 50:1173–1184.

Peel TR, Hafed ZM, Dash S, Lomber SG, Corneil BD (2016) A Causal Role for the Cortical Frontal Eye Fields in Microsaccade Deployment. PLOS Biology 14:e1002531.

Roberts MJ, Lange G, Van Der Veen T, Lowet E, De Weerd P (2019) The Attentional Blink is Related to the Microsaccade Rate Signature. Cerebral Cortex 29:5190–5203.

Rolfs M (2009) Microsaccades: Small steps on a long way. Vision Research 49:2415–2441.

Rolfs M, Kliegl R, Engbert R (2008) Toward a model of microsaccade generation: the case of microsaccadic inhibition. J Vis 8:5.1-23.

Sarter M, Givens B, Bruno JP (2001) The cognitive neuroscience of sustained attention: where top-down meets bottom-up. Brain Research Reviews 35:146–160.

Shamma SA, Elhilali M, Micheyl C (2011) Temporal coherence and attention in auditory scene analysis. Trends Neurosci 34:114–123.

Siegenthaler E, Costela FM, McCamy MB, Di Stasi LL, Otero-Millan J, Sonderegger A, Groner R, Macknik S, Martinez-Conde S (2014) Task difficulty in mental arithmetic affects microsaccadic rates and magnitudes. European Journal of Neuroscience 39:287–294.

Valsecchi M, Betta E, Turatto M (2007) Visual oddballs induce prolonged microsaccadic inhibition. Exp Brain Res 177:196–208.

Valsecchi M, Turatto M (2009) Microsaccadic responses in a bimodal oddball task. Psychological Research 73:23–33.

Veale R, Hafed ZM, Yoshida M (2017) How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. Phil Trans R Soc B 372:20160113.

Wang C-A, Blohm G, Huang J, Boehnke SE, Munoz DP (2017) Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. Biological Psychology 129:36–44.

Wang C-A, Munoz DP (2021) Differentiating global luminance, arousal and cognitive signals on pupil size and microsaccades. European Journal of Neuroscience 54:7560–7574.

White AL, Rolfs M (2016) Oculomotor inhibition covaries with conscious detection. Journal of Neurophysiology 116:1507–1521.

Widmann A, Engbert R, Schröger E (2014) Microsaccadic Responses Indicate Fast Categorization of Sounds: A Novel Approach to Study Auditory Cognition. J Neurosci 34:11152–11158.

Xue L, Huang D, Wang T, Hu Q, Chai X, Li L, Chen Y (2017) Dynamic modulation of the perceptual load on microsaccades during a selective spatial attention task. Sci Rep 7:16496.

Yablonski M, Polat U, Bonneh YS, Ben-Shachar M (2017) Microsaccades are sensitive to word structure: A novel approach to study language processing. Scientific Reports 7:3999.

Zhao S, Chait M, Dick F, Dayan P, Furukawa S, Liao H-I (2019a) Pupil-linked phasic arousal evoked by violation but not emergence of regularity within rapid sound sequences. Nat Commun 10:1–16.

Zhao S, Yum NW, Benjamin L, Benhamou E, Yoneya M, Furukawa S, Dick F, Slaney M, Chait M (2019b) Rapid Ocular Responses Are Modulated by Bottom-up-Driven Auditory Salience. J Neurosci 39:7703–7714.