Llado Gonzalez, Pedro; Hyvärinen, Petteri; Pulkki, Ville

The impact of head-worn devices in an auditory-aided visual search task

Published in:
Journal of the Acoustical Society of America

DOI:
10.1121/10.0025542

Published: 05/04/2024

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Llado Gonzalez, P., Hyvärinen, P., & Pulkki, V. (2024). The impact of head-worn devices in an auditory-aided visual search task. *Journal of the Acoustical Society of America*, *155*(4), 2460-2469. https://doi.org/10.1121/10.0025542

The impact of head-worn devices in an auditory-aided visual search task

Author's Manuscript of https://doi.org/10.1121/10.0025542

Pedro Lladó, Petteri Hyvärinen, and Ville Pulkki
pedro.llado@aalto.fi

ABSTRACT

Head-worn devices (HWDs) interfere with the natural transmission of sound from the source to the ears of the listener, worsening their localization abilities. The localization errors introduced by HWDs have been mostly studied in static scenarios but these errors are reduced if head-movements are allowed. We studied the effect of twelve HWDs on an auditory-cued visual search task, where head movements were not restricted. In this task, a visual target had to be identified in a three-dimensional space with the help of an acoustic stimulus emitted from the same location as the visual target. The results showed an increase in the search time caused by the HWDs. Acoustic measurements of a dummy head wearing the studied HWDs showed evidence of impaired localization cues, which were used to estimate the perceived localization errors using computational auditory models of static localization. These models were able to explain the search-time differences in the perceptual task, showing the influence of quadrant errors in the auditory-aided visual search task. These results indicate that HWDs have an impact on sound-source localization even when head movements are possible, which may compromise the safety and the quality of experience of the wearer.

I. INTRODUCTION

Sound localization is a fundamental aspect of hearing that helps humans to be aware of their surroundings. The ability to localize sounds may be impaired if the natural transmission of sound to the ears is interfered with or obstructed by head-worn devices (HWDs), such as headphones or hearing-protection devices (HPDs). Perceiving the direction of arrival of a sound is possible due to interaural and monaural localization cues. Interaural cues are the interaural time and level differences in the acoustic signal between the two ears (ITDs and ILDs, respectively), which allow localization in the lateral dimension (left-right). These interaural cues are, however, not enough to fully resolve the direction of the sound in a three-dimensional space, since multiple regions of the space share the same ITD and ILD values. The cone-shaped regions with identical interaural cue values are often referred to as cones of confusion, and monaural cues are needed to resolve the sound direction. These monaural cues emerge from the direction-specific filtering of the external ear, head, and torso on the spectrum of the sound between 1 kHz and 16 kHz. Monaural cues enable the localization in the polar dimension (up-down and front-back) (Algazi *et al.*, 2001; Hebrank and Wright, 1974; Wilska, 1938) and help in resolving the ambiguity within the cone of confusion.

Studies investigating the effects of HWDs on interaural cues in static environments (i.e. where dynamic cues are not available) have shown that wearing hearing protectors, headphones or hearing aids worsens the localization abilities of the users. In the lateral dimension (from the right to the left ear in the range from -90°to 90°, as in the modified interaural-polar coordinates system (Pollack *et al.*, 2022), the effects of these HWDs range from moderate increases in angular errors (Bolia *et al.*, 2001; Brungart *et al.*, 2007, 2003; Cubick *et al.*, 2018; Denk *et al.*, 2018, 2019; Gupta *et al.*, 2018; Lladó *et al.*, 2022a; Van den Bogaert *et al.*, 2008, 2006; Vause and Grantham, 1999) to more extreme cases where wearing earplugs and earmuffs together attenuate the sound below the audibility thresholds, preventing the listeners from discriminating between left and right (Brungart *et al.*, 2004; Simpson *et al.*, 2005).

In the polar dimension, localization errors caused by HWDs are even more pronounced, and they are often related to the alteration of spectral cues and high-frequency attenuation of the HWDs (Brungart *et al.*, 2007, 2003; Denk *et al.*, 2018, 2019; Lladó *et al.*, 2022b; Van den Bogaert *et al.*, 2008, 2006; Vause and Grantham, 1999; Zimpfer and Sarafian, 2014). These devices increase the probability of front-back confusion (Brungart *et al.*, 2007, 2003; Denk *et al.*, 2018, 2019; Van den Bogaert *et al.*, 2008; Vause and Grantham, 1999) and up-down confusion (Denk *et al.*, 2018; Zimpfer and Sarafian, 2014). In a horizontal plane localization experiment, Brown *et al.* (2015) found high correlation between large localization errors and the degradation caused by hearing protectors to the directional transfer functions, showing the importance of monaural cues in such tasks.

Several factors affect the degree to which a HWD may modify the acoustic properties of the sound. In case of passive headphones, the amount of attenuation and pinnae cover may play an important role on the degradation of localization cues, specially of monaural cues

(Alali *et al.*, 2011; Denk *et al.*, 2018; Lladó *et al.*, 2022b). In the case of HWDs with active components, e.g. hearing aids, hear-through headphones, etc., not only these two aspects seem to be important. The position of their microphones and their digital signal processing stages (e.g. compression, to enhance the perceived quality of sound) have an influence on the degradation of localization cues (Alali *et al.*, 2011; Best *et al.*, 2010; Denk *et al.*, 2018, 2019, 2020; Härmä *et al.*, 2004; Talcott *et al.*, 2012; Van den Bogaert *et al.*, 2011, 2006; Zimpfer and Sarafian, 2014).

When localization cues are ambiguous, head movements result in systematic changes in the localization cues. These changes in localization cues are often referred to as dynamic cues, and they improve the localization abilities of the listener (Pöntynen and Salminen, 2019; Wightman and Kistler, 1999). In case of ambiguity between front and back, head yaw rotations result in systematic changes in the interaural cues and aid in resolving the ambiguities (Pöntynen and Salminen, 2019). Thus, if a HWD impairs the localization cues, it is expected that head-movements may at least partly mitigate these negative effects to sound localization. However, they might not be sufficient to completely overcome the effects of wearing a HWD. Localizing sound sources may still take longer if additional head rotations are needed, and this may generate uncertainty in the users when trying to identify or localize sources.

To account for dynamic cues in localization, Bolia *et al.* (1999) introduced a localization-assessment method based on an auditory-aided visual search task. They measured the time to find a target in a three-dimensional space to identify the differences between localizing sounds in virtual scenarios and in real ones. They showed that the participants needed more time to find the target for virtual auditory stimuli than for those presented using actual loudspeakers. The same method has been applied in other studies to assess the participants' ability to localize targets when HWDs are used (Brungart *et al.*, 2007, 2003; Simpson *et al.*, 2005). In these experiments, the same task was conducted as in Bolia *et al.* (1999) to assess the time to find the target while wearing hearing protectors. The time to find the target increased when HPDs were used. Their results showed an increase in the time to localize the target when earplugs and earmuffs were worn, especially when worn together. Using a similar approach, Satongar *et al.* (2015) studied the effect of a circumaural open-back headphone in a dynamic localization task for sources in the horizontal plane. They showed that even circumaural open-back headphones increased the time to find the target.

