Eye Movement-Related Eardrum Oscillations (EMREOs) Do Not Have a Direct Impact on Auditory Spatial Discrimination

Authors: Nancy Sotero Silva[1], Felix Bröhl[2, 3, 4], Christoph Kayser [1],

Affiliations: [1] Cognitive Neuroscience, Faculty of Biology, Bielefeld University, Bielefeld, Germany. [2] Max Planck Institute for Human Development, Berlin, Germany. [3] Max Planck Dahlem Campus of Cognition, Berlin, Germany. [4] Chair of Biopsychology, Faculty of Biology, TU Dresden, Dresden, Germany.

Correspondence: Nancy Sotero Silva (n.sotero@uni-bielefeld.de) - Department for Cognitive Neuroscience, Faculty of Biology, Bielefeld University - Universitätsstraße 25, 33615, Bielefeld, Germany. +49 521 106-5704

Abstract

Eye-movement-related eardrum oscillations (EMREOs) are acoustic signals recorded in the ear that are thought to reflect displacements of the tympanic membrane induced by saccadic eye movements. Speculations hold that their underlying mechanisms play a role in aligning visual and acoustic signals. Yet whether and how the eardrum moves during an EMREO remains unclear. Starting from the assumption that the EMREO signal directly reflects eardrum movements, we probed human sound lateralization performance for sounds presented at different times during a saccade, hence the EMREO. Since the EMREO generation involves the two middle ear muscles, whose tension can alter sound transmission, we expected spatial sound discrimination performance to vary with the state of the ERMEO at the time of sound presentation. We contrasted perceptual sensitivity, bias, and reaction times for targets presented during either positive/negative or during small/large EMREO deflections in two tasks, one relying on free-field sounds and one presenting sounds in-ear. However, no analysis revealed any significant effect. Still, and in line with previous studies, we found that sound localization performance was dependent on the congruency of saccade and acoustic target directions. We conclude that either the eardrum does not move as directly reflected by the EMREO signal, or it does so, but the underlying changes at the tympanic membrane only have minimal perceptual impact. These results call for more refined studies to understand how the eardrum moves during a saccade and whether or how the EMREO impacts auditory or multisensory perception.

Keywords: tympanic membrane, auditory localization, eye movements, saccades, spatial discrimination

Introduction

Many objects in real life are multisensory, and our brain needs to combine the position information obtained from our different sensory modalities. This, for example, requires the alignment of eye-centered visual signals with spatially resolved sounds. Such acoustic spatial information can be obtained from interaural time (ITD) and level differences (ILD), which are particularly informative about azimuthal sound positions (Azevedo et al., 2024; Ito et al., 2020; Schut, Van der Stoep and Van der Stigchel, 2018; Tollin, 2003; Cohen and Andersen, 2000; Wightman and Kistler, 1992; Stricanne, Andersen, and Mazzoni, 1996). Recent work debates whether the multisensory alignment of visual and acoustic signals even engages the periphery of the auditory pathways (Sotero Silva, Kayser and Bröhl, 2025; King et al, 2023; Lovich et al., 2023a; Gruters et al., 2018). Studies have described eye movement-related eardrum oscillations (EMREOs), low-frequency signals recorded in the ear, which might result from deflections of the tympanic membrane precisely aligned to saccadic onsets. These EMREOs scale parametrically with the direction and amplitude of saccadic eye movements (Murphy et al., 2020; Lovich et al., 2023a). While the precise origin of EMREOs and their action on the tympanic membrane remain unclear, their generation likely involves the two middle ear muscles (MEMs) i.e., the tensor tympani and the stapedius muscles (Sotero Silva, Kayser and Bröhl, 2025; Murphy et al., 2020; Gruters et al., 2018). The tension of these can shape how fine acoustic cues, such as timing differences, are transduced (Cho, Ravicz and Puria, 2023; Lovich et al., 2023, Tasko et al., 2022). For example, Cho, Ravicz and Puria (2023) showed that the contraction of the MEMs changes the shape of the tympanic membrane and alters low-frequency middle-ear transmission magnitudes and timing delays. Hence, it is possible that sounds received during an EMREO could be encoded differently than sounds received in the absence of an EMREO and therefore result in a different perceptual estimate of a sound's location. We here test this possibility by probing participants' performance in two auditory spatial discrimination tasks in which the relevant sounds were presented at different phases of the EMREO signal.

Several studies have described the properties of EMREOs in a detailed manner. The EMREO signal recorded using in-ear microphones lasts for about 50 to 100 ms and comprises multiple deflections whose strength and sign relate to the direction and amplitude of a saccade (Lovich et al., 2023a, 2023b; King et al., 2023; Bröhl and Kayser, 2023). For example, the EMREO elicited by saccades to positions ipsi and contralateral to the ear under study exhibits deflections of opposing sign. However, it remains unclear how the tympanic membrane precisely moves during an EMREO, and whether the actual movement of the tympanic membrane indeed lasts as long and follows the time course of the recorded signal. Starting from the assumption that the EMREO is indeed indicative of the actual movement of the tympanic membrane, one would assume that sounds presented at different times relative to a saccade onset are received with the tympanic membrane either deflected in- or outwards and with different amplitudes. This should alter the transduction of a sound and affect how it is encoded in the auditory system (Cho, Ravicz and Puria, 2023). We have tested this in

a previous study using a near-threshold sound detection task (Bröhl and Kayser, 2023). However, we did not find evidence for an influence of the relation of sound presentation to the EMREO. This negative result may simply reflect the nature of the task, which did not require participants to exploit the spatial information carried by individual sounds but asked them simply to detect that a sound was heard. For the present study we implemented two tasks relying on spatial sound cues, one requiring participant to compare the relative horizontal position of two free-field sounds, and one requiring the comparison of two sounds delivered using in-ear phones at varying ITDs.

Because EMREOs are triggered by saccadic eye movements, any investigation of how EMREOs influence auditory-spatial perception must also address the relationship between the perceptual performance and saccades themselves. Several studies have reported that saccades do not impair basic auditory abilities such as sound detection or pitch discrimination; participants were reported to perform equally well for stimuli presented during a saccade and for stimuli presented in the absence of eye movements (Binda and Morrone, 2018; Harris and Lieberman, 1996). However, studies using spatial tasks have demonstrated that sounds presented around the time of a saccade are often mislocalized (Krüger et al., 2016; Collins, Heed and Röder, 2010; Pavani, Husain and Driver, 2008). Impending saccades can shift auditory attention, yielding faster reaction times when the auditory target direction matches the saccade goal (Rorden and Driver, 1999). Simultaneous eye movements and auditory-localization tasks tend to produce a bias of perceived sound location toward the direction of gaze (Lewald and Ehrenstein, 1996; Lewald, 1997; Getzmann, 2002; Razavi, O'Neill and Paige, 2007) and the saccade dynamics can modulate this bias so that the perceived auditory position is displaced opposite to the saccade direction at saccade onset and toward the saccade direction near saccade offset (Klingenhoefer and Bremmer, 2009). We hence accounted for such congruency effects between the direction of a saccade and the lateralization of a sound in our analysis.

