Effects of type of emission and masking sound, and their spatial correspondence, on blind and sighted people's ability to echolocate

Thaler[1], L., Castillo-Serrano[1], J.G., Kish[2], D. & Norman[1], L.J.

1- Department of Psychology, Durham University, South Road, Durham DH1 5AY, UK

2- World Access for the Blind, 1007 Marino Drive, Placentia, California 92870, USA

Corresponding author:

Lore Thaler
lore.thaler@durham.ac.uk
Department of Psychology, Durham University
Science Site, South Road
Durham DH1 3LE
United Kingdom

Abstract

Ambient sound can mask acoustic signals. The current study addressed how echolocation in people is affected by masking sound, and the role played by type of sound and spatial (i.e. binaural) similarity. We also investigated the role played by blindness and long-term experience with echolocation, by testing echolocation experts, as well as blind and sighted people new to echolocation. Results were obtained in two echolocation tasks where participants listened to binaural recordings of echolocation and masking sounds, and either localized echoes in azimuth or discriminated echo audibility. Echolocation and masking sounds could be either clicks or broad band noise. An adaptive staircase method was used to adjust signal-to-noise ratios (SNRs) based on participants' responses. When target and masker had the same binaural cues (i.e. both were monoaural sounds), people performed better (i.e. had lower SNRs) when target and masker used different types of sound (e.g. clicks in noise-masker or noise in clicks-masker), as compared to when target and masker used the same type of sound (e.g. clicks in click-, or noise in noise-masker). A very different pattern of results was observed when masker and target differed in their binaural cues, in which case people always performed better when clicks were the masker, regardless of type of emission used. Further, direct comparison between conditions with and without binaural difference revealed binaural release from masking only when clicks were used as emissions and masker, but not otherwise (i.e. when noise was used as masker or emission). This suggests that echolocation with clicks or noise may differ in their sensitivity to binaural cues. We observed the same pattern of results for echolocation experts, and blind and sighted people new to echolocation, suggesting a limited role played by long-term experience or blindness. In addition to generating novel predictions for future work, the findings also inform instruction in echolocation for people who are blind or sighted.

Keywords: blindness; audition; release from masking; SNR; masking

1 Introduction

Echolocation is the ability to perceive the environment using sound reflections. To achieve this, individuals often generate acoustic emissions and interpret the returning echoes to create a representation of their surroundings. This is a skill that has been described extensively in some non-human animal species, such as bats and dolphins (e.g. Jones, 2005; Schnitzler et al., 2003; Thomas et al., 2004), but by now it is well-established that humans can echolocate too (*reviews by* Kolarik et al, 2014; 2021; Thaler and Goodale, 2016). It has also been shown that humans can echolocate using artificially generated (i.e. not self-generated) emissions (e.g. Castillo-Serrano & Thaler, 2016; De Vos & Hornikx, 2018s; Steffens et al., 2022; Tirado et al., 2019, 2021) and by listening to binaural recordings of echolocation sounds (e.g. Dodsworth et al., 2020; Norman & Thaler, 2019; 2020a, 2020b; Schenkman & Nilsson, 2010; Wallmeier et al., 2013). Using relevant acoustic information from echoes human echolocators using mouth-clicks can infer object properties such as their distance, size, shape, material and position in azimuth (*reviews by* Kolarik et al, 2014; 2021; Thaler and Goodale, 2016).

In a first investigation into potential interfering effects of masking noise in human echolocation it has been shown that blind and sighted people can use echolocation to detect objects in noise (Castillo-Serrano et al., 2021). In this previous work we found that adjustments to the intensity of click emissions compensated for the potential interfering effect of broad-band masking noise when detecting sound-reflecting objects of various sizes and distances. It has also been reported that people adjust the intensity and number of emissions to detect relatively weaker echoes (Thaler et al., 2017, 2019, 2022). Such compensatory behaviour in human echolocation is perhaps not unexpected, considering adaptive behaviours observed in other echolocating species, for example bats (e.g. Amichai et al, 2015; Bates et al, 2008; Hage et al, 2013; Schnitzler et al, 2003;

Siemers and Schnitzler, 2004; Tressler and Smotherman, 2009; Luo et al, 2015a, 2015b). The current study builds on previous findings in human echolocation, in particular the finding that people can echolocate in the presence of masking noise via emission intensity adjustments (Castillo-Serrano et al., 2021), and explores the role played by the type of sonar emissions and interfering sounds, and the role played by binaural cues, i.e. spatial separation, of echoes and interfering sounds.

The extent of acoustic similarity between sounds of interest and interfering sound plays a role in signal masking (Bronkhorst, 2000; Brumm and Slabbekoorn, 2005). For example, research suggests that detection of sounds generally deteriorates in the presence of acoustically-similar interfering sounds (Kidd et al, 2002; Durlach et al, 2003). In the current study, we investigated the effect of acoustic similarity between emissions and maskers on echo perception by using two different types of sounds (i.e. clicks and broad band noise) as sonar emission and interfering sound. There are discussions within the field of acoustics as to what are best measures of acoustic similarity, with a main distinction being measures of similarity in temporal domain (i.e. signal envelope) vs. spectral domains (i.e. spectral frequency content). To be independent of this vast discussion, we chose our conditions so that similarity (or dissimilarity) applied in both temporal and spectral domains. This way, any effects we find would apply regardless of how similarity was measured. Our participants listened to binaural recordings of echolocation sounds (i.e. click and echo, or broad band noise and echo) in the presence of binaural presentations of masking sounds that could be either clicks or broad band noise. Thus, each emission was presented with masking sound that was either the same type of sound as used for the emission, or not. Based on previous research in human hearing of source sound we might expect that participants perform better (i.e. we expect lower signal to noise ratios, or SNRs) when masker and target sounds are acoustically less similar (e.g. click echolocation sound in the presence of a noise masker), as compared to when they are more similar (e.g. click echolocation sound in the presence of a click masker), i.e. we might expect an interaction effect between type of target and masking sound.

Binaural cues are important to represent auditory objects of interest in space. Spatial release from masking describes the advantage that spatial separation of signals and maskers via binaural cues offers for discrimination and localization of sounds (Litovsky, 2012). This has been reported for localization of tones in noise (Saberi et al, 1991; Good and Gilkey, 1996; Lorenzi et al, 1999) and for identification of sequences of tones (Kidd et al, 1998) and speech signals (Hawley et al, 1999). Additional work has observed greater signal interference as acoustically-similar signals and noise become spatially coincident (Freyman et al, 1999; Arbogast et al, 2002; 2005). In our work, we explore the benefit of binaural cues in human echolocation by comparing performance in a task where echoes were spatially separated from the masking noise via binaural cues (echo localization experiment), to performance in a task where they were co-located (echo audibility experiment). We might expect that participants perform better (i.e. we expect lower SNRs) when sounds are spatially separated via binaural signals, as compared to when they are spatially coincident via binaural signals, i.e. we might expect a main effect of binaural cues being available.

One may ask what is at stake here, since a lot of research has already looked at questions of sound type and binaural effects for target and masker in the context of human hearing of source sound. Yet, it is important to bear in mind that we are looking at echolocation. Thus, participants listened to sounds that contain a masker and a target, but the target contains both emission and echo. As a consequence the target itself is sort of ‘split in half’, with the echo carrying task relevant information. In fact, the emission just by itself has no information. Further, when considering spatial effects, any binaural separation would apply to the echo, but not the emission. This is why effects that have been observed in the context of human source hearing may or may not generalize to human echolocation, thus requiring separate investigation. Further to this, coming from a 'vision' perspective, for example, one may wonder if questions about similarity between masker and target or spatial release from masking address a trivial problem. Indeed, in the visual modality the problem is obvious and easy to solve. For example, whilst it is a challenge to detect one thing presented with the same thing (e.g. a red square on a red background), this task becomes easy as soon as things differ spectrally or spatially (e.g. a red square on a green background, or if the red target and background are in different locations). It is important to consider in this context that visually things in separate locations are separate on the sensor array, i.e. the retina, whilst acoustically they impinge on the same sensors (hair cells etc.), so that even spatially separate sounds are confluent in time and sensor space. This is the case also when sounds are composed of different spectral frequencies, and the brain has to work out how to separate sounds, because on the sensor array they appear simultaneously. Thus, the visual analogy of a 'green square on red background' or of 'two red things in different locations' breaks down for audition. As such, teasing apart masker from target in audition is not trivial and neither is spatial release from masking. These issues have been discussed elsewhere, but it is just to say that what may seem obvious from one sensory perspective (say vision) may be not be obvious from another (say audition).

It has been shown that performance in echolocation is better in people who are blind with long-term experience in click-based echolocation, as compared to people who are blind or sighted without experience in echolocation (e.g. Milne et al., 2014; Norman & Thaler, 2019, 2020a, 2020b; Thaler et al., 2020). Long term experience can also affect echo- perceptual judgments of size and weight in a way not observed in people without experience in echolocation (e.g. Buckingham et al., 2015; Milne et al., 2015). This suggests that expertise in echolocation plays a role for performance, rather than blindness per se. Alternatively, it has also been suggested that people who are blind are more sensitive to acoustic reverberation as compared to people who are sighted (Dufour et al., 2005; Kolarik et al., 2013), and that this may put people who are blind at a particular advantage for learning echolocation (Kolarik et al., 2016, 2021). Yet, it has been shown that people who are sighted can learn echolocation just as well as people who are blind, and perform at levels matching or approaching performance levels of blind echolocation experts (e.g. Norman et al., 2021; Teng et al., 2012). Current evidence does not suggest an advantage for people who are blind either in rate of learning or final skill levels. Interestingly, previous work about the effects of masking sound on echolocation we did not find evidence to support the idea that performance was affected by experience in echolocation, or blindness (Castillo-Serrano et al., 2021). Thus, to address the potential roles played by long-term experience with echolocation and blindness three different participant groups took part in the current study, specifically, people who were sighted or blind and who were new to click-based echolocation and people who were blind and who had long-term experience in click-based echolocation. If long-term experience with echolocation or blindness play a role for performance in our paradigm, we would expect either experts and/or people who are blind to perform better (i.e. have lower SNRs) than other participants across conditions, i.e. we might expect a main effect of group.