The proposed visual target localization task is thought to represent an ecologically valid scenario relevant to occupational settings (Simpson *et al.*, 2005). A standardized method based on the work of Bolia *et al.* (1999), proposed in (ANSI/ASA, 2019), included a visual search task to localize a target in the horizontal plane when being simultaneously presented with an acoustic stimulus. This method allowed analyzing how the time to find a target increased in a dynamic situation when wearing a HWD compared to the unobstructed listening condition.

Our study analyzes the connection between the localization ability of listeners when head movements are allowed and the degradation of static localization cues. Twelve HWDs where assessed in a perceptual task that measured the time to find a visual target aided by an acoustic stimulus (Bolia *et al.*, 1999; Simpson *et al.*, 2005). Localization performance was assessed by measuring the time to find the target and it was assumed that worse localization ability would result in increased search times. Moreover, we measured the head-related transfer functions (HRTFs) of a dummy head wearing the studied HWDs. The static localization errors wearing each of the HWDs were estimated by analyzing the HRTFs using computational auditory models. Finally, we tested the estimated static localization errors on their ability to explain the increase in search time in the auditory-aided visual search task.

II. METHODS AND MATERIALS

A. Head-worn devices

Twelve HWDs were assessed, covering a wide range of device types. All of them shared the characteristic that they allow the listener to be aware of the sound in their surroundings.This is achieved either by an open design with passive low attenuation or by using an active system. In the context of this study, active means that there is an active hear-through system that picks up sound from the environment using microphones and plays it back to the ears using an electroacoustic transducer. On the other hand, passive means that the there is no active hear-through system involved, and external sounds are only available through the possibly obstructed acoustic pathway.

The specific models and the settings at which the HWDs were used in the subjective test are listed in Table I. Devices A - C are three passive commercial headphones that could potentially be used in augmented reality applications due to their low attenuation of external sounds. A is a headphone with an open extra-aural design. B is an intra-concha headphone with an open design, since it has a circular driver with a vent, allowing the direct sound to enter the ear canal. Device C is a widely used open-back circumaural studio headphone. Devices D - F are three consumer-grade in-ear headphones with an active hear-through system. These devices were used in hear-through mode. Devices G - I are three consumer-grade circumaural headphones, also with an active hear-trough system and used in hear-through mode. Devices J - L are three HPDs for recreational or professional use with a hear-through system. J and L have a circumaural design and K has a supra-aural design. Devices K and L were used in their default settings, which combine active noise cancellation and hear-through. Device J was used with the hear-through at maximum level. While the included device types were deliberately selected to cover a wide range of HWD-induced degrees of degradation, the specific models were chosen due to their availability in the Aalto Acoustics laboratory at the time of this study.

TABLE I. Summary of the studied head-worn devices and the settings used through the whole study.

ID	Type	Active	Model	Settings
A	extra-aural headphones	no	Mysphere 3.2	open frames
B	intra-concha headphones	no	Sony linkbuds	
C	circumaural headphones	no	Sennheiser HD650	
D	in-ear headphones	yes	Apple airpods pro (1st gen.)	hear-through ON
E	in-ear headphones	yes	Sony WF-1000-XM3	hear-through ON
F	in-ear headphones	yes	Huawei freebuds pro (1st gen)	hear-through ON
G	circumaural headphones	yes	Apple airpods pro MAX	hear-through ON
H	circumaural headphones	yes	Sony WH-1000-XM4	hear-through ON
I	circumaural headphones	yes	Huawei freebuds studio	hear-through ON
J	circumaural HPD	yes	Silenta STP8000	hear-through at max. level
K	supra-aural HPD	yes	Savox Noise-COM 200	hear-through ON (default settings)
L	circumaural HPD	yes	Peltor ComTac XPI	hear-through ON (default settings)

B. Participants

Twenty participants (4 female, 16 male) between 19 and 39 years of age (median = 28.5 years) with self-reported normal hearing and normal or corrected vision participated in this experiment. All participants had participated in multiple listening experiments during the year preceding this study. All participants had experience in using the devices A and C during listening experiments, but were not familiar with the rest of HWDs. The study was done in accordance with the Declaration of Helsinki and was approved by the Research Ethics Committee of Aalto University. The participants provided written informed consent before enrolling in the experiment.

C. Apparatus

The subjective test was conducted in the multichannel anechoic chamber 'Wilska' at the Aalto University Acoustics Lab, in Espoo, Finland. Thirty-two Genelec 8331A coaxial loudspeakers were distributed in a spherical array at a distance of 2.04 m from the center. The loudspeakers were located at four different elevations. In the horizontal plane, $\phi = 0°$, twelve loudspeakers were distributed evenly with an angular distance between consecutive loudspeakers in azimuth $\theta_d = 30°$. In the planes at $\phi = \pm\ 30°$, the loudspeakers were distributed evenly with an angular distance $\theta_d = 45°$. In the plane at $\phi = 60°$, the loudspeakers were distributed evenly with an angular distance $\theta_d = 90°$.

A 2×2 LED matrix (15 mm × 15 mm for the LED centers) was mounted directly in front of each loudspeaker, providing a visual target in the search task. The target loudspeaker's LED matrix always had an even number of LEDs illuminated in red (two or four LEDs randomly between trials). All non-target loudspeakers had an odd number of LEDs illuminated (one or three LEDs randomly for each loudspeaker). The LED system was controlled using an Arduino UNO WiFi Rev2, which was controlled using Max 8 via serial communication. A synchronous auditory stimulus was emitted from the target loudspeaker to help the task of finding the visual target. The sound stimulus was an intermittent pink noise (250 ms on, 250 ms off, 10 ms onset and offset ramps, following the description of the Method 3 in ANSI/ASA S3.71(ANSI/ASA, 2019)) with an A-weighted level of 65 dB SPL measured at the listener's position. The stimulus lasted until the listener gave the response. The lights of the room were dimmed to facilitate the visual search task.

D. Listening test procedure

The procedure used in this study was derived from the experiments by Bolia *et al.* (1999) and Simpson *et al.* (2005), and from the third method in ANSI/ASA (2019). The participants controlled the test and gave their responses using two hand-held response buttons, one in each hand. Each trial began with the participant facing the $(\theta, \phi) = (0°,0°)$ loudspeaker and demonstrating readiness by pressing both buttons at the same time. After the buttons were released, the stimulus presentation started after a 1-s interval. During the stimulus presentation, the auditory stimulus was presented from one target loudspeaker, and the LED matrices in all loudspeakers were illuminated such that only the target loudspeaker had an even number of LEDs illuminated and all the other loudspeakers had an odd number of LEDs illuminated.

In each trial, the participant had to find the target and respond if the number of illuminated LEDs was two or four by pressing the left or right button, respectively. The search time was determined from the time between the stimulus onset and the participant response. After pressing the button, the participants were asked to again face the loudspeaker located at $(\theta, \phi) = (0°,0°)$ before starting the next trial by pressing the two buttons.

The first four targets of each round were located always in the same positions $(\theta, \phi) =$ (0°,0°), (180°,0°), (-90°,30°) and (90°,30°) and marked the beginning of a new listening block (these four targets were discarded from the analysis and repeated during the actual test trials). The 32 following targets were located in random order on each of the loudspeakers. After the 36 trials were completed, the round ended, and there was a short pause during which the participants were asked to change the HWD themselves. In the case of devices B, D, E and F, various sizes of eartips were offered for the participant to choose what was more comfortable. Before the following round started, the researcher in charge of conducting the test confirmed visually that there were no obvious fitting issues.