In the present study the same participants performed two auditory spatial discrimination tasks in which they compared the horizontal position of two sounds presented around horizontal saccades. Each trial required a comparison of the perceived horizontal position of two subsequent sounds. The first sound was always presented centrally and served as a reference, while the second sound was positioned to the left or right. Importantly, we jittered the presentation times of the second sounds relative to the saccade, so that the presentation of this randomly fell into different phases of the concurrently recorded EMREOs. To test whether task performance i.e., perceptual sensitivity, bias or reaction times, was affected by the timing of the sound relative to the EMREO, we performed two analyses: First, we contrasted trials in which the targets were presented during EMREO deflections of different signs, and second, during deflections of different magnitude. Overall, our results reveal no influence of the target-EMREO relation on participants' performance in either task or analysis.

Materials and Methods

Participants

The data presented here were collected as part of a two-experiment study. The first part of that study was designed to characterize EMREOs during visually guided and acoustically guided saccades and has been published (Sotero Silva, Kayser and Bröhl, 2025). The second part is reported here. 34 participants were recruited among the university and community members (24 females, 10 males, mean ± SD age = 27,3 ± 5,3 years) without self-reported hearing or vision impairments and neurological or psychiatric disorders. Each participant provided written informed consent and received 15 euros per hour as compensation. The study was approved by the ethics committee of Bielefeld University (#2021-218). Initially, participants were visually screened for obstructions in their ear canals and performed pure tone audiometry to establish air conduction hearing thresholds. They underwent tympanometry and were tested for acoustic reflexes at 500, 1000, 2000 and 4000 Hz. Individuals with an audiometry quadritone average threshold below 20 dB were considered to have normal hearing (World Health Organization, 2020) and were included in the actual experiment. Data was obtained in two sessions taking place on two different days. We ensured that the functioning of the ear was comparable between sessions based on the hearing screening and middle ear assessments. For each participant, respective sessions tested two experiments described in the following.

Experimental Setup

The experiments were conducted in the same set-up as described in Sotero Silva et al. (2025). Recordings took place in a soundproofed and electrically shielded booth with no light sources other than a projector. Participants remained seated in front of a table with their head on a chinrest, aligned to the center of a sound-permeable screen 100 cm away (Screen International Modigliani, 2 × 1 m). In-ear microphone recordings were collected using two Etymotic ER-10C (Etymotic Research) systems, via probes that served both for sound presentation and microphone recording (we name the phones/microphone assembly as “probes” when we refer to the device in general; as “in-ear phones” to indicate the stimuli presentation; and as “microphone” to indicate the recordings). Microphone signals were pre-amplified using the ER-10C DPOAE amplifiers (Etymotic Research), with gain set to +40 dB. Then, the signal was digitized through an ActiveTwo AD-box (BioSemi) at a sampling rate of 2048 Hz. All visual stimuli were projected (Acer Predator Z650, Acer; 60 Hz refresh rate) onto the screen. The sound stimuli were created using Sound Blaster Soundcards with sampling rate 44100 Hz and were presented differently depending on the experiment. In one experiment they came from an array of 5 free-field speakers (Monacor MKS-26/SW, MONACOR) located behind the screen and positioned between -22 and 22 degrees (spaced in 11 degrees steps) azimuth and 8 cm from the screen border (then around 108 cm away from the subject). In the other experiment they were amplified with a HeadAMP 4 (ARTcessories) amplifier and presented via the ER-10C in-ear phones. The Psychophysics toolbox (Brainard, 1997) for MATLAB (The MathWorks, 2022) was

used to control the stimulus presentation, which was synchronized to the recording system using TTL pulses. Figure 1A represents the set-up montage.

Eye movements were recorded at 1000 Hz from the left eye using a desktop mounted infrared-based EyeLink 1000 Plus Eye Tracker (SR Research Ltd., 2015) running Host Software vs 5.15. Eye-tracking calibration and validation using the built-in 5-point grid was performed for each block. The parameters for saccade detection in the EyeLink system ("cognitive" setting) were a velocity threshold of 30°/s and an acceleration threshold of 8000°/s. Eye positions were recorded as horizontal and vertical coordinates in degrees of visual angle.

Control measurements were performed at the beginning of each session to ensure the quality of the microphone and in-ear recordings. First a recording was made in the empty booth. Microphone responses for 15 seconds of silence followed by 20 tone stimuli presented from the speakers (10 stimuli at 100 Hz, 10 stimuli at 200 Hz, duration 400 ms) were measured. Secondly, after the insertion of the probes in the participants' ears, the same measurement was repeated. Adjustments for the microphone insertion and signal corrections across ears for data analysis were performed based on these measurements as described in Sotero Silva et al. (2025).

Experimental Paradigms

During each session we first obtained data during visually and auditory guided saccades as described in our previous study (Sotero Silva, et al., 2025). Subsequently, participants were instructed about the respective acoustic tasks, performed training trials, completed a threshold estimation procedure, and then performed four blocks of the actual tasks.

Each trial involved the presentation of two tones (1200 Hz 15 ms sine burst, cosine ramped and presented at 65 dB SPL; inter-sound interval of 240 ms +/- 20 ms jitter) presented either from two different free-field speakers ("free-field task") or using varying ITDs presented via the in-ear phones ("in-ear task"; see Fig. 1B,C). For the free-field task, sounds were presented over the array of five speakers. A continuous sound localization was achieved by linearly interpolating sound amplitudes between neighboring speakers. For the in-ear task, sounds were delivered via the ER-10C systems and were lateralized by adjusting their relative ITDs to achieve the required perceptual auditory position (sound locations were converted from azimuth angles into ITD values based on a generic head-radius function of 8.76 cm). The first tone was always presented in the center position, the second either to the left or right. Participants' task was to indicate whether the second tone was to the right or left relative to the first tone by pressing one of two response buttons. During the familiarization phase, participants were initially presented with the tones together with an indication of the correct response on the screen. Subsequently they had to perform the task without feedback. Once they reached 70% correct responses, they performed a block to determine their Minimum Audible Angle (MAA). This was measured using three interleaved 3-down 1-up staircases from which we averaged the individual MAA's to obtain the participants threshold. For the actual task we then presented tones with

a position difference near threshold, adding a +/-5% jitter on the MAA for each trial. Participants completed 4 blocks of 150 trials for each task.