2 Methods

All procedures were approved by the Ethics committee at Durham University Psychology Department (REF 16/19) and were carried out in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki) and the British Psychological Society. The participant letter of information and consent form were provided in accessible format to all participants with vision impairments. All participants gave written informed consent prior testing. Sighted participants were compensated with £6 per hour or participant pool credits. Visually impaired participants received £10 per hour, with the higher compensation compensating for more complex logistics to attend testing sessions.

2.1. Echo Localization

This Experiment explored people's ability to use echolocation in the presence of a masking sound to localize a 50-cm diameter disk placed at +/-20° in azimuth relative to the echolocators' straight ahead orientation, i.e. either 20° to the left or to the right. Thus, the target echolocation sound had binaural echo cues. The masking sound was a mono-aural sound, i.e. the same sound was presented to right and left ears. This was done to distinguish the masker from the target via binaural cues.

2.1.1.Participants

Three different groups of participants took part: blind expert echolocators (BEs), blind controls (BCs), and sighted controls (SCs). All participants had normal hearing levels appropriate for their age group (ISO 7029:2017) as assessed with pure tone audiometry (250-8000Hz; Hughson Westlake; Interacoustics AD629 audiometer, Interacoustics, Denmark) and no history of neurological disease. BEs had long-term experience using click-based echolocation on a daily basis, whereas BCs and SCs indicated either no previous experience with echolocation or no regular use of echolocation in order to meet the criteria to be considered experts. Table 1 lists details of blind participants. Not all participants took part in both experiments, thus table 1 also lists which person took part in which experiment. In the localization experiment, three blind expert echolocators (Participant identification: BE1-BE3 in Table 1. Mean age: 37.3 years, SD:

14.04; all male), six blind participants with no experience in echolocation (Participant identification: BC1-BC6 in Table 1. Mean age: 48.8 years, SD: 16.98; one female.), and 20 sighted people with no experience in echolocation with normal or corrected to normal vision (SC, mean age: 35.9 years, SD: 15.6; min: 18; max: 70; median: 31; twelve female), participated.

Table 1 *- Details of blind expert echolocators and blind control participants who took part in the study. BE – Blind Echolocation Expert; BC – Blind Control Participant. Blind M -Male; F – Female*

ID	age in years at time of testing	Gender	Degree of vision impairment at time of testing	cause and age at onset of vision loss	History of use of click-based echolocation	Echo Localization	Echo Audibility
BE1	36	M	Total blindness	Severe childhood glaucoma	Daily; since age 12	X	
BE2	24	M	Total blindness	Unknown cause; vision loss at age 12; eyes removed at age 19	Daily since age 12	X	
BE3	52	M	Total blindness	Retinoblastoma; onset at birth; enucleation at age 13 months	Daily; since early childhood/no exact age remembered	X	X
BE4	43	F	Total blindness	Leber's congenital amaurosis; birth	Daily; since age 31 years		X
BE5	59	M	Total blindness	Retinal detachment; birth	Daily; since age 6 years		X

BC1	46	M	bright light detection	ocular albinism; birth	None	X	X
BC2	54	M	Bright light detection	Retinitis Pigmentosa; birth with increasing severity	None	X	X
BC3	18	M	Right eye total blindness (enucleation); left eye 20/200 acuity	Retinopathy of prematurity; accident at 11 years (leading to enucleation of right eye)	No regular use	X	X
BC4	60	F	Total blindness left eye; some peripheral vision in the right eye	Stichler’s syndrome; retinal sciasis, from birth with increasing severity	None	X	X
BC5	67	M	Bright light perception	Leber's Amaurosis; from birth	None	X	
BC6	48	M	Total blindness in left eye; residual bright light perception in the right eye	Severe childhood glaucoma; 3 months old	None	X	X
BC7	36	F	Bright light detection	Unknown cause; from birth	None		X

2.1.2. Experimental sounds

Echolocation emissions

Separate click and noise sound files were digitally created at a sampling rate of 96 kHz and 24-bit resolution using Matlab R2015b (The Mathworks, Natick, MA). To build the click, a 4.5-kHz tone of 10-ms duration was generated and then all values were multiplied up until the first half period by 0.6; these characteristics simulate the rising intensity of a natural click. Then, all values after the first 1.5 periods were multiplied by the output of the decaying exponential function $y=e^{-6x}$, where x is a series of linear equally spaced values between 0 and 1 that is equal in length to the number of values in the sinusoid between the first 1.5 periods and its end; this is comparable to the fall in intensity of a natural click. This artificial click approximates the waveform of an actual mouth click produced by human expert echolocators (de Vos & Hornikx, 2017; Martinez-Rojas et al, 2009; Thaler et al, 2017) and it has been used in other echolocation studies (e.g. Thaler & Castillo-Serrano, 2016; Norman and Thaler, 2018). The noise emission was a 500-ms broadband noise with energy between 0.2 and 20 kHz. This type of emission has been used successfully in previous investigations about human echolocation (e.g. Schenkmann & Nilsson, 2010)

Masking Sounds

Two types of masking sound were used in this study: Broad band noise and clicks (without echo). The Broadband noise masker was a 60-s broadband noise with energy between 0.2 and 20 kHz. The click masker was a 60-s train of click samples identical to the click emission used in this study. The click train contained clicks at 10 Hz, with random jitter applied to the onset of each click (drawn randomly and uniformly between -0.025 and +0.025 seconds).

Sound recording equipment and setup

Recordings of all sounds used in this study were obtained in a sound-insulated and echo-acoustic dampened room (approx. 2.9 m x 4.2 m x 4.9m; 24 dBA noise floor) lined with foam wedges (315 Hz cut-off frequency). Binaural sound recordings were produced at a 96-kHz sampling rate and 24-bit resolution using in-ear microphones (Bruel & Kjaer model 4101, Denmark) that were attached to a portable digital recorder (Tascam DR-100 MK2, TEAC Corporation, Japan). The microphones were placed in the ears of a custom-made manikin, consisting of a head and torso. Five-millimetre diameter holes were drilled inside the manikin's ears to act as artificial ear canals and these made possible to insert and keep the in-ear microphones steady. See Norman & Thaler (2018) for anthropometric details of this manikin. Sound recordings across all test conditions were made using a constant level of amplification in all electronic equipment. The echolocation emissions were played individually through a loudspeaker (Fostex FE103En) that was fixed to the mouth of the manikin. Masking sounds were played individually from the same loudspeaker but mounted on a metal pole standing 100 cm away from the left ear of the manikin. The loudspeaker was controlled using a Dell Latitude E7470 laptop (Intel Core i56300U CPU 2.40 GHz, 8 GB RAM, 64-bit Windows 7 Enterprise) via a USB Sound Card (Creative Sound Blaster X-Fi HD Sound Card; Creative Technology Ltd., Creative Labs Ireland, Dublin, Ireland) and amplifier (Kramer 900N; Kramer Electronics Ltd., Jerusalem, Israel).

We obtained individual sound recordings of the click emission with echo, and the broadband noise with echo in the presence of a 0.8-mm thick disk (50 cm diameter) made from plywood and covered with matte emulsion paint and placed at 1 m from the manikin displaced in azimuth by either -20° (i.e., to the left) or +20° (i.e., to the right). The flat side of the disk was angled towards the manikin at either location, to facilitate performance (Rowan et al., 2017). Figure 1 (a) and (b) present illustrations of the recording setup for each echolocation sound.

Individual recordings of each of the masking sounds were obtained by playing each sound from the loudspeaker into the left ear of the manikin at a distance of 1 m. For these recordings the room was empty, i.e. no object was presented. The recording set up for maskers is illustrated in Figure 1 (c). Note that whilst microphones in the ears of the manikin made separate recordings for the right and left channel for each masking sound, during the experiment only the sound recorded on the left channel was presented to both ears of each listener. This was done to remove binaural cues from the masking sounds.

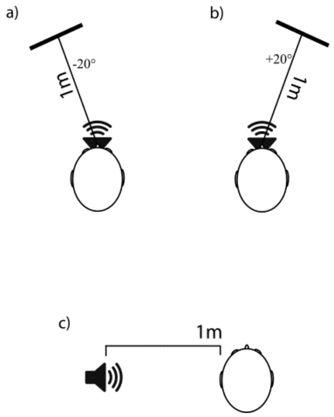

Figure 1 – Top view illustrations of setup used to generate sound recordings for the experiment. In all cases sounds were recorded by in-ear microphones placed inside the manikins head's ears. **(a)** Illustrates the recording setup for echolocation sounds with the sound reflecting 50-cm diameter disk facing the manikin from 1m distance at an azimuth angle -20° (i.e. to the left) from straight ahead. The manikin faced front and the sound emitting loudspeaker was fixed to the manikin's mouth. **(b)** Same as in (a), but the sound reflecting 50-cm diameter disk was facing the manikin at an azimuth angle +20° (i.e. to the right) from straight ahead. **(c)** Illustrates the recording setup for masking sounds. The

manikin faced front and the sound emitting loudspeaker stood facing the left ear of the manikin at 1-m distance. Sound reflecting disks were absent during recordings of masking sounds (i.e. the room was empty).