The experiment consisted of thirteen trial sets: twelve sets with an HWD and one without any HWD (open ears, OE) in random order. Before starting the experiment, three rounds of training, each the same length as the actual test, were conducted. Each subject performed one training round in open ears, one with a random passive HWD (A - C), and one with a random active HWD (D - L). Due to the test procedure, the participants were aware of the headphone they were wearing at all times.

E. Perceptual data analysis

All participants had an overall correct response rate above 98% throughout the experiment. The fastest responses were for frontal locations, and none of them were faster than 386 ms. Therefore, the incorrect responses were interpreted as accidental wrong-button presses, and no responses were discarded. The incorrect responses were also included in determining the response times. For each participant and condition, the median response time was computed and used in group-level analysis.

The group-level data, consisting of the median response time for each participant in each condition, was tested for normality using the Shapiro-Wilk test. Since the data were not normally distributed, the non-parametric Friedman test was used with the Fisher's Least Significant Differences (LSD) test for post-hoc testing with Bonferroni-Holm correction.

F. Acoustic measurements procedure

The HRTFs of a G.R.A.S. KEMAR 45BC head and torso simulator with antropometric pinnae (Shore 00-35) and ear simulators (RA0045) were measured wearing the studied devices. The measurements took place in the multichannel anechoic room 'Wilska' in the Aalto Acoustics Lab, the same room as the perceptual test. The measurements were performed using the multiple exponential sweep method (Majdak *et al.*, 2007).

The impulse response measurement of the loudspeaker located in the angle $(\theta, \phi) = (0°,0°)$ was repeated three times, removing and repositioning the HWD between measurements. If these three transfer functions fell within a 2 dB difference at all frequencies in the range 250 to 4000 Hz for both ears, the positioning was considered correct. If this was the case, the HRTF was measured every 5° in the horizontal plane using a digitally controlled turntable, and every 30° in the median plane. If the positioning was considered incorrect, three new measurements were taken until they matched the self-imposed criterion. The HWDs with active hear-through were also measured using the same stimulus as in the perceptual experiment (A-weighted level of 65 dB SPL, 250 ms on, 250 ms off, 10 ms onset and offset ramps). These additional measurements aimed at analyzing potential non-linearities that might have affected the impulse response analysis. The temporal and spectral analysis of the measurements did not differ substantially, and therefore impulse responses from the exponential sine sweep measurements were used for the rest of the study (see Sec. VIII).

G. Acoustic measurements analysis

The acoustic measurements were analyzed to see if they had explanatory power on the perceptual results. The HRTFs were free-field equalized to minimize the effect of the measurement microphones and loudspeakers (Møller, 1992). An analysis of the ITD and ILD differences were computed for the horizontal plane measurements to analyze the effect of HWDs in the left-right dimension. The impulse responses were filtered by a gammatone filterbank (Glasberg and Moore, 1990; Lyon, 1997) with a frequency spacing corresponding to one equivalent rectangular bandwidth. The ITD estimates for each filter output were computed by finding the maximum in the interaural cross-correlation (Jeffress, 1948) in the range $[-670, 670]\,\mu$s. The ILD was computed as the RMS difference between left and right ears in dB after applying the same filtering. The broadband ITD and ILD values were computed as the average in the frequency range from 50 Hz to 1.5 kHz for ITD, and from 1 kHz to 8 kHz for ILD.

However, translating frequency dependent ITD and ILD estimates into estimated angles is not trivial. Thus, the model introduced in Lladó *et al.* (2022a) was used to estimate the perceived lateral angle, e_{lat}, and localization blur e_{blur} in the horizontal plane, where the lateral angle is described by $\vartheta \in [-90°, 90°]$ from the right to left ear (Pollack *et al.*, 2022). This model is based on a neural network that was trained to predict the effect of HWDs on localization in the frontal horizontal plane. This model was not previously verified for all the tested conditions, but the cross-validation in Lladó *et al.* (2022a) suggests that it is able to predict unseen data to some extent. The model is available in the Auditory Modelling Toolbox Majdak *et al.* (2022).

The model from Lladó *et al.* (2022a) returns for each target angle, an estimate of the perceived lateral angle, e_{lat}, and the localization blur, e_{blur}. The median absolute deviation (MAD) from the OE condition was computed for each HWD, h:

$$\mathrm{MAD}^h = \mathrm{median}(|e_\vartheta^h - e_\vartheta^{\mathrm{OE}}|), \tag{1}$$

where the median is computed over target lateral angles ϑ.

On the other hand, the possibility of localizing sound sources in sagittal planes, i.e. parallel to the median plane that divides the body in left and right, relies on cues extracted from the magnitude spectrum of the sound Macpherson and Middlebrooks (2002). Thus, the direction-dependent effect of each HWD on the magnitude spectrum of the sound was computed in the median plane every $\varphi_{\mathrm{d}} = 30°$, where $\varphi \in [-90°, 270°)$ is the polar angle from front below to back below (Pollack *et al.*, 2022).

Similar to the interaural cues, translating the spectral information into localization in sagittal planes is not obvious. Baumgartner *et al.* (2014) proposed a model that estimates quadrant-error rates (QE) and polar errors (PE) in sagittal plane localization given a binaural impulse response. The model is available at Majdak *et al.* (2022). This model has proven to be useful to estimate impairments induced by HWDs (Lladó *et al.*, 2022b). Thus, it is used here as a tool to understand the relation between the magnitude response of the HRTF measurements and the perceptual test results. The HRTF in the OE condition is used as the template, and the HRTFs of the KEMAR wearing each HWD are used as the targets. The model parameters are adopted from Lladó *et al.* (2022b) (degree of selectivity $\Gamma = 17$ dB^{-1}, sensitivity $\mathrm{S} = 0.35$ and motor response scatter $\varepsilon = 27°$; see Baumgartner *et al.* (2014) for a detailed description of the parameters). This model was not previously verified for all the tested conditions but the model parameters were obtained from the OE condition only.

The model from Baumgartner *et al.* (2014) returns a single pair of estimates for QE and PE, e_{QE} and e_{PE} respectively, for all target directions in the median plane together. The absolute differences to the OE condition were computed as $\Delta e^h_{\mathrm{QE}} = |e^h_{\mathrm{QE}} - e^{\mathrm{OE}}_{\mathrm{QE}}|$ and $\Delta e^h_{\mathrm{PE}} = |e^h_{\mathrm{PE}} - e^{\mathrm{OE}}_{\mathrm{QE}}|$.

A linear model is used to describe the median search times from the listening test, T^h, as a function of the model-estimated localization metrics:

$$T^h = \beta_0 + \beta_1 \mathrm{MAD}^h_{\mathrm{lat}} + \beta_2 \mathrm{MAD}^h_{\mathrm{blur}} + \beta_3 \Delta e^h_{\mathrm{QE}} + \beta_4 \Delta e^h_{\mathrm{PE}}, \tag{2}$$

III. RESULTS

A. Perceptual experiment results

The distribution of the median search times for each participant are summarized in Figure 1. The median time in the OE condition was 0.96 s. The median search time in the slowest condition, K, was 1.68 s. A Friedman test was conducted to compare the effect of HWDs on the median time to find the target. The effect of the HWDs reached statistical significance, with $\chi^2(12,\ 228) = 154.55$, $p < 0.01$. The Fisher's LSD post-hoc tests with Bonferroni-Holm correction showed statistically significant differences (see Table II).