During the actual task participants had to perform both visually-guided saccades and the acoustic discrimination task. Visually guided saccades were cued using a white dot (0.2-degree radius). The dot first appeared in the center of the screen for 400 ms and participants had to fixate this. Subsequently, the fixation dot jumped to a random horizontal position to the right or the left (target positions were between -9 and -3 degrees to the left, and +3 and +9 degrees azimuth to the right, in one-degree steps) and participants were instructed to follow this dot as quickly as possible. The dot remained present for 500 ms. Note that participants also performed similar visually guided saccades in a separate block designed to just analyze EMREO properties (Sotero, Kayser and Bröhl, 2025) and were therefore familiar with this task. Inter-trial intervals were uniformly sampled from the interval between 900 to 1400 ms.

The timing of the two sounds was adjusted to the saccade behavior of each participant with the aim to sample different positions during the saccade-induced EMREO with the second tone. To achieve this, we estimated the average saccade latency and its variability for each participant based on the separately obtained visual saccade task. Based on this we derived an estimate for when to present the sequence of tones, which also included estimates of the delays between auditory and visual presentation times in the setup. For most participants the average saccade latency was on the order of 250 ms and hence the first tone was usually presented prior to the saccade and the second tone during the saccade. However, given the trial-to-trial variability of saccade latencies, this resulted in some target sounds to be presented prior to the EMREO, some during, and some after the EMREO. Example data for one participant are shown in Figure 2A. Each trial hence comprised a horizontal visually guided saccade starting from a central fixation dot, and two tones, the second of which was designed to be perceptually to the left or right of the central fixation dot. The direction of the instructed saccade and the direction of the relative sound positions, as well as their respective amplitudes, were pseudo-randomized. As a result, the relative acoustic position could be in the same direction (termed congruent) or the opposite direction (termed incongruent) as the visual saccade (shown in Figure 1B).

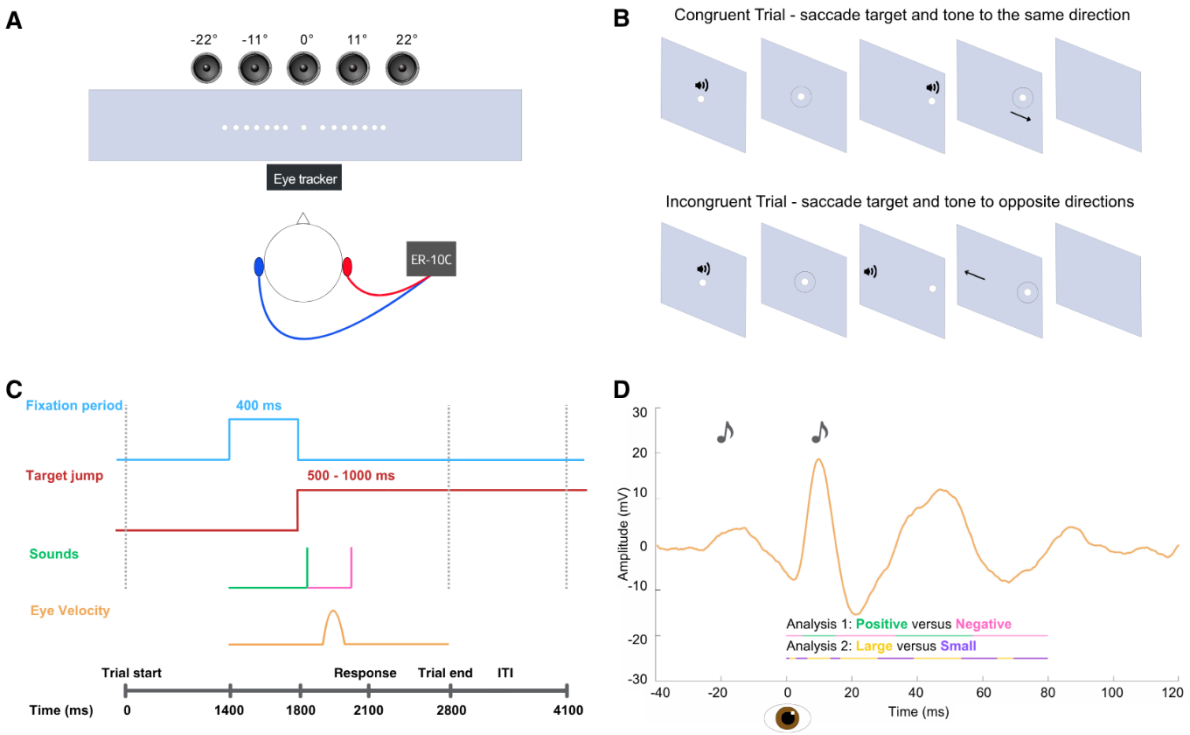

Figure 1. Experimental Paradigms and Logic. (**A**) shows a representation of the experimental setup, indicating the position of the participant, the screen for the projection of the visual targets, the eye tracker and the speakers positioned behind the screen. In-ear microphones for the left (blue) and right (red) connected to the ER-10C amplifier are also displayed. (**B**) shows the sequence and direction of visual and auditory targets for trials in which the direction of the saccade and the lateralization of the acoustic target are congruent and incongruent. Arrows represent the expected response for the sound lateralization task. (**C**) represents the timeline for one trial. Fixation period lasted 400 ms (light blue) followed by target jump lasting 600 ms (red), period for the saccade detection (orange). Tones were presented around the fixation (first tone, in green) and after jump (second tone, in pink). Time events are represented in the black thick line. (**D**) represents possible incidences of the tones relative to a schematic EMREO (orange line). Based on this incidence we devised two different analyses, one in which we contrasted behavioral performance for tones presented during positive or negative EMREO deflection (green and pink) and one contrasting tones presented by a deflection of large or small magnitude (yellow and purple). In this example, the first tone is presented during fixation (20ms before the target jump), while the second tone falls during a large positive deflection (around 15 ms after jump). Valid saccades would be detected in the 1000 ms interval after the start of the target jump, represented by the red line.

Data Processing and Analysis

For each participant we computed the trial-averaged EMREOs during the acoustic task as reported previously (Sotero Silva, Kayser and Bröhl, 2025; Bröhl and Kayser, 2023). We aligned the microphone recordings to the onset of the largest saccade following the jump of the fixation point. To avoid artifacts, we only included saccades whose amplitude was within the range 2 to 16 degrees visual angle, for which the initial fixation prior to the saccade was stable and for which the entire horizontal and vertical traces of the eye position were contained in an 18 x 18 degrees window around the central fixation point. The EMREO epochs for the included saccades were demeaned and screened for extreme values, and epochs for which the maximal Z-score of the EMREO signal at any point in time exceeded 4.5 SDs of the overall signal were removed. In addition, we visually inspected the EMREO time courses. Some recordings presented

an easily recognizable periodic oscillation, while others reflected the contamination with external noise and some expressed excessively many trials with values exceeding the Z-score threshold. Of the 34 participants, we retained 23 for further analysis based on the quality of the EMREO signal. The resulting trial-averaged EMREOs are shown for the group-data in Figure 2C and 2D and illustrate the opposing deflections for saccades ipsi- and contralateral to the recorded ear.