Sound stimuli

For echolocation stimuli two conditions were used for each emission: one corresponded to the sound recording made when the 50-cm disk stood at 1 m -20° in azimuth (i.e. to the left; the reference sound), and the other corresponded to the sound recording obtained when the same disk was placed at +20° in azimuth (i.e. to the right; the comparison sound). Figure 2 presents the waveforms illustrations of these sounds.

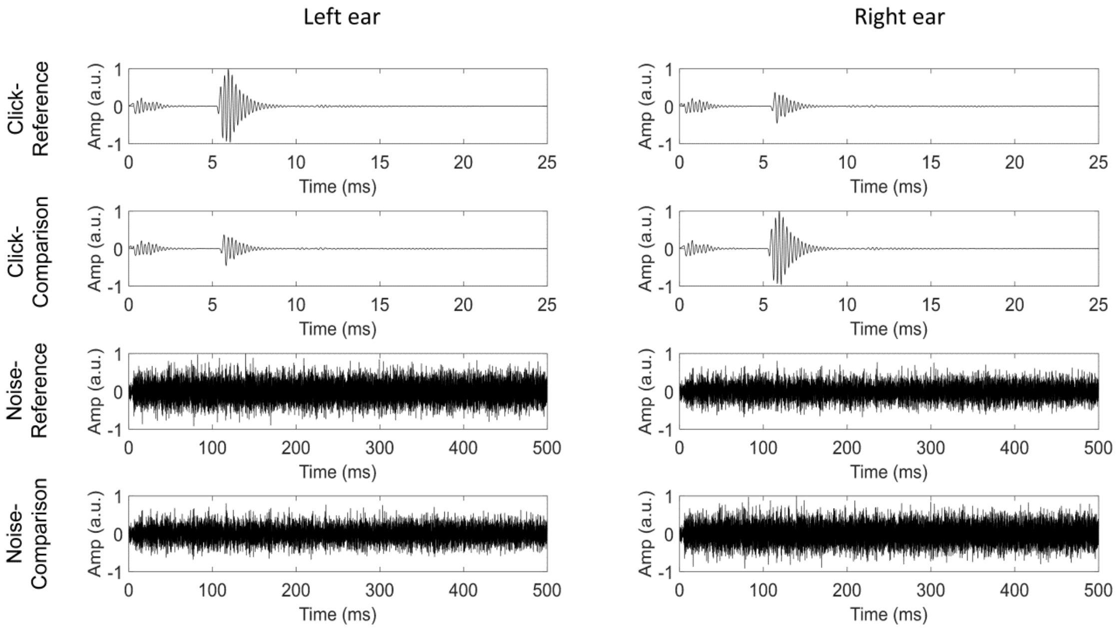

Figure 2 – Waveform plots of echolocation sounds presented in the Echo Localization Experiment. From top to bottom: illustrations of binaural recordings of click sounds recorded when the target disk was placed -20° in azimuth (i.e. to the left of the manikin; reference sound) (row 1), and when the target disk was placed +20° in azimuth (i.e. to the right of

the manikin; comparison sound) (row 2); binaural recordings of noise burst when the target disk was placed -20° in azimuth (reference sound) (row 3), and when it was placed +20° in azimuth (comparison sound) (row 4). The emission and echo are temporally separated in the click recordings, and they overlap temporally in the longer-duration noise burst recordings. The abbreviation a.u. refers to “arbitrary units.” In click recordings it is particularly evident that echoes are of higher intensity than emissions. This is because we made our recordings using in-ear microphones placed behind the loudspeaker, leading to a lower intensity of emissions measured at the ear.

For masker stimuli, two different conditions were used: Broadband noise and clicks. The clicks masker on each trial was a 1-s randomly chosen sample from the recording of the train of clicks, with clicks the same as used for the click emission. The noise masker on each trial was a 1-s sample randomly chosen from the recording of broadband noise. Figure 3 presents waveform illustrations of these sounds.

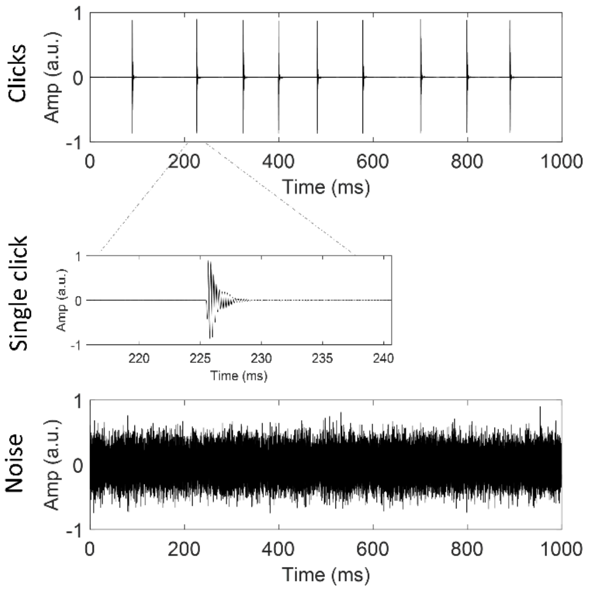

Figure 3 – Waveform plots of masking sounds presented in the Echo Localization Experiment. In all conditions the same sound was presented to both ears, hence only one channel is plotted for each sound. The abbreviation a.u. refers to "arbitrary units."

Figure 4 shows power spectra (1/3 Octave Bands) for the experimental sounds used in the Echo Localization experiment. It is evident that power spectra for clicks masker and click echolocation sounds are very similar, and the same for noise masker and noise echolocation sounds. Due to the nature of the clicks and click-masker (10Hz) the masker rarely overlaps the clicks and echoes in time, though.

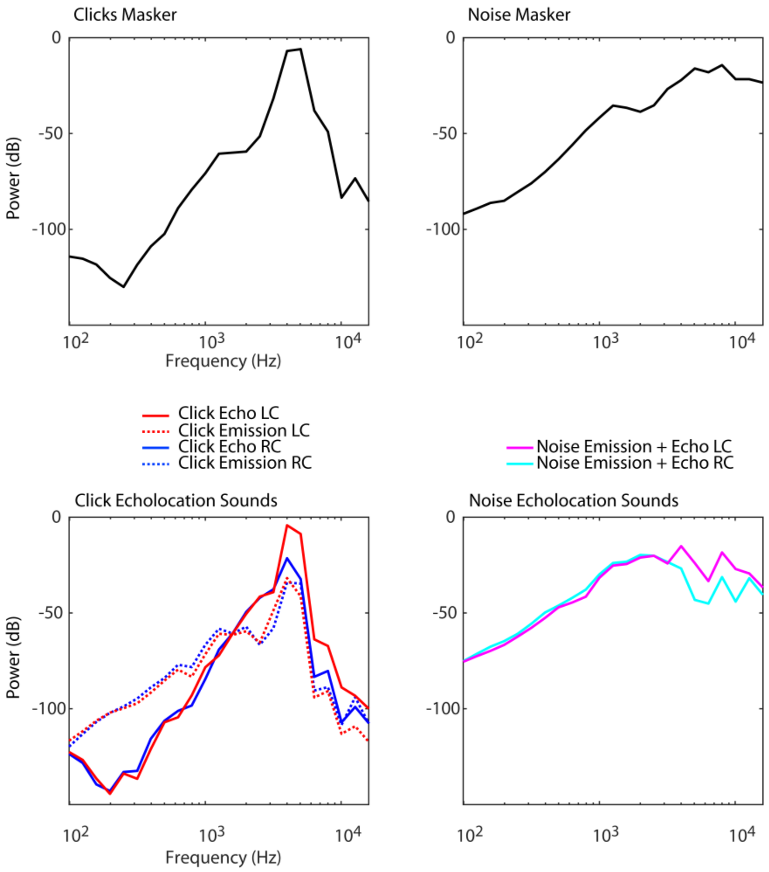

Figure 4 – Power spectra (1/3 octave bands with respect to total power) for sounds used in the Echo Localization Experiment. Top panels show masking sounds, bottom panels echolocation sounds. Different line colours and styles denote spectra for the various components, i.e. masker, emission, echoes. LC – left channel. RC – Right Channel. Spectra shown are for the reference sound (sound reflecting object placed on left side), but they are equivalent for the comparison sound, except that RC and LC are reversed.

2.1.3. Set up and apparatus for behavioural task

All experimental sessions were performed in a sound-insulated and echo-acoustic dampened test room (approx. 3 m x 2.5 m x 3.3 m) in Durham University Psychology department. The experimental sounds were played from a PC (Intel Core i56600 CPU 3.30 GHz, 16 GB RAM, 64-bit operating system, x64-based processor, Windows 10 Home) and participants listened to the experimental sound through in-ear headphones (Etymotic Research ER4B MicroPro) that were connected to a USB (Sound Card (Creative Sound Blaster X-Fi HD Sound Card; Creative Technology Ltd., Creative Labs Ireland, Dublin, Ireland) attached to the PC. Participants sat on a chair in the test room and performed a computer-based experiment; they used a keyboard to enter their responses. All sighted and other participants who had residual vision wore a blindfold during the experiments. The experiments were programmed in Matlab R2015b (The Mathworks, Natick, MA) and Psychtoolbox (v3.0.12; Brainard, 1997). Sounds were played to participants at a level at which the sound file with the highest peak intensity was presented at 80 dB SPL.