For a clear and concise presentation of the directional-dependent analysis of the results, a subset of three exemplary HWDs has been selected. However, the results of the same analysis for all HWDs is included as supplementary material (see Section VIII). The HWDs included in the subset are A (extra-aural headphones), which yielded the best performance of all studied HWDs, F (in-ear headphones with hear-through), which yielded the best performance among active hear-through HWDs, and K (supra-aural HPD with hear-through), which yielded the worst performance overall.

The median search time over the studied directions is shown in Figure 2 for the subset of HWDs. The effect of the source location in the OE condition seems to be mild, since the times to find the target have similar values both over azimuth and over elevation angles. For the HWD A, which was the closest in the median search time to the OE condition, the influence of the azimuth and elevation angles is also mild, showing a compact distribution of times similar to the OE condition. In the other extreme, the device that produced the largest median time to find the target, K, shows a different behavior. The responses are similar to OE at $(\theta,\phi) = (90°,0°)$ and $(270°,0°)$ only, suggesting increases in front-back confusions and/or elevation errors at all other locations. The results from the HWD F show an intermediate step between A and K, where the results in the horizontal plane for $\theta = 0°$, 90° and 270° are similar to the results of A. For the rest of angles, the values increased, but the directional effect of A seems milder than the effect of K for all the studied angles except for $(\theta,\phi) = (180°,0°)$. An ad-hoc two-way ANOVA test was conducted to verify the findings, and revealed a statistically significant interaction between the effects of the HWD and the direction ($F(372,\ 7904) = 2.17$, $p < .01$).

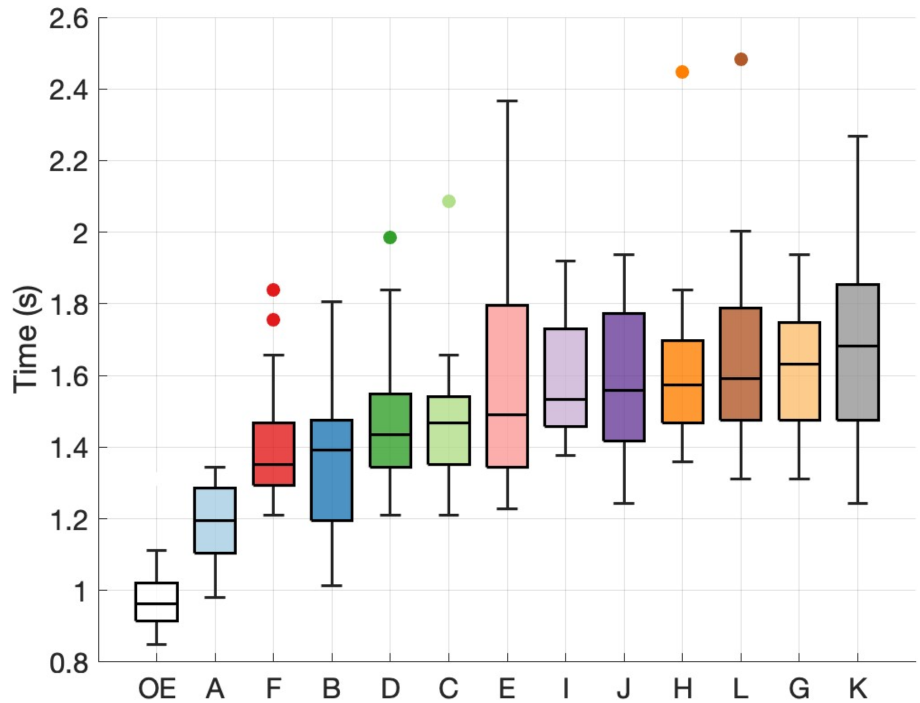


FIG. 1. (Color online) Median time to find the target for each participant when wearing each studied HWD and in OE condition. The boxes represent interquartile ranges. The devices are shown sorted by the median search time.

B. Acoustic measurements analysis results

The ITD and ILD estimates computed from the measurements are shown in Figure 3ITD and Figure 3ILD. The ITD and ILD for the OE condition show the expected behavior, which enables an accurate localization in lateral angle. Despite some differences, the ITD and ILD analysis for the HWDs A and F show a similar behaviour than for the OE conditions, both broadband and for the presented frequency bands. In the case of K, the ITD seems to be preserved, but the computed ILD estimates show an inconsistent behaviour over angle and over frequency.

The model of auditory localization introduced in Lladó *et al.* (2022a) was used to estimate the lateral angle from the frequency dependent ITD and ILD estimates. The estimated angles are shown in 3Estimated angle, and the $\mathrm{MAD_{lat}}$ and $\mathrm{MAD_{blur}}$ for each HWD are

TABLE II. Summary of the statistical analysis. The asterisks denote statistically significant differences among conditions using Fisher's LSD posthoc test with Bonferroni-Holm corrections ($p < 0.01$).

	OE	A	F	B	D	C	E	I	J	H	L	G	K
OE					*		*	*	*	*	*	*	*
A							*	*	*	*	*	*	*
F											*		*
B								*		*	*	*	*
D	*												
C													
E	*	*											
I	*	*		*									
J	*	*											
H	*	*		*									
L	*	*	*	*									
G	*	*		*									
K	*	*	*	*									

shown in Table III. The OE condition follows a compressed range of angles compared to the actual sound source locations ranging from about -60° $< \vartheta <$ 60°, instead of the covering the whole frontal horizontal plane. Despite this compression, the estimated angles follow the expected trends when compared to the source locations. The HWDs A and F followed a similar response as in the OE condition. On the other hand, the estimated angles for the HWD K presented a larger amount of error when compared to the actual sources locations.

The direction-dependent effect of each HWD on transmitted magnitude spectra is summarized in Figure 4. From the magnitude of the HRTFs for the front direction, it is shown that the main characteristics of the HRTF structure are not well maintained for all the studied HWDs. The devices A and F seem to keep the magnitude response similar to the OE condition up to the most prominent peak at around 3 kHz. However, the profiles of the HRTF at the location of the most prominent notch already present clear differences. The