Based on the participant-wise trial-averaged EMREOs we analyzed the performance in the acoustic discrimination task in two ways: one probing the relevance of the direction of the EMREO deflection and one probing the magnitude of the EMREO deflection. Each analysis was performed separately for trials in which the direction of the acoustic target was congruent or incongruent with the direction of the saccade target, given that previous studies have suggested a potential influence of saccade-target congruency on auditory spatial performance (Klingenhoefer and Bremmer, 2009; Getzmann, 2002; Rorden and Driver, 1999; Lewald and Ehrenstein, 1996; Lewald, 1997). For each analysis we split trials based on the state of the trial-averaged EMREO at the trial-wise latency of the second tone relative to saccade onset (i.e. time zero in the EMREO time axis). Since for each participant we obtained one EMREO for each ear per ipsi- and contralateral saccade direction, we combined these signals based on the following logic. For the analysis probing the relevance of the direction of the EMREO deflection we focused on the EMREO in the ear contralateral to the saccade direction. The EMREO in the contralateral ear is generally stronger than that in the ipsilateral ear (as shown in Figure 2C and 2D; also shown in Sotero Silva, Kayser and Bröhl, 2025; King et al., 2023; Lovich et al., 2023a; Gruters et al., 2018), but both vary in an anti-phase manner so that a positive deflection in one ear corresponds to a negative deflection in the other. We also repeated this analysis based on the EMREO in the ipsilateral ear but obtained essentially the same results. For the analysis probing the magnitude of the EMREO deflection, we inverted the sign of the contralateral EMREO and averaged both. We then split trials for each analysis as follows (Fig. 1D): for the direction of the EMREO deflection we split trials depending on the sign (positive, negative) of the relevant EMREO signal. For the magnitude of the EMREO deflection we split trials based on the median of the absolute magnitude of the relevant EMREO deflection (large, small).

For each analysis (free-field or in-ear tasks, trials with saccade and acoustic target congruent or incongruent, and either focusing on EMREO direction or magnitude) this resulted in two groups of trials. We contrasted reaction times, perceptual sensitivity and perceptual bias between these, with sensitivity and bias computed based on signal detection theory (Macmillan and Creelman, 2005).

To ensure meaningful definition of the EMREO we focused this analysis only on trials in which the second tone was presented between 0 ms and 80 ms from saccade onset, which is the period containing the prominent EMREO signal. Given the pseudorandom presentation time of the tones relative to the saccadic eye movements (Figure 2A), the resulting number of trials differed between trial groups and between participants. For the final analysis we only included participants for whom we could retain more than 10

trials in any group and for any of the analyses. The final dataset comprised data from 14 participants with 179 +/- 22 trials in the free-field task and 21 participants 213 +/- 22 trials in the in-ear task. Figure 2B shows the distribution of target times relative to saccade onset across trials and participants for each of the tasks.

Statistical Testing

The main hypothesis concerned a potential difference in behavioral performance (reaction times, sensitivity or bias) between trial groups. We implemented these analyses using paired t-tests for which we report p-values, Cohen's D effect size and the associated Bayes factors (BF) obtained using the JZS prior (Rouder et al., 2009). Given the focus on two auditory tasks, two ways of splitting trials per EMREO (direction and magnitude) and treating trials with congruent (incongruent) saccade and auditory target directions separately, this amounted to a larger number of tests. As this may raise concerns about the interpretation of p-values, we focus on the associated effect sizes and BF for interpretation. Since our main conclusion is that no analysis yielded considerable effects (as judged based on either Cohen's D or the BF) we do not see the need to correct for multiple statistical comparisons.

Results

Spatial Discrimination Performance Depends on the Congruency of Saccade Direction and Target Lateralization

We first asked whether participants' performance in discriminating the lateralization of the target sounds depends on the congruency of the saccade direction and the horizontal shift of the target sound relative to the central reference. Previous studies suggest that performance should be better when these two directions are congruent, and our data confirms this. Participants sensitivity was significantly higher for congruent saccade-target directions both in the free-field task (congruent d' 0.88 +/- 0.15 vs incongruent d' -0.31 +/- 0.18; $p < 10\text{-}5$, Cohen 1.34, BF = 135) and the in-ear task (congruent d' 0.89 +/- 0.16 vs. incongruent d' -0.19 +/- 0.23; $p = 0.003$, Cohen 0.75, BF = 15). Similarly, participants' behavior exhibited different biases in both conditions, both for the free-field task (congruent -0.38 +/- 0.13 vs. incongruent 0.12 +/- 0.11; $p = 0.002$, Cohen 1.05, BF = 24) and the in-ear task (congruent -0.47 +/- 0.10 vs. incongruent 0.13 +/- 0.12; $p = 0.001$, Cohen 0.88, BF = 50) reflecting participants tendency to partly align their responses with the direction of the saccade. Reaction times were nearly identical in both conditions in the free-field task (congruent 0.92 +/- 0.03 s vs. incongruent 0.93 +/- 0.02 s: $p = 0.641$, Cohen 0.13, BF = 0.30) and the in-ear task (congruent 0.97 +/- 0.03 s vs. incongruent 0.96 +/- 0.02; $p = 0.819$, Cohen 0.05, BF = 0.23).

Group-level EMREOs

As reported previously for the same participants, the EMREO signals followed the expected pattern: an initial oscillation appearing around the start of the eye movement, signals with opposite phases for saccades towards ipsi- and contralateral targets, and robust peaks between 10 and 40 ms after saccade onset. Figure 2A shows example data from one individual obtained during the free-field task illustrating that participants EMREO together with the presentation times of the two sounds during individual trials. Figure 2C and 2D shows the group level EMREO time courses for both tasks. The latter is shown for both ears and saccades towards visual targets ipsi- and contralateral to each ear separately.