2.1.4. Procedure for behavioural task

Participants' task was a 2-interval forced choice task, in which they listened to two sounds in succession, separated by 500 ms of silence, and identified which of the two sounds (first or second) contained the echo from the object presented on the left side, i.e. the reference sound. Sound presentation order was random on each trial. Participants entered their responses using a computer keyboard. They pressed the 'z' and 'm' keys to indicate that their judgement corresponded to the first or the second sound, respectively. Participants completed training and test sessions for all conditions. During training, participants were made familiar with the tasks with no masking sound presented. In the test sessions, participants completed the tasks in the presence of masking sound. Here, an adaptive staircase procedure adjusted the intensity of emissions, relative to the intensity of masking sound. Specifically, emissions' dB SNR increased or decreased based on participants' ability to respond correctly.

It typically took participants 3 hours to complete all training and testing sessions. Breaks were provided to all participants in between experimental conditions in order to prevent fatigue, and participants had the option to complete their participation on separate days. As far as possible, presentation of echolocation sounds was counterbalanced across participants. Half of the participants in each group used click emissions first, and then noise emissions, and the order of masking sounds was chosen at random. The other half of participants in each group used noise first, and then used click emissions.

Training procedure

Participants learned the task during the training sessions. In the first part of the training, participants completed blocks of 40 trials in which they heard feedback for each trial and they trained until they reached an accuracy level of at least 90% correct responses when feedback was presented. A high pitch tone (i.e. 1200 Hz) indicated that they gave correct responses, and they heard a lower pitch tone (i.e. 600 Hz) for incorrect responses. Once participants performed the task with 90% accuracy when they heard feedback, they proceeded to complete blocks of 40 trials without feedback. Participants were expected to give at least 90% correct responses when no feedback was presented before they were allowed to perform the test sessions.

Testing procedure

Participants completed four test conditions that included masking sounds. These conditions corresponded to the combination of emissions (clicks and noise) and maskers (clicks, and noise). On each trial, each echolocation sound was presented in the middle of a 1-s masking segment (randomly chosen for each trial). This segment also included an additional 250-ms linear ramped onset (from zero to the desired sound level for that trial). So, the sequence of each test trial was:

250 ms linearly ramped masker, 1000 ms masker (including echolocation sound), 500 ms silence, 250 ms linearly ramped masker, 1000 ms masker (including echolocation sound).

On each trial, the levels of the emission and masker were determined using a 2-up-1-down adaptive staircase procedure, in which the signal-to-noise ratio (SNR) varied based on participants' accuracy. Specifically, the SNR was defined as the ratio (in dB) of the emission, without any echo present, relative to masker. For SNR values below 0, the intensity of the masker remained constant, whilst the intensity of the emission decreased after two consecutive correct responses and increased after one incorrect response. For SNR values above 0, the intensity of the emission remained constant, whilst the intensity of the masker increased after two consecutive correct responses and decreased after one incorrect response. The magnitude of the intensity increment/decrement was 6 dB until 6 staircase reversals had been made, after which it was 2 dB. Four interleaved adaptive staircases were included in the test sessions and each one was assigned a different starting SNR value (-20, -10, 0 and +10). Staircases continued to be presented as long as the SNR values were within its limits (i.e., -70 and +40 dB SNR). Each staircase terminated after 14 direction reversals occurred (i.e., from correct to incorrect, or vice versa). Feedback was not provided during the test sessions.

2.1.5. Data analysis

Psychometric curves describing proportion correct as a function of SNR were fitted to data for each condition in each experimental task. Matlab R2015b (The Mathworks, Natick, MA) and the Palamedes toolbox (Prins & Kingdom, 2018) were used to fit psychometric functions (cumulative normal, with threshold and slope as free parameters) with a maximum likelihood criterion to describe proportion correct as a function of signal-to-noise ratio. The point on the function at which proportion correct was 0.75 was taken as threshold i.e. the SNR at which people are expected to obtain 75% correct responses. Further statistical analyses on the group level were conducted with SPSS v26. Specifically, SNR results on each condition were analysed in a mixed model ANOVA with within-subjects factors of ‘emission’ (2) and ‘masking sound’ (2). ‘Group’ (SC, BC and BE) was the between-subject factor. Threshold for significance was set at .05 and Bonferroni correction was applied for multiple comparisons, and corrected thresholds are reported as appropriate in the text.

2.2. Echo Audibility

In this experiment , we wanted to test the effects of acoustic similarity between emission and masker in the absence of binaural cues, so that by direct comparison to the localization experiment we could also assess the role played by binaural cues for echolocation in the presence of a masking sound. Thus, we designed a paradigm where echolocation sounds and maskers were yoked to those used in the localization experiment, but that did not contain any binaural cues. Like the echo localization task, the echo audibility task was also a 2-interval forced choice task. In the echo audibility task participants judged which of two echolocation sounds was more audible, i.e. which one they could hear better, in the presence of masking sound. Just as in the localization experiment, participants first trained the task with feedback, before masking noise was introduced. More details are described below.

2.2.1. Participants

Three blind expert echolocators (Participant identification: BE3-BE5 in Table 1. Mean age: 51.3 years, SD: 8.02; one female), six blind participants with no experience in echolocation (Participant identification: BC1, BC2, BC3, BC4, BC6, BC7 in Table 1. Mean age: 43.67 years, SD: 14.94; two female), and 20 sighted people with no experience in echolocation with normal or corrected to normal vision (SC, mean age: 36.8 years, SD: 16.9; min: 19; max: 71; median: 36; 15 female), participated in this experiment. All participants had normal hearing levels appropriate for their age group (ISO 7029:2017) as assessed with pure tone audiometry (250-8000Hz; Hughson Westlake; Interacoustics AD629 audiometer, Interacoustics, Denmark) and no history of neurological disease.

2.2.2. Experimental Sounds

For masking sounds the exact same sounds used in the echo localization experiment were used. Echolocation sounds were based on those used in the echo localization experiment but instead of using binaural recordings, we used right and left channels separately to create two mono-aural sounds for each echo emission. Figure 5 presents the waveforms illustrations of these sounds. Spectral properties of sounds used are like those for echo localization (compare Figure 4), with the only difference that left and right channel sounds (from the echo localization experiment) correspond to reference and comparison sounds (in the echo audibility experiment), respectively. Thus, both in the temporal and the spectral domain stimuli were yoked to those in the echo localization experiment in terms of their temporal and spectral similarity, but they did not contain binaural cues. Just like for echo localization, due to the nature of the clicks and click-masker (10Hz) the clicks masker rarely overlapped the clicks and echoes in time.

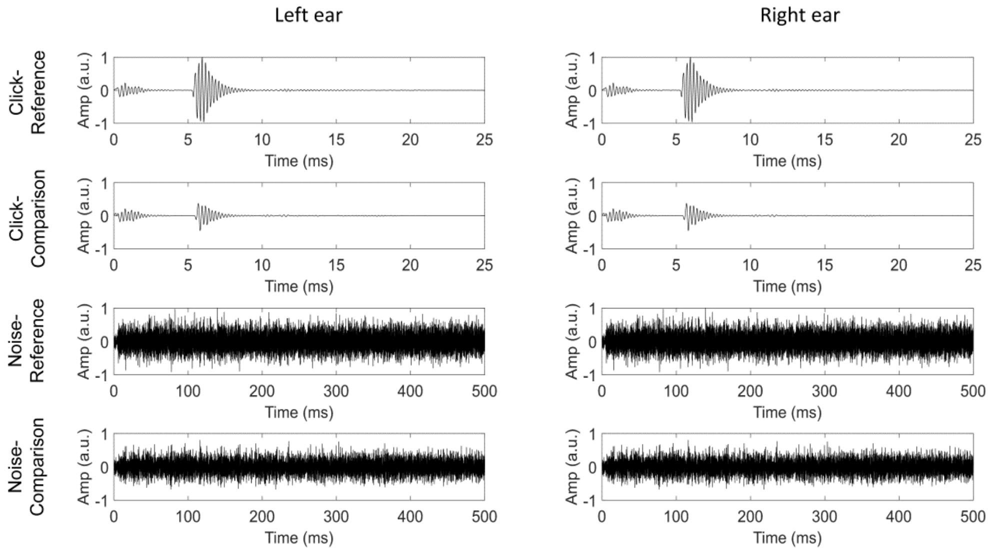

Figure 5 – Waveform plots of echolocation sounds used in the Echo Audibility Experiment. Left and right panels illustrate sound played to the left and right ear respectively (which were identical in this experiment), and the different conditions are shown in different rows. From top to bottom: illustrations of click reference sounds (row 1) and click comparison sound (row 2); illustration of noise reference sound (row 3), and noise comparison sound (row 4). The emission and echo are temporally separated in the click recording, while they overlap temporally in the noise recordings. The abbreviation a.u. refers to “arbitrary units”.

2.2.3. Set up and apparatus for behavioural task

The same set-up and apparatus as used for the echo localization experiment was also used for the echo audibility experiment.

2.2.4. Procedure for Behavioural task

Participants’ task was to listen to two sounds in succession, separated by 500 ms of silence, and to state which of the two sounds (first or second) contained the echo that they could hear better (i.e. the target sound). Apart from this everything was the same as for the localization experiment. The reason that we asked people to judge how well they could hear echoes was that both sounds always contained an echo. Thus, we did not feel it was appropriate to instruct people to ‘detect’ an echo (in particular at higher SNR values this instruction would have been confusing). Importantly, the exact same task was chosen for all conditions of the experiment. Participants reported that this task felt intuitive to them, and the training and testing data show that they performed well (see Results).