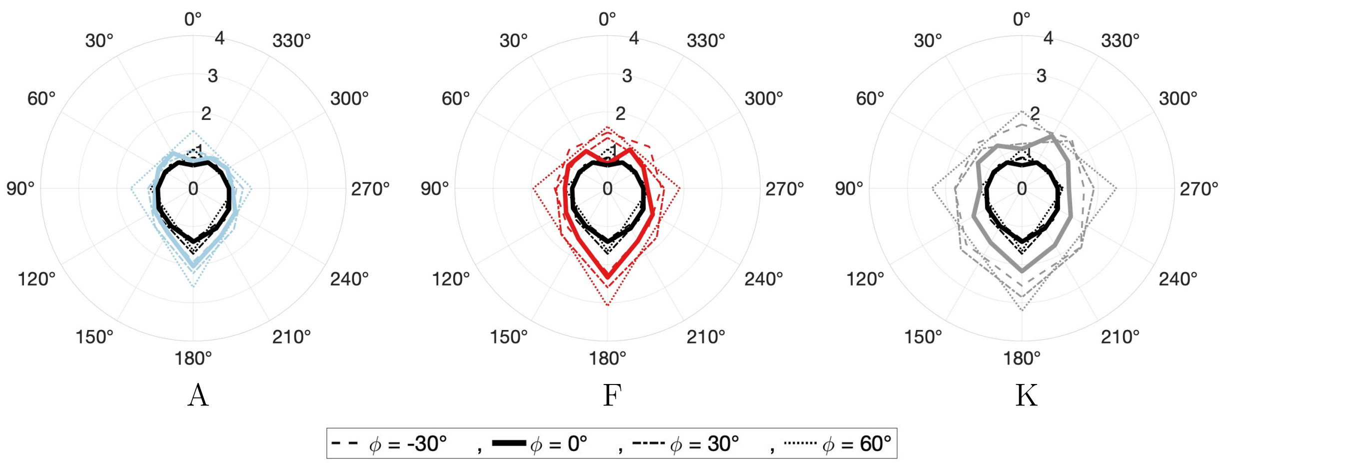
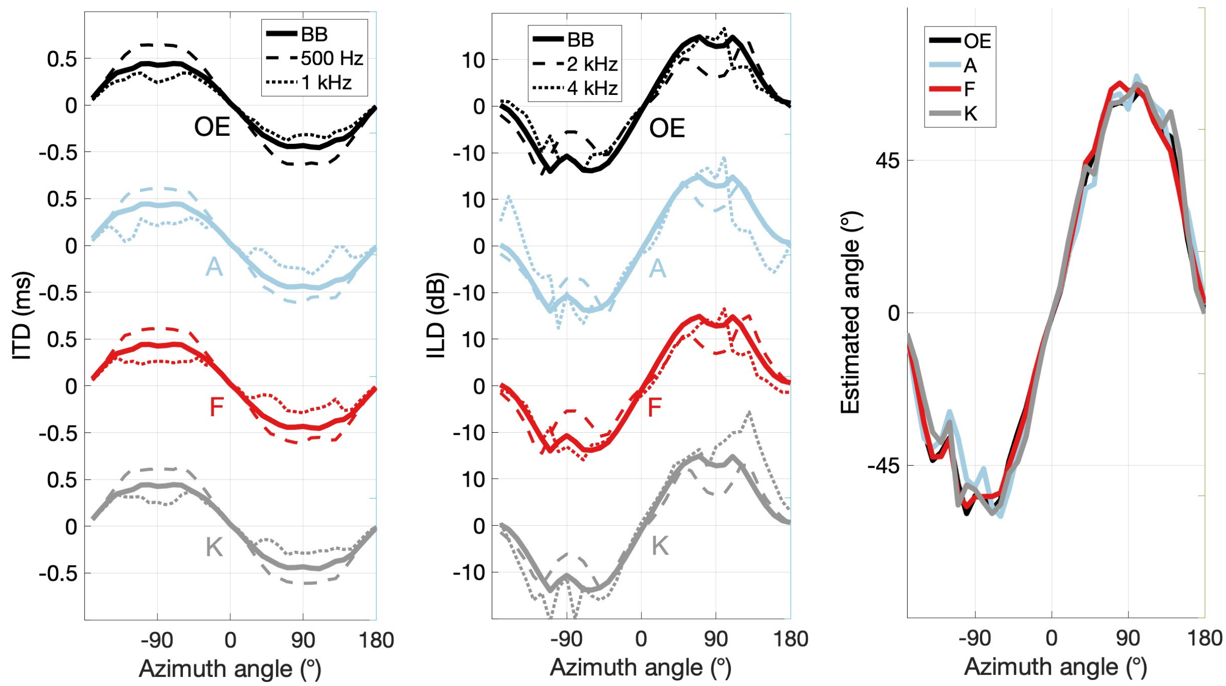


FIG. 2. (Color online) Median time to find the target across participants for each studied angle for the subset of HWDs (color) and the OE condition (black). The radius indicates the time in seconds.

devices F and K present non-negligible attenuation at frequencies above 4 kHz, which may prevent the listeners from using important information for sagittal-plane localization.

Differences in the magnitude spectrum of the HRTF among front, top and back directions are present in the OE case, which allow sagittal plane localization. Even though differences over direction are also present for the device A, they differ both in magnitude and frequency to the OE case. This may generate problems when resolving the angle of arrival in a sagittal plane. For devices F and K, the differences seem to be smaller across source locations. Thus, the differences among directions are reduced, which makes it more challenging to resolve the location of a sound source in sagittal plane.

The effect of the HWDs on median plane localization was estimated using the model from Baumgartner *et al.* (2014). The results of these estimates for the subset of devices are shown in Figure 5. The estimates show that in the OE a diagonal pattern is apparent, which predicts that the estimated elevation corresponds to the elevation of the source. The estimated behaviour in condition A is similar to the OE condition with subtle errors. For the device F, the diagonal is partially preserved, but the localization errors increase. In the condition K, no diagonal pattern can be observed and a bias is present at polar angle 120°, suggesting that the median plane localization may be impaired. The Δe_{QE} and Δe_{QE} for each HWD are shown in Table III.



ITD ILD Estimated angle

FIG. 3. (Color online) Left and center: interaural time (left) and level (center) differences computed broadband and for two exemplary frequency bands for sources in the horizontal plane. Right: estimated angle computed from the ITD and ILD estimates using the model in Lladó *et al.* (2022a).

The coefficients of the fitted linear model are: $\beta_0 = 0.953, \beta_1 = -0.004, \beta_2 = 0.370, \beta_3 = 0.012, \beta_4 = -0.008$. The adjusted $R^2 = 0.93$ shows that these estimated localization metrics are able to explain most of the variance of the listening test median search times. The Δe_{QE} is the variable that alone is able to explain the largest amount of variance (adjusted $R^2 = 0.79$).

IV. DISCUSSION

The results of the auditory-aided visual search task show that HWDs affect localization even in dynamic situations. The results confirm that HWDs have a significant impact on