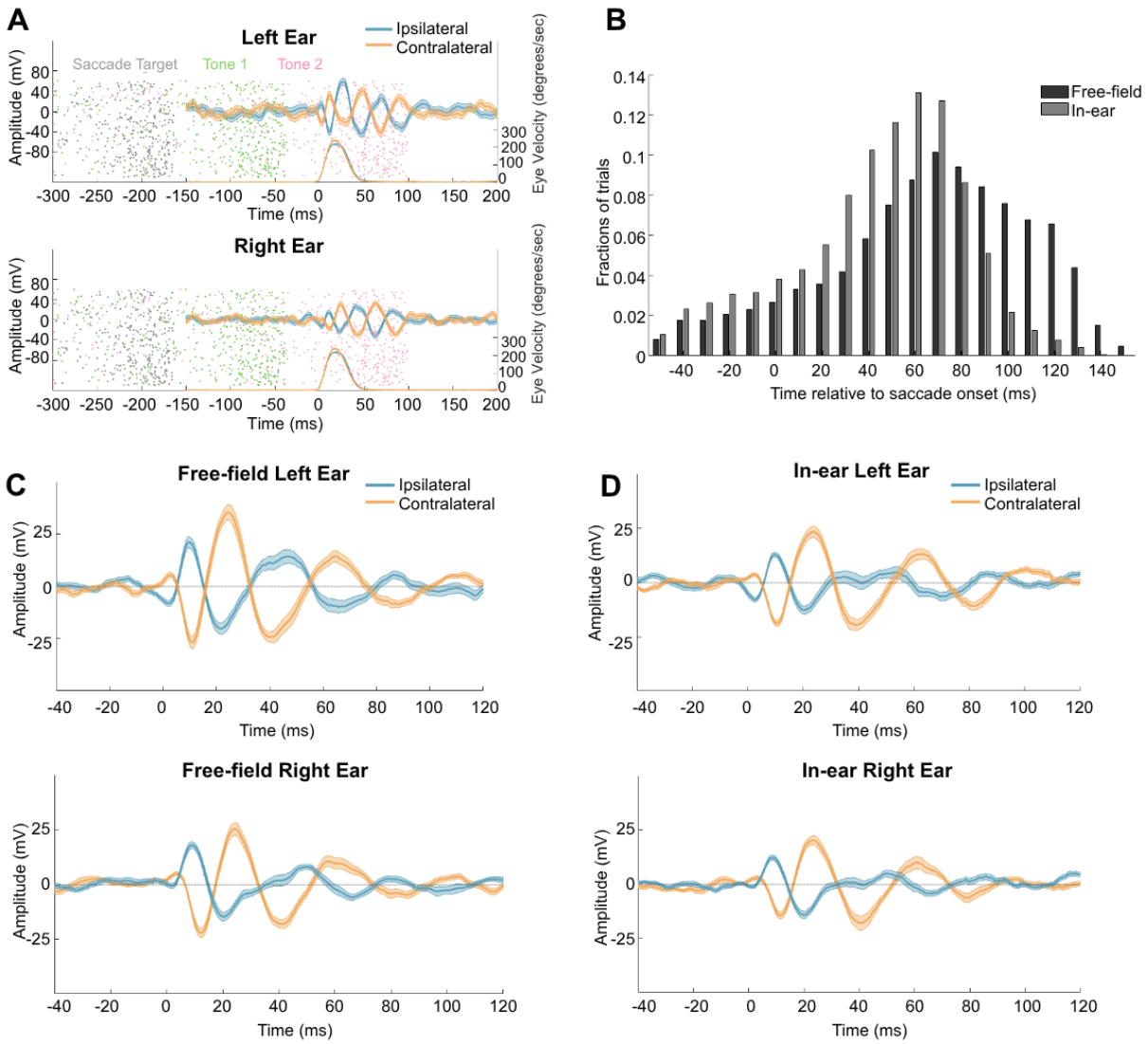

Figure 2. Example and Group-Level EMREOs. (**A**) shows example data for one participant, illustrating the EMREO together with eye velocity during the free-field task. Dots display the presentation times for the Saccade Target (grey), the first tone (green) and the second tone (pink) relative to the EMREO signal in individual trials. (**B**) shows the variability of the presentation of tones per trial within the free-field (dark grey) and in-ear (light grey) tasks across all trials of all participants. (**C**) and (**D**) show group-average EMREOs for the free-field (n=23, 202.6 +/- 31.8 trials) and in-ear (n=23, 197.9 +/- 30.3 trials) tasks. EMREO time courses are shown for the left and right ears, split for ipsi- (blue) and contralateral (orange) saccades. Thicker lines represent the mean EMREO, and shading reflect the SEM.

Spatial Discrimination Performance Is Not Affected by the Sound-EMREO Relation

We performed two analyses to test whether the state of the ear at the time of the target tone, as reflected by the EMREO signal, is relevant for participants' performance. In one analysis we split trials by whether the target sound was presented during a positive versus a negative deflection of the contralateral EMREO; in the second analysis we split trials by whether the target sound was presented during a large or small deflection of the EMREOs. Each analysis focused on three aspects of the behavioral data (participants sensitivity, bias and reaction times) and was applied separately to trials in which the direction of the saccade and the lateralization of the target sound were congruent or incongruent. Each analysis was applied to the data from the free-field task (n = 14) and the in-ear task (n = 21).

The respective results are shown in Figure 3 and Tables 1 and 2. Almost all effect sizes were small, and all BF reflect indecisive or anecdotal evidence against a condition difference. Only the sensitivity for trials split by EMREO magnitude during the in-ear task featured a statistically significant result ($p < 0.05$), but the BF for this effect was not supportive of a clear effect. That is, even across multiple analyses probing a possible relation of EMREO direction or magnitude on behavior we were unable to support such an effect.

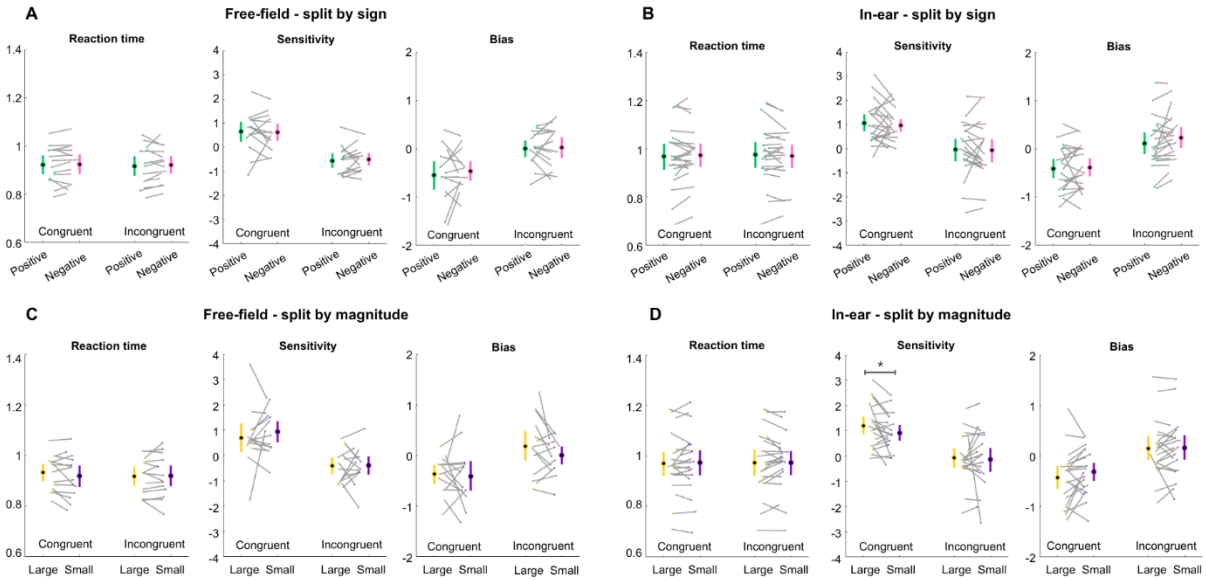

Figure 3. Sound Lateralization Performance Is Independent of the EMREO. Reaction times, sensitivity and bias for congruence and incongruent trials. (**A**) and (**C**) show the data for the free-field task (n=14), Panels (**B**) and (**D**) for the in-ear task (n=21). Data split is shown in the rows: by sign (positive deflection, green; negative deflections, pink) or magnitude (large deflections, yellow, small deflections, purple). Means and SD are shown in the thicker vertical bars, thin lines the individual data. * represents $p < 0.05$ (paired t-test).