Training and testing procedure

The same training and testing procedures as used for the echo localization experiment were used, with the only difference that participants task was to determine which interval contained the echo sound that they could hear better. Just as for echo localization, masking sounds were only presented during testing.

2.2.5. Data analysis

Data were analysed in the same way as for the echo localization experiment.

3. Results

Data availability statement

Data are available as Supplemental Material S1.

3.1. Echo Localization

Training sessions

All expert echolocators were 100% accurate after a single training session with and without feedback for both echolocation emissions. Participants new to echolocation reached the 90% correct response criterion after an average of 1.5 training blocks with feedback (SD: 1.04) and after 1.12 training blocks without feedback (SD: 0.32). For sessions with feedback, using ANOVA with emission type (click vs. noise) as repeated variable and group (blind vs. sighted) as between subjects factor, there was no significant difference between click and noise emissions (F(1,24)=.295; p=.592; η_p^2=.012), or between blind and sighted groups (F(1,24)=.779; p=.386; η_p^2= .031) in terms of the numbers of training sessions, and there was also no significant interaction (F(1,24)=1.016; p=.323; η_p^2 =.041). The same analysis applied to sessions without feedback also revealed no significant effects (emission: F(1,24)=.399;p=.534; η_p^2=.016; group: F(1,24)=.117;p=.735; η_p^2= .005; emission x group: F(1,24)=1.374; p=.253; η_p^2=.054). Thus, our data suggest that blind and sighted participants did not differ in the amount of training required for both click and noise emissions.

Test sessions

Figure 6 (a) presents threshold signal-to-noise ratios for each group and test condition. It is evident that SNRs were similar across groups, but differed across conditions. Specifically, in contrast to what one may have expected based on known effects of acoustic similarity in source hearing where masking effects are driven by acoustic similarity, in the echo localization experiment SNRs were consistently highest for the noise masker, and lowest for the click masker, regardless of which emission type was used. Thus acoustic similarity does not appear to play a role.

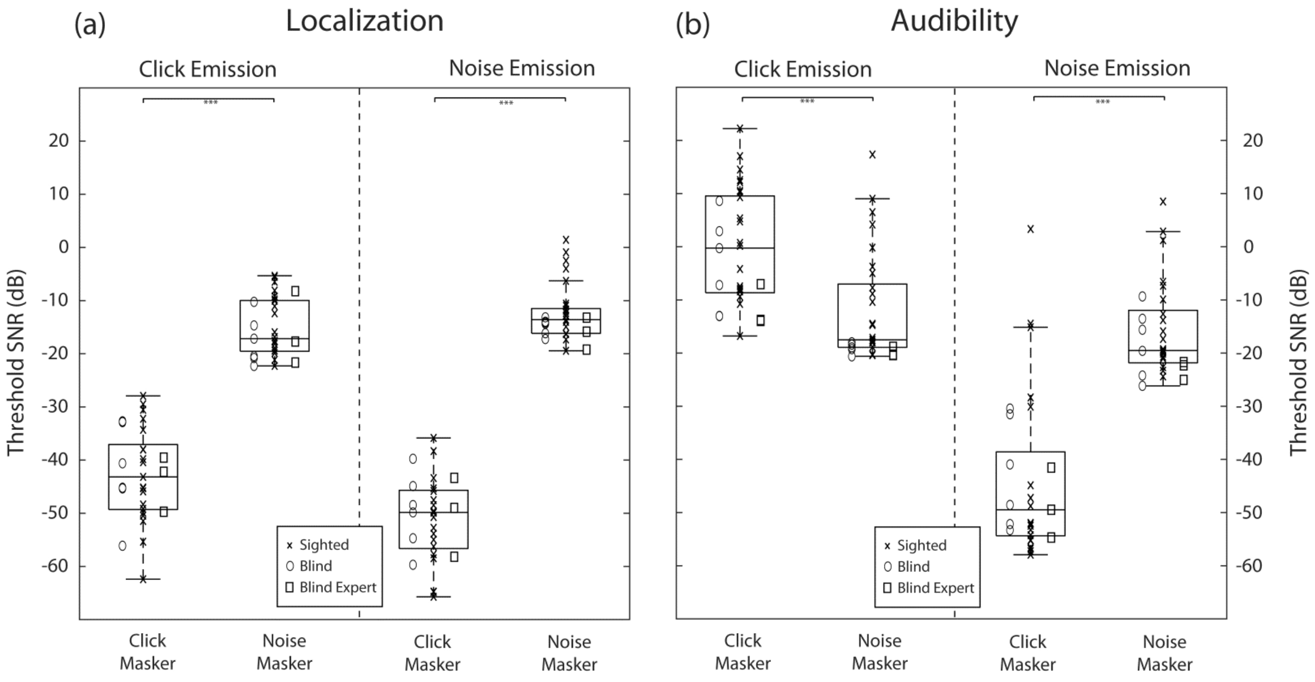

Figure 6 – SNRs at threshold (75%) for the (a) Echo Localization Experiment and (b) Echo Audibility Experiment. SNRs for intensity of click emissions and noise emissions presented along with each masking sound, plotted separately for the three participant groups . Box and Whisker plots with horizontal bars and lower/upper box boundaries representing median and 25th/75th percentile, respectively. Whiskers extend to 1.5 IQR, drawn back to the closest data point within that range. Symbols denote data from individual participants and are broken down into the different participant groups by shape. Asterisks indicate results of post-hoc tests (Bonferroni corrected) *** p<.001. For details see main text.

An ANOVA with emission and masker as repeated variables and group as between subjects variable showed significant effects of ‘emission’ ($F(1, 26)=5.130$; $p<.032$, $\eta_p^2 = .165$), where click emissions had generally higher SNRs (mean: -28.96; SD: 15.69) than noise emissions (mean: -

31.77; SD: 20.56), and ‘masker’ ($F(1, 26)=134.215$; $p<.001$, $\eta_p^2 = .838$), where the click masker had generally lower SNRs (Mean: -46.95; SD: 9.1) than the noise masker (mean: -13.78; SD: 5.5). There was also a significant ‘emission’ and ‘masker’ interaction ($F(1,26)=21.181$; $p<.001$, $\eta_p^2 = .449$). The main effect of group was not significant ($F(2,26)=.556$; $p=.580$; $\eta_p^2= .041$), neither was the interaction between emission, masker and group ($F(2,26)=.671$;$p=.520$; $\eta_p^2= .049$), or the interaction between emission and group ($F(2,26)=.042$;$p=.958$; $\eta_p^2= .003$), or the interaction between masker and group ($F(2,26)=.521$;$p=.600$; $\eta_p^2= .039$). It can also be seen in Figure 6 (a), that all three participant groups showed the same pattern of results. Thus, for subsequent analyses to follow up the interaction between emission and masker we conducted paired t-tests across masking conditions for each emission separately (Bonferroni corrected threshold for significance was .025). For the click emission, SNRs for the click masker (mean:-42.81; SD: 8.59) were significantly lower than for the noise masker (mean: -15.11; SD: 5.51; $t(28)=12.467$; $p<.001$). The same pattern of results was observed for the noise emission, where SNRs for the click masker (mean: -51.09; SD: 7.69) were significantly lower than for the noise masker (mean: -12.44; SD: 5.25; $t(28)=18.437$;$p<.001$). Thus, our results suggest that when target and masker were separated by binaural cues, acoustic similarity did not seem to play a role, i.e. SNRs for both click and noise emissions were lower for click masker as compared to the noise masker, making the noise masker the more efficient masker regardless of emission type.

3.2. Echo Audibility

Training sessions

All expert echolocators were 100% accurate after a single training session with and without feedback for both echolocation emissions. Participants new to echolocation reached the 90%

correct response criterion after an average of 2.13 training blocks with feedback (SD: 2.28) and after 1.21 training blocks without feedback (SD: 0.57). For sessions with feedback, using ANOVA with emission type (click vs. noise) as repeated variable and group (blind vs. sighted) as between subjects factor, people needed significantly fewer training sessions for noise emissions (mean:1, SD:0) than for click emissions (mean: 3.27, SD: 2.82) (F(1,24)=12.315; p=.002; η_p^2=.339), but there was no difference between blind and sighted groups (F(1,24)=.050; p=.825; η_p^2= .002) in terms of the numbers of training sessions, and there was also no significant interaction (F(1,24)=.050; p=.825; η_p^2 =.002). For sessions without feedback, people also needed significantly fewer training sessions for noise emissions (mean:1, SD:0) than for click emissions (mean: 1.42, SD: 0.76) (F(1,24)=6.274; p=.019; η_p^2=.207), but there was no difference between blind and sighted groups (F(1,24)=.077; p=.783; η_p^2 =.003) in terms of the numbers of training sessions, and there was also no significant interaction (F(1,24)=.077; p=.783; η_p^2 =.003). Thus, our data suggest that noise emissions were learned more quickly, but that blind and sighted participants did not differ in the amount of training required.

Test sessions

Threshold signal-to-noise ratios for the different groups and test conditions are shown in Figure 6 (b). SNRs are similar across groups, but differ across conditions. Specifically, SNRs were lowest when target and masker were different (i.e. click in noise, or noise in click), but higher when they were the same (i.e. clicks in clicks, noise in noise). Thus, even though for the click masker, click and echo rarely overlapped the masking clicks, the clicks masker was the more efficient masker. This is what one may have expected based on known effects of acoustic similarity in source hearing where masking effects are driven by acoustic similarity.