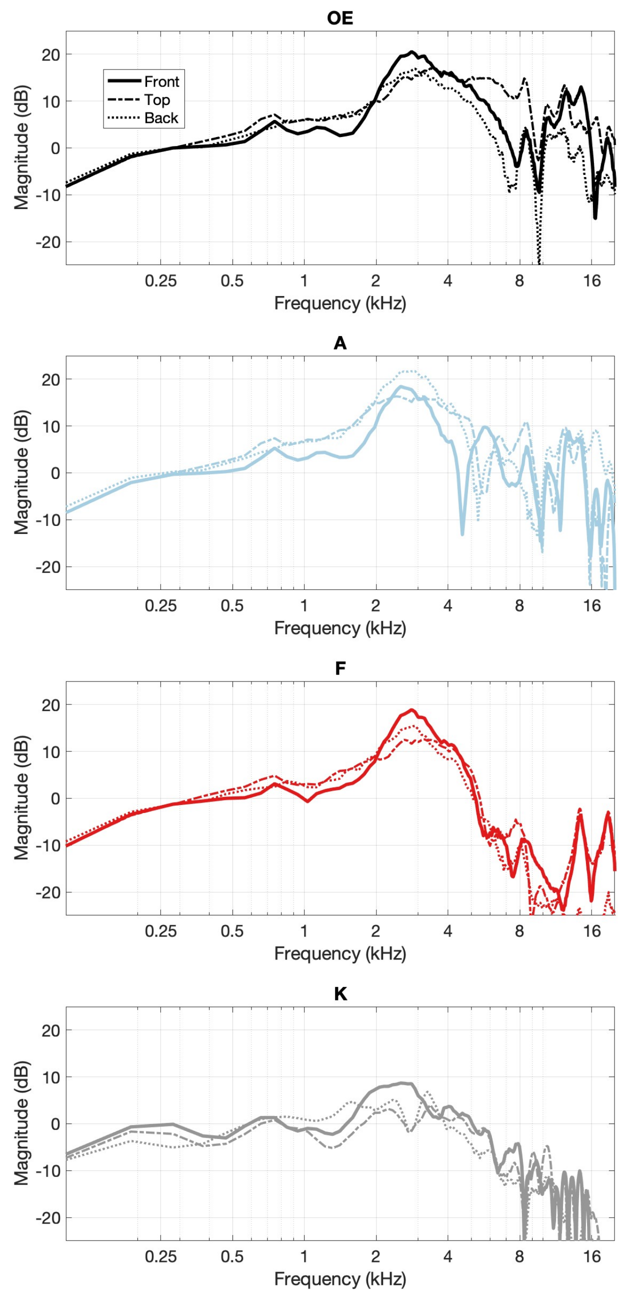


FIG. 4. (Color online) Magnitude spectrum of the HRTF in the directions front, top and back in the median plane for a subset of HWDs.

the response times in a visual search task aided by an acoustic stimulus. Head movements are, therefore, not enough to completely overcome the detrimental effect caused by the HWDs on localization. It is important to note that the stimulus in this experiment was an intermittent pink noise of 250 ms-intervals following ANSI/ASA (2019) and consistent with previous studies (Bolia *et al.*, 1999; Simpson *et al.*, 2005). Binaural cues are considered

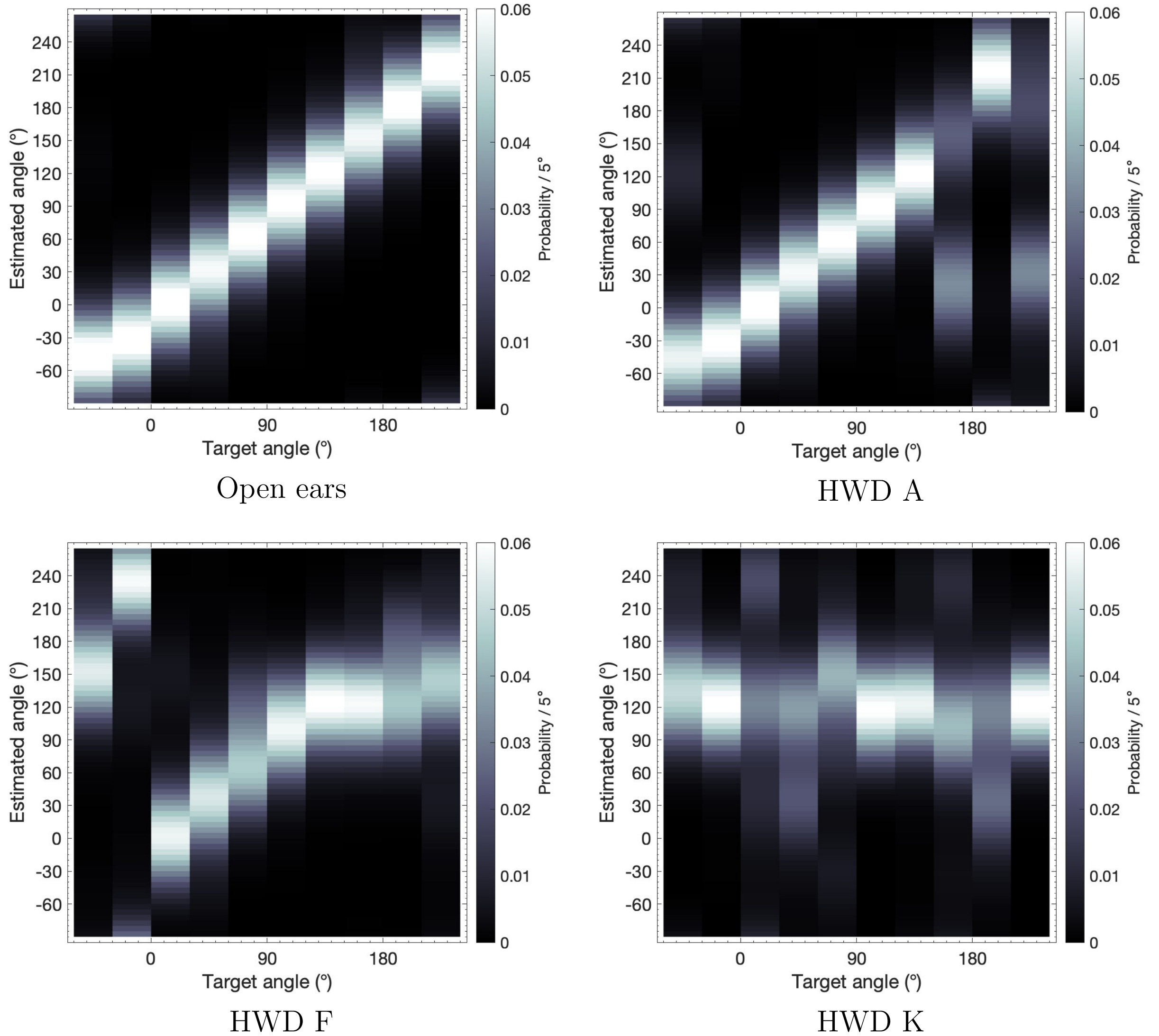


FIG. 5. Localization estimation in the median plane using the model proposed in Baumgartner *et al.* (2014). The brighter the color of the plot, the higher the probability of the model prediction.

useful when the duration of the stimulus is about 300 ms or longer (Blauert, 1997). Thus, it is not clear whether the times in the search task would remain the same for continuous noise or if the periodicity of the intermittent noise changed. This remains an open question that should be addressed for a better understanding of the temporal integration of dynamic cues.

The stimuli used in the test included both visual and auditory cues as in Bolia *et al.* (1999); Simpson *et al.* (2005). Visual search has been extensively studied (Eckstein, 2011; Wolfe, 1994, 2021; Wolfe and Horowitz, 2017). In such a task, the time to find the target

TABLE III. Differences in the auditory-model estimated localization metrics of each HWD compared with respect to the OE condition.

HWD	$\mathrm{MAD_{lat}}$ (°)	$\mathrm{MAD_{blur}}$ (°)	Δe_{QE} (%)	Δe_{PE} (°)
A	2.72	0.52	11.43	2.45
F	3.27	0.51	23.13	13.26
B	1.47	0.28	37.13	11.89
D	5.00	0.72	42.33	22.30
C	2.99	0.71	29.15	12.05
E	23.74	0.67	45.82	21.70
I	11.89	0.87	34.88	15.79
J	10.29	0.78	46.73	21.99
H	12.67	0.76	49.50	23.28
L	4.42	0.57	44.59	19.77
G	4.13	1.04	34.95	17.73
K	12.33	0.91	48.90	21.32

depends on multiple factors, such as the number of distractors, their shape, their size, etc. However, in our experiment the visual information was kept constant over conditions, and we assume that the increase in search time should be explained by means of the auditory cues degradation. Nonetheless, the increase in search time depends on the number of visual distractors, specially when the auditory information is degraded (Rudmann and Strybel, 1999; Simpson *et al.*, 2005).