Table 1. Statistical results for the free-field task. Values are p-values obtained from paired t-tests, Cohen's D effect size and Bayes factors (BF) (n=14).

		Congruent Trials			Incongruent Trials		
		p	Cohen	BF	p	Cohen	BF
Data split by deflection for contralateral EMREO	Reaction time	0.773	0.08	0.28	0.480	0.19	0.34
	Sensitivity	0.848	0.05	0.27	0.568	0.16	0.31
	Bias	0.486	0.19	0.34	0.747	0.09	0.28
Data split by magnitude	Reaction time	0.150	0.41	0.70	0.780	0.08	0.28
	Sensitivity	0.387	0.24	0.38	0.931	0.02	0.27
	Bias	0.730	0.09	0.29	0.107	0.46	0.89

Table 2. Statistical results for the in-ear task. Values are p-values obtained from paired t-tests, Cohen's D effect size and Bayes factors (BF) (n=21). * represents p < 0.05 (paired t-test).

		Congruent Trials			Incongruent Trials		
		p	Cohen	BF	p	Cohen	BF
Data split by deflection for contralateral EMREO	Reaction time	0.480	0.16	0.29	0.397	0.19	0.32
	Sensitivity	0.506	0.15	0.28	0.808	0.05	0.23
	Bias	0.702	0.08	0.24	0.093	0.38	0.85
Data split by magnitude	Reaction time	0.630	0.11	0.25	0.965	0.01	0.23
	Sensitivity	**0.032 ***	**0.50**	**1.94**	0.761	0.07	0.24
	Bias	0.123	0.35	0.69	0.840	0.04	0.23

Discussion

Eye movement-related eardrum oscillations (EMREOs) describe pressure changes in the outer ear following saccadic eye movements. Despite several recent studies, their origin and function remain unclear. The generation of the EMREOs has been postulated to involve the two MEMs (the tensor tympani and stapedius) and possibly the outer hair cells. The MEMs implement critical functions in protecting hearing by increasing the stiffness of the ossicular chain. Small changes in the stiffness have been associated with potentially shifting the phase of the transmitted sound. Such a phase shift could distort the encoding of temporal cues that are important for spatial hearing (Cho, Ravicz and Puria, 2023; Racca et al., 2022; Lawrence, 1965). Hence, one could expect an influence of the EMREO on the encoding of acoustic spatial signals. Addressing this, we probed whether the relative presentation of sounds in different phases of the EMREO signal affects how well their position can be judged. We tested this in two tasks performed near perceptual threshold and consistently across both the evidence does not point in favor of such an effect.

We take this to suggest that either the interpretation of the EMREO signal as directly reflecting movements of the tympanic membrane is not true; or alternatively, that the effective changes at the eardrum are too short lived or have only small perceptual effects. The assumption that the temporally extended counter-phase pattern of the EMREO reflects movements of the tympanic membrane on the same time scale is central to the logic of the present experiments. Under this assumption, and the assumption that the EMREO involve the MEMs introducing interaural timing differences, which correspond to sound localization differences on the scale of saccade amplitudes utilized in this study (Cho, Ravicz and Puria, 2023), it is reasonable to expect a difference in behavioral performance between sounds presented at different deflections of the EMREO. However, the lack of behavioral consequences observed here is incompatible with this hypothesis. Alternatively, the eardrum may oscillate just like the EMREO signal but the resulting changes in the membrane stiffness may not produce a sufficiently strong or lasting change to influence sound localization performance under the conditions tested here.

Understanding the Origin of the EMREO Signal

The EMREO waveform is likely associated with the coordinated contraction of the tensor tympani and stapedius, with a possible contribution from the outer hair cells (Gruters et al., 2018; Lovich et al., 2023a). However, the EMREO signal is an acoustic signal recorded using in-ear microphones and only reflects the change of pressure at the microphone membrane. Whether the tympanic membrane moves directly in proportion to the EMREO signal remains unknown. One possibility is that the tympanic membrane only moves very briefly, or changes its tension, and the resulting temporally extended signal recorded is only an acoustic reverberation resulting from this brief change. In this interpretation the EMREO could still be a peripheral signature of a central motor command involved in oculomotor function, but that command may translate into only a brief action on the tympanic membrane (Lovich et al. 2023b, Sotero Silva, Kayser and Bröhl, 2025). This does not rule out that the EMREO has a detectable impact on spatial sound encoding and perception, or that it plays a role in multisensory coordinate alignment (Gruters et al, 2018; Murphy et al., 2020; Lovich et

al., 2023a). However, this action could be limited to a short time directly around the actual action of the MEMs, e.g. directly at saccade onset, and may easily be missed for sounds not presented at the suitable time point.

A further possibility is that the eardrum moves in line with the EMREO signal, but the actual changes in membrane stiffness and sound transduction delays caused by the middle ear muscles are irrelevant for sound localization. Relative eardrum stiffness, modulated for example by the acoustic reflex, alters the perceived properties of sounds (Racca et al., 2022; Avan et al., 2000), such as attenuating low-frequency transmission and shortening the middle-ear transmission group delay (O'Connor, Cai and Puria, 2017). At the same time, the magnitude of an individual's EMREO may be related to eardrum compliance, whereby lower compliance, therefore larger stiffness, is associated with a higher EMREO peak amplitude (Sotero Silva, Kayser and Bröhl, 2025). The assumption that the action of the middle ear muscles during the EMREO has a meaningful impact on sound transduction under everyday settings still is unclear, as previous data supporting sound transduction being mediated by MEM tension were obtained ex-vivo (Cho, Ravicz and Puria, 2023). However, obtaining direct information on the actual movement of the eardrum in humans in vivo during task performance is a challenge and would require measurements that are not affected by concurrently presented acoustic stimuli (e.g. using laser doppler vibrometry).