Consistent with this observation, an ANOVA with emission and masker as repeated variables and group as between subjects variable showed a significant ‘emission’ and ‘masking sound’ interaction (F(1,26)=74.726; p<.001, η_p^2 = .742). This was accompanied by a significant effect of ‘emission’ (F(1, 26)=85.341; p<.001, η_p^2 = .766), where noise emissions had generally lower SNRs (mean: -28.95; SD: 18.81) than click emissions (mean: -6.14; SD: 12.05), and a significant main effect of ‘masker’ (F(1, 26)=18.935; p<.001, η_p^2 = .421), where the clicks masker had lower SNRs (mean: -22.22; SD: 25.65), than the noise masker (mean: -13.88; SD: 9.71). The main effect of group was not significant (F(1,26)=2.104; p=.142; η_p^2= .139), neither was the interaction between emission, masker and group (F(2,26)=.194;p=.825; η_p^2= .015), or the interaction between emission and group (F(2,26)=1.404;p=.264; η_p^2= .097), or the interaction between masker and group (F(2,26)=1.028;p=.372; η_p^2= .073). It can also be seen in Figure 6 (b), that all three participant groups showed the same pattern of results. Thus, to follow up the interaction between ‘emission’ and ‘masker’ we combined data cross groups and conducted paired t-tests across masking conditions for each emission separately (Bonferroni corrected threshold for significance was .025). In line with data shown in Figure 6 (b), for the click emission, SNRs for the noise masker were lower (mean: -11.93; SD: 10.38) than for the click masker (mean: -.35; SD: 10.88; t(28)=5.225; p<.001). The reverse pattern of results was observed for the noise emission, where SNRs for the click masker (mean: -44.08; SD: 15.16) were lower than for the noise masker (mean: -15.82; SD: 8.75; t(28)=14.685;p<.001).

In sum, our results suggest that in the echo audibility experiment , where the target and the masker have no difference in terms of binaural cues, acoustic similarity between signal and masker plays an important role, i.e. performance is worst (and threshold SNRs are highest) for both click and noise emissions for the masker most similar to the emission.

3.3. Direct Comparison between Audibility and Localization– Binaural Release from Masking

To directly assess the role played by binaural release from masking, we compared performance in all conditions across the echo audibility experiment (without binaural difference between target and masker) and the echo localization experiment (with binaural difference between target and masker). In our study, different sighted participants had performed in each of the experiments. Yet, for BCs and BEs, some had performed both experiments (compare Table 1). Thus, we split data for those participants who had done both experiments, so that we pseudo-randomly assigned 6 participants to each experiment. Thus, data from BE1, BE2, BE3, BC4, BC5 and BC6 were analyzed for the echo localization experiment, and data from BE4, BE5, BC1, BC2, BC3 and BC7 were analyzed for the echo audibility experiment. The whole data set was then analysed using mixed ANOVA (with emission and masking sound as repeated variables, and binaural cue as between subjects factor). Figure 6 (a) and (b) show data from both experiments, and it seems that there is binaural release from masking (i.e. better performance when binaural cues are available in the localization experiment compared to when they are not available in the audibility experiment), but more so for click emissions in click maskers, than for any of the other conditions. Consistent with this observation, the ANOVA revealed a main effect of binaural cue ($F(1, 50)=41.357$, $p<.001$, $\eta_p^2 = .453$), i.e. people had overall lower SNRs when binaural cues were available (mean: -30.03; SD: 18.3) as compared to when they were not available (mean: -17.92; SD: 20.37), but this was moderated by a significant interaction between emission, masker and binaural cue ($F(1, 50)=53.989$, $p<.001$, $\eta_p^2 = .519$). Follow-up analyses with independent samples t-tests (df corrected for unequal variances as appropriate; Bonferroni corrected threshold of significance was .0125) across the two experiments for each click and masking sound combination showed that there was a significant binaural advantage of 42 dB for click emissions in click masker (t(47.007)= 15.22; p<.001; mean difference: -42.15; SE of difference: 2.77), but none of the other comparisons were significant (click emission in noise masker: t(37.212)= 1.570; p=.125; mean difference: -3.69; SE of difference: 2.35; noise emissions in click masker (t(36.98)= 1.803; p=.079; mean difference:-6.16; SE of difference: 3.42; noise emission in noise masker: t(40.695)=1.692; p=.098; mean difference: 3.54; SE of difference: 2.09). As expected from previous analyses for echo audibility and echo localization experiments, in the overall ANOVA there were also significant effects for emission (F(1, 50)=170.897, p<.001, $\eta_p^2 = .774$) and masking sound type (F (1, 50)=265.694, p<.001, $\eta_p^2 = .842$), as well as the interaction between the two factors (F(1, 50)=176.306, p<.001, $\eta_p^2 = .779$), as well as significant interactions between emission and binaural cue (F(1, 50)=102.675, p<.001, $\eta_p^2 = .673$) and between masker and binaural cue (F(1, 50)=86.9, p<.001, $\eta_p^2 = .635$). Since we have analyzed effects of emission and masking sound separately for each experiment earlier in previous sections we will not follow these up further. In sum, there was an advantage for performance when binaural cues distinguished the echo from the masker, i.e. binaural release from masking, but only when clicks were used as emissions and masking sounds.

4. Discussion

In our study blind and sighted people localized echoes in azimuth and discriminated audibility of echoes in the presence of interfering sound across two separate tasks. Previous research had shown that people can echolocate in the presence of masking sound, and that this is facilitated by increased emission intensity (Castillo-Serrano et al., 2021). The present work extends our previous findings on human echolocation ability by considering the effects of type of emission and masking sound, and their binaural There are discussions within the field what are best measures of acoustic similarity, with a main distinction being measures of similarity in temporal (i.e. signal envelope) vs. spectral domains. Importantly, we had chosen our conditions so that similarity (or dissimilarity) applied in both temporal and spectral domains, so that our results apply regardless of which measure of similarity would be chosen. Future research may possibly address the issue of what is a best measure of similarity in the context of human echolocation.

Importantly, sounds used in the echo audibility experiment (no binaural cues) had been yoked to those used in the echo localization experiment (binaural cues). In this way, direct comparison between experiments enabled us to assess the role played by binaural cues for echolocation in the presence of a masking sound.

We found that when no binaural cues were present (Echo Audibility Experiment), acoustic similarity drove performance, so that participants needed higher intensity echolocation sounds to perceive echoes in the presence of masking sound that was the same type of sound as the emission (i.e. click emissions in clicks masker and noise emissions in noise), as compared to when they were not the same type of sound (i.e. clicks in noise masker, and noise in clicks masker). This is what would be expected based on effects observed for source hearing in the context of masking sound. Yet, results changed dramatically when binaural cues were available, in which case acoustic similarity did not play a role and noise was always the most efficient masker. This was unexpected based on results obtained in source hearing. Further, we found that only clicks in clicks masker experienced a binaural release from masking, with an SNR reduction of 42 dB, which is an incredibly large advantage. Notably, none of the other conditions experienced any release from masking via binaural cues. Since all other conditions contained noise either as emission or masker, this may suggest that echolocation using clicks may differ in sensitivity to binaural cues, as compared to echolocation using noise.

One may ask if the task we used actually required people to echolocate. In our paradigm participants listened to two sounds in two separate intervals. Each sound contained masker and reference or comparison sound, and reference and comparison always contained emission and echo. Any SNR adjustment via adaptive staircase always applied to both reference and comparison. Thus, to perform the task participants had to work with the echo, i.e. they had to echolocate, because emissions etc. were not informative.

Research on sound source localisation suggests that listeners can use lower SNRs to localize sounds when the direction in azimuth of target sounds and interfering noise is different, as compared to when they are spatially coincident (Saberi et al, 1991; Good and Gilkey, 1996). Other work has also reported that the perceived spatial separation of signal and masker results in improved perception and identification of speech signals in noise (Peissig and Kollmeier, 1997; Hawley et al, 1999; Freyman et al. 1999, Arbogast et al. 2002; Hawley et al, 2004; Litovsky, 2012), and tones in noise (Kidd et al, 1998; Lorenzi et al, 1999; Kopčo and Shinn-Cunningham, 2003). In the context of echolocation, research on bats has documented the role of spatial separation between a target surface and the source of noise for object detection. Sümer and colleagues (2009) found that bats accurately detected a wire as the position of the target and the source of interfering sounds became spatially separated in the horizontal plane. They did not present masking noise playbacks, but the source of noise was an object placed at various azimuth angles which reflected bats' own sonar pulses. Signal design by echolocating bats for target localization in azimuth may extend to adaptations other than intensity. For example, timing of bats' emitted pulses increased with increasing interference caused by a distracter (a metal rod) in the horizontal plane (Aytekin et al, 2010); they also observed that emitted pulse duration decreased, and remained short, as the position of the target and the source of interfering sounds became coincident. Though methodological differences do not allow direct comparisons between our findings and those by other studies, our observations indicate that perceived differences in the direction of echoes and masking sound facilitated localization of sounds of interest (i.e. echoes) in the presence of high levels of masking sound. But, as noted above, this was especially true for click emissions and when clicks were used as maskers. Thus, whilst overall our results are consistent with previous work investigating effects of binaural cues for masking, most importantly they also suggest that for human echolocation the type of emission as well as the masker play a role for the effects of binaural cues.