The results of the linear model show the relation of localization errors in static conditions with the search time in the auditory-aided visual search task. In particular, increases in search times seem to be connected to increases in the model-estimated quadrant error rates. This result agrees with the findings from Brown *et al.* (2015), where the amount of front-back confusions could be explained by the degradation of spectral cues. Moreover, our results suggest that spectral cues may also play a pivotal role beyond static localization.

Headphones A and B present an open design, trying to minimize the amount of attenuation that the headphones introduce to external sound. This explains the preservation of localization cues that are still available to the listener in condition A, which translated into a reduced time to find the target (condition B is discussed later due to large inter-subject differences). The differences between A and C are in accordance with those found in a previous study that analyzed the localization errors for the static condition only (Lladó *et al.*, 2022b).

It has been shown in the past that HWDs with active hear-through introduced larger localization errors than the passive ones (Zimpfer and Sarafian, 2014). An interesting outcome from our study is that HWDs D - F (in-ear headphones with an active hear-through system) resulted in similar performance as with HWDs A - C (passive headphones). This may be explained by the location of the microphones on the in-ear earphones, possibly resulting in ITD and ILD values more similar to OE condition (see Figure 3) (Denk *et al.*, 2019). Moreover, the in-ear headpone microphones may pick up some cues caused by the filtering of the pinnae, reducing the amount of front-back confusions and elevation errors (see Figure 4 and Figure 5) (Best *et al.*, 2010; Denk *et al.*, 2019; Van den Bogaert *et al.*, 2011). Thus, the degradation of localization cues is not necessarily related to the device type and it is necessary to analyze the acoustic characteristics to estimate its effect.

The search time results for HWDs G - L were similar. Both the lack of pinna cues and the active hear-through strategies are similar in these six devices, and therefore a similar result was expected. Even though the in-ear headphones also use active hear-through, the effect on the time to find the target seems to be milder than for the HWDs that cover the pinnae. From the similarities in the HWDs G - L, one could think that there is a ceiling effect in the perceptual test results. However, in the pilot phase of the study we included a condition where no acoustic stimulus was presented, and the LED pattern had to be found without auditory guidance, as in Simpson *et al.* (2005). The median search time for the four participants in the pilot was over 7 s, and was therefore left out of the actual test, since it provided little information and increased the experiment time considerably. In any case, it is possible that there is a ceiling effect, once the listeners have access to left-right auditory information.

Listeners' adaptation to a new set of localization cues is expected to some extent when wearing a HWD (Audet *et al.*, 2022; Casali and Robinette, 2015). This adaptation process can be faster if there is feedback or explicit training involved (Mendonça *et al.*, 2012).

However, this adaptation may not be possible when these new cues are ambiguous (Carlile *et al.*, 2014; Denk *et al.*, 2018), i.e., do not lead to a single location in the space, e.g. when the spectral cues do not help to resolve the polar angle (see Figure 4 and Figure 5). Thus, maintaining the cues as natural as possible may be the best strategy to ensure a correct localization ability of the listeners. If that is not possible, then it seems necessary to provide at least enough difference over angle for the wearer to learn.

Each device may have a unique effect on each subject's localization performance, especially on elevation localization and the rate of front-back confusions, due to the idiosyncratic nature of the pinna cues. Thus, future work could consider measuring acoustically the effect of each HWD fitting on each subject, together with individual data of their HRTF with and without the studied devices. Moreover, head movements could be tracked to understand the rotations that subjects perform when trying to localize a sound source. This would help understanding the acoustic factors that induce the impairment in the dynamic localization task.

V. CONCLUSIONS

An experiment was conducted to measure the time to find a target when a head-worn device is worn using an auditory-aided visual search task. The increase in localization errors when wearing head-worn devices has been previously reported in numerous studies due to the degradation of localization cues. Our results are aligned with these findings and suggest that when head movements are possible, the detrimental effects caused by head-worn devices on localization are not completely overcome. The search time to locate the source significantly increased compared to when no head-worn device was worn. Devices that degrade the localization cues less, e.g. due to their physical design or microphone position, had a lesser impact.

The acoustic measurements with a binaural manikin showed that the interaural cues are often altered by HWDs. In addition, the spectral profile of the magnitude response was found to be affected by most of the studied devices in the measurements. Based on our model-based analysis, the increases in search time seem to be related to the degradations of spectral cues, especially for locations susceptible to front-back confusions and elevated sources. Thus, this experiment supports the importance of maintaining accurate spectral cues even when the task allows head movements.

VI. AUTHORS DECLARATIONS

The authors have no conflicts of interest to declare.

VII. DATA AVAILABILITY

The measurement data and the code to reproduce the model-based analysis will be available online at the time of publication.

VIII. SUPPLEMENTARY MATERIAL

See supplementary material at [URL will be inserted by AIP] for extended information of the studied head-worn devices for Figure 2, Figure 3, Figure 4 and Figure 5. Two additional figures are included for an analysis of potential non-linearities of the HWDs with active hear-through.

ACKNOWLEDGMENTS

This research was supported by Business Finland as part of the Human Optimised XR (HumOR) Project.

Alali, K. A., Casali, J. G. *et al.* (**2011**). “The challenge of localizing vehicle backup alarms: effects of passive and electronic hearing protectors, ambient noise level, and backup alarm spectral content,” Noise and Health **13**(51), 99.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001**). “Elevation localization and head-related transfer function analysis at low frequencies,” The Journal of the Acoustical Society of America **109**(3), 1110–1122.

ANSI/ASA (**2019**). "Methods for measuring the effect of head-worn devices on directional sound localization in the horizontal plane," ANSI/ASA S3.71, 2019 Edition .

Audet, D. J., Gray, W. O., and Brown, A. D. (**2022**). "Audiovisual training rapidly reduces potentially hazardous perceptual errors caused by earplugs," Hearing Research **414**, 108394.

Baumgartner, R., Majdak, P., and Laback, B. (**2014**). "Modeling sound-source localization in sagittal planes for human listeners," The Journal of the Acoustical Society of America **136**(2), 791–802.

Best, V., Kalluri, S., McLachlan, S., Valentine, S., Edwards, B., and Carlile, S. (**2010**). "A comparison of cic and bte hearing aids for three-dimensional localization of speech," International Journal of Audiology **49**(10), 723–732.

Blauert, J. (**1997**). *Spatial Hearing: the Psychophysics of Human Sound Localization* (MIT press), pp. 178–191.

Bolia, R. S., D'Angelo, W. R., and McKinley, R. L. (**1999**). "Aurally aided visual search in three-dimensional space," Human factors **41**(4), 664–669.

Bolia, R. S., D'Angelo, W. R., Mishler, P. J., and Morris, L. J. (**2001**). "Effects of hearing protectors on auditory localization in azimuth and elevation," Human Factors **43**(1), 122–128.

Brown, A. D., Beemer, B. T., Greene, N. T., Argo IV, T., Meegan, G. D., and Tollin, D. J. (**2015**). "Effects of active and passive hearing protection devices on sound source localization, speech recognition, and tone detection," PLoS One **10**(8), e0136568.

Brungart, D. S., Hobbs, B. W., and Hamil, J. T. (**2007**). "A comparsion of acoustic and psychoacoustic measurements of pass-through hearing protection devices," in *2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, IEEE, pp. 70–73.