Alternatively, the eardrum could move for a prolonged time in line with the EMREO signal, but this may impact sound localization performance only under conditions different than those tested here. For example, since the amplitude of the EMREO scales with saccade size, it could be that any relation of EMREOs to sound perception emerges only for very large EMREO amplitudes, hence large saccades. Larger saccades also induce larger spatial incongruencies between visual and acoustic signals, increasing the need to align multisensory information. However, the data published in previous studies suggests that the amplitude of the EMREO signal does not scale linearly with saccade size and rather seems to asymptote for saccades larger than about 10 to 15 degrees (Gruters et al., 2018). This casts doubts on the possibility that the EMREO induced by very large saccades (>20 degrees) has a more pronounced effect than an EMREO for saccades around 5 to 10 degrees, as tested here.

Another possibility is that the movement of the eardrum reflected in the EMREO matters only for sounds other than those tested here. The present experiments relied on a single brief tone of 1200 Hz. The tone was presented in a free-field condition and in an in-ear condition, with the latter only varying the stimulus ITDs to generate the percept of different azimuthal positions. ITDs are mostly relevant for estimating azimuthal positions for low frequency sounds, i.e., below 2000 Hz (Azevedo et al., 2024; Wightman and Kistler, 1992). While the present sounds fall in this range, it could be that the perceptual effect becomes relevant only at much lower sound frequencies and future studies may consider a wider range of sounds.

Finally, the laboratory setting used to study EMREOs do not necessarily replicate the multifaceted conditions present in natural environments, where EMREOs may possibly play a more pronounced role for hearing. For example, the studies to date rely on a fixed head position and eye movements that do not bear any relation to any acoustic

task. However, if the EMREO reflects a mechanism genuinely involved in aligning visual and acoustic signals, one may expect this mechanism to operate in a task-specific manner. Many everyday conditions involve multiple visual and acoustic stimuli, and attention and task-relevance are critical factors shaping multisensory binding under such conditions (Noppeney, 2021). To address this, future studies could employ continuous spatial tasks, such as tracking moving sounds, or tasks detecting subtle changes in a sound's properties to increase ecological validity.

Congruency of Saccade and Acoustic Target Modulates Performance

Our paradigm required participants to make rapid judgments about positions for sounds presented during a saccade that were either displaced in the direction congruent or incongruent with the saccade. Perceptual sensitivity was significantly higher for congruent trials in both tasks, consistent with previous studies on saccades and auditory perception (Krüger et al., 2016; Collins, Heed and Röder, 2010). Auditory targets presented at the same direction as the saccade were previously reported to yield shorter reaction times (Yamagishi and Furukawa, 2020) and a perceptual bias towards the saccade direction (Lewald and Ehrenstein, 1996; Lewald, 1997; Getzmann, 2002). This is exactly what we observed here.

Such a congruency effect between saccades and sensory targets may also be expected based on work on multisensory alignment. For example, Rorden and Driver (1999) found that auditory localization was most accurate in the direction of the upcoming saccade, indicating that the motor plan biases auditory processing towards future gaze directions. Similar cross-modal benefits appear for tactile-motor interactions (Juravle and Deubel, 2009) or visual-tactile tactile discrimination (Spence et al., 2004; Rorden et al., 2002) and further studies underscore that even the smallest oculomotor adjustments can influence cross-modal attentional processing (Kliegl et al., 2009). This points to a shared mechanism of cross-modal attentional and sensorimotor integration that adjusts perceptual estimates based on multisensory signals and planned or executed movements. Since their first description, EMREOs have been framed in the context of multisensory spatial perception, but whether they index a similar interaction between auditory perception and saccadic eye movements remains to be determined.

Author contributions

Nancy Sotero Silva: Performed experiments, Analyzed data, Interpreted results of experiments, Prepared figures, Drafted manuscript, Edited and revised the manuscript. **Felix Bröhl:** Conceived and designed research, Analyzed data, Interpreted results of experiments, Drafted manuscript, Edited and revised the manuscript, Approved final version of the manuscript. **Christoph Kayser:** Conceived and designed research, Analyzed data, Interpreted results of experiments, Drafted manuscript, Approved final version of the manuscript.

References

1. Azevedo ACPFO, Pellegrino TG, Peña JL, Marin B, Pavão R. Sufficiency of ITD for (biased) human auditory azimuthal perception. *bioRxiv* 2024.11.05.622135, 2024. https://doi.org/10.1101/2024.11.05.622135
2. Ito S, Si Y, Feldheim DA, Litke, AM. Spectral cues are necessary to encode azimuthal auditory space in the mouse superior colliculus. *Nature Communications* 11:1087, 2020. https://doi.org/10.1038/s41467-020-14897-7
3. Schut MJ, Van der Stoep N, Van der Stigchel S. Auditory spatial attention is encoded in a retinotopic reference frame across eye-movements. *PLOS One* 13(8): e0202414, 2018. https://doi.org/10.1371/journal.pone.0202414
4. Tollin DJ. The Lateral Superior Olive: A Functional Role in Sound Source Localization. *Neuroscientist* 9(2): 127-143, 2003. https://doi.org/10.1177/1073858403252228
5. Cohen YE, Andersen RA. Reaches to sounds encoded in an eye-centered reference frame. *Neuron* 27: 647-652, 2000. https://doi.org/10.1016/S0896-6273(00)00073-8
6. Wightman FL, Kistler DJ. The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am.* 91: 1648–1661, 1992. https://doi.org/10.1121/1.402445
7. Stricanne B, Andersen RA, Mazzoni P. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J. Neurophysiol.* 76: 2071–2076, 1996. https://doi.org/10.1152/jn.1996.76.3.2071
8. Sotero Silva N, Kayser C, Bröhl F. Unraveling eye movement-related eardrum oscillations (EMREOs): how saccade direction and tympanometric measurements relate to their amplitude and time course. *Hearing Research* 461: 109276, 2025. https://doi.org/10.1016/j.heares.2025.109276
9. King CD, Lovich SN, Murphy DLK, Landrum R, Kaylie D, Shera CA, Groh JM. Individual similarities and differences in eye-movement-related eardrum oscillations (EMREOs). *Hearing Research* 440: 108899, 2023. https://doi.org/10.1016/j.heares.2023.108899
10. Lovich SN, King CD, Murphy DLK, Abbasi H, Bruns P, Shera CA, Groh JM. Conserved features of eye movement related eardrum oscillations (EMREOs) across humans and monkeys. *Philosoph. transact. Royal Soc. London Series B, Biol. sci.,* 378: 20220340, 2023. https://doi.org/10.1098/rstb.2022.0340
11. Lovich SN, King CD, Murphy DLK, Landrum RE, Shera CA, Groh JM. Parametric information about eye movements is sent to the ears. *PNAS* 120(48): e2303562120, 2023. https://doi.org/10.1073/pnas.2303562120
12. Gruters KG, Murphy DLK, Jenson CD, Smith DW, Shera CA, Groh JM. The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. *PNAS* 115(6): E1309-E1318, 2018. https://doi.org/10.1073/pnas.1717948115