Our study used computer generated emissions, and sound adaptations were limited to modification to the intensity of echolocation sounds. Previous work in the context of active human echolocation (i.e. when people make their own clicks) has highlighted dynamic adjustments in emission intensity and number of emissions, in the absence of adjustments in spectral content, pulse duration or inter-click-intervals (Thaler et al., 2018, 2019, 2022). These studies did not use masking noise, however. Thus, future work should explore the possibility that signal modifications other than intensity can compensate for masking noise in human echolocation, similar to observations of noise-induced pulse adjustments in echolocating bats (Tressler and Smotherman, 2009; Hage et al, 2013; Luo et al, 2015; Lu et al, 2020) and modifications to human speech signals in noise (Lane and Tranel, 1971; Brumm and Zollinger, 2011; Hotchkin and Parks, 2013). A novel prediction that future work could also investigate is, if head movements that introduce binaural differences between target and non-target sound, may be an efficient strategy to improve SNRs for human echolocation using click emissions.

Previous research has shown that people who are blind and have long-term experience in click-based echolocation perform better compared to people who are blind or sighted without experience in echolocation (e.g. Milne et al., 2014; Norman & Thaler, 2019, 2020a, 2020b; Thaler et al., 2020). Long term experience can also affect echo- perceptual judgments of size and weight (e.g. Buckingham et al., 2015; Milne et al., 2015). This suggests that expertise in echolocation rather than blindness drives performance. Alternatively, it has also been suggested that people who are blind are at a particular advantage for learning echolocation (Kolarik et al., 2016, 2021). Contradicting this latter view, and more in line with the idea that experience is key, people who are sighted can learn echolocation just as well as people who are blind and can perform at levels matching or approaching performance levels of blind echolocation experts (e.g. Norman et al., 2021; Teng et al., 2012). The current study, which used a sample size comparable to or exceeding those used in previous work, did not find evidence supporting the idea that the pattern of results differed across participant groups. This suggests that blindness or experience in echolocation play only a limited role for the effects we found. This replicates what we found in our previous study on effects of masking sound on echolocation (Castillo-Serrano et al., 2021). It is possible that the training participants did as part of the experiment (which was the same in our previous and current work), minimized effects of long term experience and/or blindness on performance in the tasks we used. Alternatively, it is possible that the effects we found represent a general principle of human echo-acoustic processing that applies to anyone regardless of visual status or experience with echolocation.

It is important to address whether the results of the present study might generalize to echolocation in more ecologically valid settings. The object size we used was relevant to people who use echolocation in everyday life (e.g. to detect side panel of a bus shelter, a large tree or a person). The click emissions we used were similar to natural human mouth clicks for echolocation (De Vos & Hornikx, 2017; Thaler et al., 2017; Zhang et al., 2017). It was a necessity in the design of this study, however, that participants did not actively generate their own emissions, as otherwise we would have lacked control over acoustics of emissions. It has been shown in a previous study (Thaler & Castillo-Serrano, 2016), however, that when expert echolocators use clicks to detect a target of the same size used here, there is no difference in their performance when they create their own emissions compared to when they use artificial ones similar to those used here. In terms of maskers, these are also expected to have relevance for everyday situations. For example, people echolocating in the presence of other echolocators clicking would be presented with interfering clicks. Such clicks may also be generated by cane tips or footsteps impinging on hard surfaces. Broad band noise could be considered akin to noise created by traffic or rain in terrestrial settings, even though spectral composition of these sounds varies with traffic/precipitation volume, recording position and impingement surface. In sum, we expect, that the current results with click emissions would generalize to active echolocation in ecologically valid settings.

Echolocation can provide real life advantages for people who are blind in terms of mobility, independence and wellbeing (Norman et al., 2021; Thaler, 2013). Importantly, we replicated previous findings (Castillo-Serrano & Thaler, 2021) that even in the presence of masking noise click-based echolocation is an effective sensing tool, and provided an important extension demonstrating that adjustments of the intensity of emissions are sensitive to the type of background sound and binaural information present. Our results exemplify that for successful echolocation people need dynamic control of the signals that carry relevant acoustic information to support their behaviour. This information will be useful for instruction and of guidance for new users.

Declarations of interest statement

Declarations of interest: none.

Acknowledgments

We thank all our participants for taking part. This work was partially supported by BBSRC grant to LT (BB/M007847/1) and by a PhD studentship to JGCS by National Council of Science and Technology (CONACyT) Mexico (600799).

References

Amichai, E., Blumrosen, G., Yovel, Y., 2015. Calling louder and longer: how bats use biosonar under severe acoustic interference from other bats. Proc. R. Soc. B. 282: 20152064. https://doi.org/10.1098/rspb.2015.2064.

Arbogast, T.L., Mason, C.R., and Kidd, Jr, G., 2002. The effect of spatial separation on informational and energetic masking of speech. J. Acoust. Soc. Am. 112, 2086–2098. https://doi.org/10.1121/1.1510141.

Arbogast, T.L., Mason, C.R., Kidd, Jr, G., 2005. The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America, 117, 2169–2180. https://doi.org/10.1121/1.1861598.

Aytekin, M., Mao, B., and Moss, C.F., 2010. Spatial perception and adaptive sonar behavior. J. Acoust. Soc. Am. 128, 3788–3798. https://doi.org/10.1121/1.3504707.

Bates, M.E., Stamper, S.A. and Simmons, J.A., 2008. Jamming avoidance response of big brown bats in target detection. J. Exp. Biol. 211, 106-113 https://doi.org/10.1242/jeb.009688.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Bronkhorst, A., 2000. The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. Acustica 86, 117–128.

Brumm, H., Slabbekoorn, H., 2005. Acoustic communication in noise. Advances in the Study of Behavior 35, 151–209.

Brumm, H., Zollinger, S.A., 2011. The evolution of the Lombard effect: 100 years of psychoacoustic research. Behaviour. 148: 1173 – 1198. https://doi.org/10.1163/ 000579511X605759.

Brungart, D.S., 2001. Informational and energetic masking effects in the perception of two simultaneous talkers. J. Acoust. Soc. Am. 109, 1101–1109. https://doi.org/10.1121/1.1345696.

Buckingham, G., Milne, J.L., Byrne, C.M., Goodale, M.A. 2014. The size-weight illusion induced through human echolocation. Psychol Sci 2014, 26:237–242. https://doi.org/10.1177/0956797614561267.

Castillo-Serrano J.G., Norman L.J., Foresteire D., & Thaler L., 2021. Increased emission intensity can compensate for the presence of noise in human click-based echolocation. *Sci Rep.* 11, 1750. https://doi.org/10.1038/s41598-021-81220-9.

de Vos R, Hornikx M., 2017. Acoustic properties of tongue clicks used for human echolocation. Acta Acust United Acust. 103, 1106–1115. https://doi.org/10.3813/AAA.919138.

de Vos, R., & Hornikx, M. (2018). Human ability to judge relative size and lateral position of a sound reflecting board using click signals: Influence of source position and click properties. Acta Acustica united with Acustica, 104(1), 131-144.

Dodsworth, C., Norman, L. J., & Thaler, L. (2020). Navigation and perception of spatial layout in virtual echo-acoustic space. Cognition, 197, 104185.

Dufour, A., Després, O., Candas, V., 2005. Enhanced sensitivity to echo cues in blind subjects. Exp Brain Res. 165: 515–519. https://doi.org/10.1007/s00221-005-2329-3.

Durlach, N.I., Mason, C.R., Shinn-Cunningham, B.G, Arbogast, T.L., Colburn, H.S., Kidd, JR, G., 2003. Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. J. Acoust. Soc. Am. 114, 368–379. https://doi.org/10.1121/1.1577562.

Freyman, R., Helfer, K., McCall, D., Clifton, R., 1999. The role of perceived spatial separation in the unmasking of speech. J. Acoust. Soc. Am. 106, 3578–3587. https://doi.org/10.1121/1.428211.

Good, M.D., Gilkey, R.H., 1996. Sound localization in noise: The effect of signal-to-noise ratio. J Acoust Soc Am. 99:1108–1117. https://doi.org/10.1121/1.415233.

Hage, S.R., Jiang, T., Berquist, S.W., Feng, J., Metzner, W., 2013. Ambient noise induces independent shifts in call frequency and amplitude within the Lombard effect in echolocating bats. Proceedings of the National Academy of Sciences, 110, 4063–4068. https://doi.org/10.1073/pnas.1211533110.

Hawley, M.L., Litovsky, R.Y., Colburn, H.S., 1999. Speech intelligibility and localization in complex environments. J. Acoust. Soc. Am. 105, 3436–3448. https://doi.org/10.1121/1.424670.

Hawley, M., Litovsky, R., Culling, J., 2004. The benefit of binaural hearing in a cocktail party: Effects of location and type of interferer. J. Acoust. Soc. Am. 115, 833–843. https://doi.org/10.1121/1.1639908.

Hiryu, S., Hagino, T., Riquimaroux, H. & Watanabe, Y., 2007. Echo-intensity compensation in echolocating bats (Pipistrellus abramus) during flight measured by a telemetry microphone. J. Acoust. Soc. Am. 121, 1749–1757. https ://doi.org/10.1121/1.2431337.

Hotchkin, C., Parks, S., 2013. The Lombard effect and other noise-induced vocal modifications: insight from mammalian communication systems. Biol. Rev., 88 (4), pp. 809-824. https://doi.org/10.1111/brv.12026.

Jones, G., 2005. Echolocation. Curr Biology. 15(13): 484–488.

Kidd, Jr, G., Mason, C.R., Arbogast, T.L., 2002. Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. J. Acoust. Soc. Am. 111, 1367–1376. https://doi.org/10.1121/1.1448342.

Kidd, Jr, G., Mason, C.R., Rohtla, T.L., Deliwala, P.S., 1998. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. J. Acoust. Soc. Am. 104, 422–431. https://doi.org/10.1121/1.423246.