Brungart, D. S., Kordik, A. J., Eades, C. S., and Simpson, B. D. (**2003**). "The effect of microphone placement on localization accuracy with electronic pass-through earplugs," in *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No. 03TH8684)*, IEEE, pp. 149–152.

Brungart, D. S., Kordik, A. J., and Simpson, B. D. (**2004**). "The effects of single and double hearing protection on the localization and segregation of spatially-separated speech signals (l)," The Journal of the Acoustical Society of America **116**(4), 1897–1900.

Carlile, S., Balachandar, K., and Kelly, H. (**2014**). "Accommodating to new ears: the effects of sensory and sensory-motor feedback," J. Acoust. Soc. Am. **135**(4), 2002–2011.

Casali, J. G., and Robinette, M. B. (**2015**). "Effects of user training with electronically-modulated sound transmission hearing protectors and the open ear on horizontal localization ability," International Journal of Audiology **54**(sup1), S37–S45.

Cubick, J., Buchholz, J. M., Best, V., Lavandier, M., and Dau, T. (**2018**). "Listening through hearing aids affects spatial perception and speech intelligibility in normal-hearing listeners," The Journal of the Acoustical Society of America **144**(5), 2896–2905.

Denk, F., Ewert, S. D., and Kollmeier, B. (**2018**). "Spectral directional cues captured by hearing device microphones in individual human ears," The Journal of the Acoustical Society of America **144**(4), 2072–2087.

Denk, F., Ewert, S. D., and Kollmeier, B. (**2019**). "On the limitations of sound localization with hearing devices," The Journal of the Acoustical Society of America **146**(3), 1732–1744.

Denk, F., Schepker, H., Doclo, S., and Kollmeier, B. (**2020**). "Acoustic transparency in hearables - technical evaluation," Journal of the Audio Engineering Society **68**(7/8), 508–521.

Eckstein, M. P. (**2011**). "Visual search: A retrospective," Journal of vision **11**(5), 14–14.

Glasberg, B. R., and Moore, B. C. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hearing research **47**(1-2), 103–138.

Gupta, R., Ranjan, R., He, J., and Woon-Seng, G. (**2018**). "Investigation of effect of vr/ar headgear on head related transfer functions for natural listening," in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society.

Härmä, A., Jakka, J., Tikander, M., Karjalainen, M., Lokki, T., Hiipakka, J., and Lorho, G. (**2004**). "Augmented reality audio for mobile and wearable appliances," Journal of the Audio Engineering Society **52**(6), 618–639.

Hebrank, J., and Wright, D. (**1974**). "Spectral cues used in the localization of sound sources on the median plane," The Journal of the Acoustical Society of America **56**(6), 1829–1834.

Jeffress, L. A. (**1948**). "A place theory of sound localization.," Journal of comparative and physiological psychology **41**(1), 35.

Lladó, P., Hyvärinen, P., and Pulkki, V. (**2022**a). "Auditory model-based estimation of the effect of head-worn devices on frontal horizontal localisation," Acta Acustica **6**, 1.

Lladó, P., McKenzie, T., Meyer-Kahlen, N., and Schlecht, S. J. (**2022**b). "Predicting perceptual transparency of head-worn devices," Journal of the Audio Engineering Society .

Lyon, R. F. (**1997**). "All-pole models of auditory filtering," Diversity in auditory mechanics 205–211.

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited," The Journal of the Acoustical Society of America **111**(5), 2219–2236.

Majdak, P., Balazs, P., and Laback, B. (**2007**). "Multiple exponential sweep method for fast measurement of head-related transfer functions," Journal of the Audio Engineering Society **55**(7/8), 623–637.

Majdak, P., Hollomey, C., and Baumgartner, R. (**2022**). "Amt 1. x: A toolbox for reproducible research in auditory modeling," Acta Acustica **6**, 19.

Mendonça, C., Campos, G., Dias, P., Vieira, J., Ferreira, J. P., and Santos, J. A. (**2012**). "On the improvement of localization accuracy with non-individualized hrtf-based sounds," Journal of the Audio Engineering Society **60**(10), 821–830.

Møller, H. (**1992**). "Fundamentals of binaural technology," Applied acoustics **36**(3-4), 171–218.

Pollack, K., Kreuzer, W., and Majdak, P. (**2022**). "Perspective chapter: Modern acquisition of personalised head-related transfer functions–an overview," Advances in Fundamental and Applied Research on Spatial Audio .

Pöntynen, H., and Salminen, N. H. (**2019**). "Resolving front-back ambiguity with head rotation: The role of level dynamics," Hearing research **377**, 196–207.

Rudmann, D. S., and Strybel, T. Z. (**1999**). "Auditory spatial facilitation of visual search performance: Effect of cue precision and distractor density," Human Factors **41**(1), 146–160.

Satongar, D., Pike, C., Lam, Y. W., Tew, A. I. *et al.* (**2015**). "The influence of headphones on the localization of external loudspeaker sources," Journal of the Audio Engineering Society **63**(10), 799–810.

Simpson, B. D., Bolia, R. S., McKinley, R. L., and Brungart, D. S. (**2005**). "The impact of hearing protection on sound localization and orienting behavior," Human Factors **47**(1), 188–198.

Talcott, K. A., Casali, J. G., Keady, J. P., and Killion, M. C. (**2012**). "Azimuthal auditory localization of gunshots in a realistic field environment: effects of open-ear versus hearing protection-enhancement devices (hpeds), military vehicle noise, and hearing impairment," International journal of audiology **51**(sup1), S20–S30.

Van den Bogaert, T., Carette, E., and Wouters, J. (**2011**). "Sound source localization using hearing aids with microphones placed behind-the-ear, in-the-canal, and in-the-pinna," International Journal of Audiology **50**(3), 164–176.

Van den Bogaert, T., Doclo, S., Wouters, J., and Moonen, M. (**2008**). "The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids," The Journal of the Acoustical Society of America **124**(1), 484–497.

Van den Bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., and Wouters, J. (**2006**). "Horizontal localization with bilateral hearing aids: Without is better than with," The journal of the acoustical society of America **119**(1), 515–526.

Vause, N. L., and Grantham, D. W. (**1999**). "Effects of earplugs and protective headgear on auditory localization ability in the horizontal plane," Human Factors **41**(2), 282–294.

Wightman, F. L., and Kistler, D. J. (**1999**). "Resolution of front–back ambiguity in spatial hearing by listener and source movement," The Journal of the Acoustical Society of America **105**(5), 2841–2853.

Wilska, A. (**1938**). "Studies on directional hearing," Ph.D. dissertation. English translation, Aalto University School of Science and Technology, Dept. of Signal Processing and Acoustics, 2010. Ph.D. dissertation originally published in German as "Untersuchungen über das Richtungshören", University of Helsinki .

Wolfe, J. M. (**1994**). "Guided search 2.0 a revised model of visual search," Psychonomic bulletin & review **1**, 202–238.

Wolfe, J. M. (**2021**). "Guided search 6.0: An updated model of visual search," Psychonomic Bulletin & Review **28**(4), 1060–1092.

Wolfe, J. M., and Horowitz, T. S. (**2017**). "Five factors that guide attention in visual search," Nature Human Behaviour **1**(3), 0058.

Zimpfer, V., and Sarafian, D. (**2014**). "Impact of hearing protection devices on sound localization performance," Frontiers in Neuroscience **8**, 135.