13. Murphy DLK, King CD, Lovich SN, Landrum RE, Shera CA, Groh JM. Evidence for a system in the auditory periphery that may contribute to linking sounds and images in space. *bioRxiv* 2020.07.19.210864, 2020. https://doi.org/10.1101/2020.07.19.210864
14. Cho NH, Ravicz ME, Puria S. Human middle-ear muscle pulls change tympanic-membrane shape and low-frequency middle-ear transmission magnitudes and delays. *Hearing Research* 430: 108721, 2023. https://doi.org/10.1016/j.heares.2023.108721
15. Tasko SM, Deiters KK, Flamme GA, Smith MV, Murphy WJ, Jones HG, Greene NT, Ahroon WA. Effects of unilateral eye closure on middle ear muscle contractions. *Hearing Research* 424: 108594, 2022. https://doi.org/10.1016/j.heares.2022.108594
16. Bröhl F, Kayser C. Detection of Spatially Localized Sounds Is Robust to Saccades and Concurrent Eye Movement-Related Eardrum Oscillations (EMREOs). *Journal of Neuroscience* 43(45): 7668-7677, 2023. https://doi.org/10.1523/JNEUROSCI.0818-23.2023
17. Binda P, Morrone MC. Vision During Saccadic Eye Movements. *Annual Review of Vision Science* 4: 193-213, 2018 https://doi.org/10.1146/annurev-vision-091517-034317
18. Harris LR, Lieberman L. Auditory stimulus detection is not suppressed during saccadic eye movements. *Perception* 25: 999–1004, 1996. https://doi.org/10.1068/p250999
19. Krüger HM, Collins T, Englitz B, Cavanagh P. Saccades create similar mislocalizations in visual and auditory space. *J Neurophysiol* 115: 2237–2245, 2016. https://doi.org/10.1152/jn.00853.2014
20. Collins T, Heed T, Röder B. Eye-movement-driven changes in the perception of auditory space. *Attention, Perception, & Psychophysics* 72(3): 736-746, 2010. https://doi.org/10.3758/APP.72.3.736
21. Pavani F, Husain M, driver J. Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Experimental brain research* 189(4): 435-449, 2008. https://doi.org/10.1007/s00221-008-1440-7
22. Rorden C, Driver J. Does auditory attention shift in the direction of an upcoming saccade? *Neuropsychologia* 37(3): 357-77, 1999. https://doi.org/10.1016/S0028-3932(98)00072-4
23. Lewald J, Ehrenstein WH. The effect of eye position on auditory lateralization. *Exp Brain Res* 108: 473–485, 1996. https://doi.org/10.1007/BF00227270
24. Lewald J. Eye-position effects in directional hearing. *Behav Brain Res* 87(1): 35-48, 1997. https://doi.org/10.1016/S0166-4328(96)02254-1
25. Getzmann S. The effect of eye position and background noise on vertical sound localization. *Hear Res* 169: 130–139, 2002. https://doi.org/10.1016/S0378-5955(02)00387-8
26. Razavi B, O'Neill WE, Paige GD. Auditory spatial perception dynamically realigns with changing eye position. *Journal of Neuroscience* 27(38): 10249-10258, 2007. https://doi.org/10.1523/JNEUROSCI.0938-07.2007
27. Klingenhoefer S, Bremmer F. Perisaccadic localization of auditory stimuli. *Exp Brain Res* 198: 411–423, 2009. https://doi.org/10.1007/s00221-009-1869-3
28. World Health Organization. World report on hearing (Online). https://www.who.int/publications/i/item/9789240020481 [2021 March 3rd]
29. Brainard DH. The Psychophysics Toolbox. *Spatial Vision* 10(4): 433–436, 1997. https://doi.org/10.1163/156856897X00357
30. The MathWorks Inc. *MATLAB version: 9.13.0 (R2022b)*, 2022. https://www.mathworks.com
31. SR Research Ltd. *EyeLink 1000 Plus.* [Apparatus and software], 2013. https://www.sr-research.com/eyelink-1000-plus/
32. Macmillan NA, Creelman CD. *Detection theory: A user's guide* (2nd ed). Lawrence Erlbaum Associates Publishers, 2005.

33. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review* 16: 225–237, 2009. https://doi.org/10.3758/PBR.16.2.225
34. Racca JM, Delgado RE, Gifford RH, Ramachandran R, Hood LJ. The Effects of Middle-ear Stiffness on the Auditory Brainstem Neural Encoding of Phase. *J Assoc Res Otolaryngol* 23(6): 859-873, 2022. https://doi.org/10.1007/s10162-022-00872-0
35. Lawrence M. Middle ear muscle influence on binaural hearing. *Archives of Otolaryngology* 82(5): 478-482, 1965.
36. Avan P, Büki B, Maat B, Dordain M, Wit HP. Middle ear influence on otoacoustic emissions. I: noninvasive investigation of the human transmission apparatus and comparison with model results. *Hear Res* 140(1–2): 189–201, 2000. https://doi.org/10.1016/S0378-5955(99)00201-4
37. O'Connor KN, Cai H, Puria S. The effects of varying tympanic-membrane material properties on human middle-ear sound transmission in a three-dimensional finite-element model. *The Journal of the Acoustical Society of America* 142(5): 2836–2853, 2017. https://doi.org/10.1121/1.5008741
38. Noppeney U. Perceptual inference, learning, and attention in a multisensory world. *Annual review of neuroscience* 44(1): 449-473, 2021. https://doi.org/10.1146/annurev-neuro-100120-085519
39. Yamagishi S, Furukawa S. Factors Influencing Saccadic Reaction Time: Effect of Task Modality, Stimulus Saliency, Spatial Congruency of Stimuli, and Pupil Size. *Front Hum Neurosci* 26(14): 571893, 2020. https://doi.org/10.3389/fnhum.2020.571893
40. Juravle G, Deubel, H. Action preparation enhances the processing of tactile targets. *Exp Brain Res* 198: 301–311, 2009. https://doi.org/10.1007/s00221-009-1819-0
41. Spence C, Pavani F, Maravita A, Holmes N. Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *Journal of Physiology-Paris*, 98(1-3): 171-189, 2004. https://doi.org/10.1016/j.jphysparis.2004.03.008
42. Rorden C, Greene K, Sasine GM, Baylis GC. Enhanced tactile performance at the destination of an upcoming saccade. *Curr Biol* 12: 1–6, 2002. https://doi.org/10.1016/S0960-9822(02)01039-4
43. Kliegl R, Rolfs M, Laubrock J, Engbert R. Microsaccadic modulation of response times in spatial attention tasks. *Psychological Research* 73: 136–146, 2009. https://doi.org/10.1007/s00426-008-0202-2