Kolarik, A.J., Cirstea, S., Pardhan, S., Moore, B.C., 2014. A summary of research investigating echolocation abilities of blind and sighted humans. Hear Res, 310C:60–68. https://doi.org/10.1016/j.heares.2014.01.010.

Kolarik, A.J., Pardhan, S., Cirstea, S., Moore, B.C.J., 2013. Using acoustic information to perceive room size: effects of blindness, room reverberation time, and stimulus. Perception 42, 985-990. https://doi.org/10.1068/p7555.

Kolarik, A. J., Pardhan, S., & Moore, B. C. J. (2021). A framework to account for the effects of visual loss on human auditory abilities. Psychological Review, 128(5), 913–935

Kopčo, N., Shinn-Cunningham, B., 2003. Spatial unmasking of nearby pure tone targets in a simulated anechoic environment. J. Acoust. Soc. Am. 114: 2856–2870. https://doi.org/10.1121/1.1616577.

Lane, H., Tranel, B., 1971. The Lombard sign and the role of hearing in speech. J. Speech Hear. Res. 14, 677–709. https://doi.org/10.1044/jshr.1404.677.

Litovsky, R.Y., 2012. Spatial release from masking. Acoust. Today 8, 18–25. https://doi.org/10.1121/1.4729575.

Lorenzi, C., Gatehouse, S., Lever, C., 1999. Sound localization in noise in normal-hearing listeners. J Acoust Soc Am. 105:1810–1820. https://doi.org/10.1121/1.426719.

Luo, J., Goerlitz, H., Brumm, H., 2016. Linking the sender to the receiver: vocal adjustments by bats to maintain signal detection in noise. Sci Rep 5, 18556. https://doi.org/10.1038/srep18556.

Luo, J., Siemers, B.M., Koselj, K., 2015. How anthropogenic noise affects foraging. Glob Change Biol 29, 3278–3289. https://doi.org/10.1111/gcb.12997.

Martinez-Rojas, J.A., Hermosilla, J.A., Montero, R.S., Espi-Lopez, P.L., 2009. Physical analysis of several organic signals for human echolocation: oral vacuum pulses. Acta Acust. United Acust. 95, 325-330. https://doi.org/10.3813/AAA.918155.

Milne, J.L., Anello, M., Goodale, M.A., Thaler, L., 2015. A blind human expert echolocator shows size constancy for objects perceived by echoes. Neurocase. 21:465–470. https://doi.org/10.1080/13554794.2014.922994.

Milne, J.L., Goodale, M.A., Thaler, L., 2014. The role of head movements in the discrimination of 2-D shape by blind echolocation experts. Atten Percept Psychophys. 76:1828–1837. https://doi.org/10.3758/s13414-014-0695-2.

Norman, L.J., Thaler, L., 2018. Human echolocation for target detection is more accurate with emissions containing higher spectral frequencies, and this is explained by echo intensity. I-Perception, 9 (3), pp. 1-18. https://doi.org/10.1177/2041669518776984.

Norman, L.J., Thaler L., 2019. Retinotopic-like maps of spatial sound in primary ‘visual’ cortex of blind human echolocators. Proc. R. Soc. B. 286: 20191910. https://doi.org/10.1098/rspb.2019.1910.

Norman, L. J., Thaler, L., 2020a. Stimulus uncertainty affects perception in human echolocation: Timing, level, and spectrum. *Journal of Experimental Psychology: General, 149*(12), 2314-2331. https://doi.org/10.1037/xge0000775.

Norman, L., Thaler, L., 2020b. Perceptual constancy with a novel sensory skill. *Journal of Experimental Psychology: Human Perception and Performance, 47*(2), 269-281.DOI: 10.1037/xhp0000888

Norman, L., Dodsworth, C., Foresteire, D., Thaler, L., 2021. Human click-based echolocation: Effects of blindness and age, and real-life implications in a 10-week training program. PLoS ONE 16(6): e0252330. https://doi.org/10.1371/journal.pone.0252330

Peissig, J., Kollmeier, B., 1997. Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. J. Acoust. Soc. Am. 35, 1660–1670. https://doi.org/10.1121/1.418150.

Prins, N., Kingdom, F.A.A., 2018. Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. Front. Psychol. 9, 1250. https://doi.org/10.3389/fpsyg .2018.01250.

Rowan, D., Papadopoulos, T., Archer, L., Goodhew, A., Cozens, H., Lopez, R. G., . . . Allen, R., 2017. The detection of ‘virtual’ objects using echoes by humans: Spectral cues. Hearing Research, 350, 205–216. https://doi.org/10.1016/j.heares.2017.04.008.

Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., Perrot, D.R., 1991. Free - field release from masking. J. Acoust. Soc. Am. 90, 1355–1370. https://doi.org/10.1121/1.401927.

Schenkman, B. N., & Nilsson, M. E. (2010). Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. Perception, 39(4), 483-501.

Schnitzler, H-U., Moss, C.F., Denzinger, A., 2003. From spatial orientation to food acquisition in echolocating bats. Trends Ecol. Evol. 18, 386–394. https://doi.org/10.1016/S0169-5347(03)00185-X.

Siemers, B., Schnitzler, H-U., 2004. Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. Nature 429, 657–661. https://doi.org/10.1038/nature02547.

Steeneken, H.J., Geurtsen, F.W. 1988. Description of the RSG-10 Noise Database. TNO Institute Perception, Technical report IZF 1988–3. http://www.steen eken.nl/7-noise -data-base/ (accessed April 2017).

Steffens, H., Schutte, M., & Ewert, S. D. (2022). Auditory orientation and distance estimation of sighted humans using virtual echolocation with artificial and self-generated sounds. JASA Express Letters, 2(12).

Sümer, S., Denzinger, A., Schnitzler, H.-U., 2009. Spatial unmasking in the echolocating Big Brown Bat, Eptesicus fuscus. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 195, 463. https://doi.org/10.1007/s00359-009-0424-9.

Teng, S., & Whitney, D. (2011). The acuity of echolocation: Spatial resolution in sighted persons compared to the performance of an expert who is blind. Journal of Visual Impairment & Blindness, 105(1), 20-32.

Thaler, L., 2013. Echolocation may have real-life advantages for blind people: an analysis of survey data. Front Physiol 4:98. https://doi.org/10.3389/fphys.2013.00098.

Thaler, L., Castillo-Serrano, J.G., 2016. People's ability to detect objects using click-based echolocation: a direct comparison between mouth-clicks and clicks made by a loudspeaker. PLoS One 11:e0154868. https://doi.org/10.1371/journal.pone.0154868.

Thaler, L., De Vos, R., Kish, D., Antoniou, M., Baker, C., Hornikx, M., 2018. Human echolocators adjust loudness and number of clicks for detection of reflectors at various azimuth angles. Proc. R. Soc. B. 285: 2017273520172735. http://doi.org/10.1098/rspb.2017.2735.

Thaler, L., De Vos, R., Kish, D., Antoniou, M., Baker, C., & Hornikx, M., 2019. Human click-based echolocation of distance: Superfine acuity and dynamic clicking behaviour. Journal of the Association for Research in Otolaryngology, 20(5), 499–510. https://doi.org/10.1007/s10162-019-00728-0.

Thaler, L., Zhang, X., Antoniou, M., Kish, D. & Cowie, D., 2020. The flexible Action System: Click-based Echolocation may replace certain visual Functionality for adaptive Walking. *Journal of Experimental Psychology: Human Perception and Performance,* 46(1), 21-35. DOI: 10.1037/xhp0000697.

Thaler, L., Norman, L., De Vos, R., Kish, D., Antoniou, M., Baker, C., Hornikx, M. (2022). Human Echolocators have better Localization Off Axis. Psychological Science. 33(7): 1143–1153. doi: 10.1177/09567976211068070

Thaler, L., Goodale, M.A., 2016. Echolocation in humans: an overview. Wiley Interdiscip. Rev. Cogn. Sci. 7, 382–393. https://doi.org/10.1002/wcs.1408.

Thaler, L., Reich, G.M., Zhang, X., Wang, D., Smith, G.E., Tao, Z., … Antoniou, M., 2017. Mouth-clicks used by blind expert human echolocators: signal description and model based signal synthesis. PLoS Comput. Biol. 13(8): e1005670. https://doi.org/10.1371/journal.pcbi.1005670.

Thomas, J. A., Moss, C. F., & Vater, M. (Eds.). (2004). *Echolocation in bats and dolphins*. University of Chicago press.

Tirado, C., Gerdfeldter, B., Kärnekull, S. C., & Nilsson, M. E. (2021). Comparing echo-detection and echo-localization in sighted individuals. Perception, 50(4), 308-327.

Tirado, C., Lundén, P., & Nilsson, M. E. (2019). The Echobot: An automated system for stimulus presentation in studies of human echolocation. Plos one, 14(10), e0223327.

Tressler, J., Smotherman, M.S., 2009. Context-dependent effects of noise on echolocation pulse characteristics in free-tailed bats. J Comp Physiol A 195, 923–934. https://doi.org/10.1007/s00359-009-0468-x.

Wallmeier, L., Geßele, N., & Wiegrebe, L. (2013). Echolocation versus echo suppression in humans. Proceedings of the Royal Society B: Biological Sciences, 280(1769), 20131428.

Citation on deposit: Thaler, L., Castillo-Serrano, J. G., Kish, D., & Norman, L. (2024). Effects of type of emission and masking sound, and their spatial correspondence, on blind and sighted people's ability to echolocate. Neuropsychologia, Article 108822. https://doi.org/10.1016...psychologia.2024.108822