Hearing Research 430 (2023) 108722

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Sound localization acuity of the common marmoset (*Callithrix jacchus*)

Chenggang Chen[1], Evan D. Remington[1], Xiaoqin Wang*

Department of Biomedical Engineering, Johns Hopkins University School of Medicine, 720 Rutland Ave., Traylor 410, Baltimore, MD 21025, United States

ARTICLE INFO

Article history:
Received 13 August 2022
Revised 3 February 2023
Accepted 10 February 2023
Available online 12 February 2023

Keywords:
Marmoset
Primate hearing
Minimum audible angle
Spatial acuity
Spectral cues
Likelihood ratio test
Bayesian inference

ABSTRACT

The common marmoset (*Callithrix jacchus*) is a small arboreal New World primate which has emerged as a promising model in auditory neuroscience. One potentially useful application of this model system is in the study of the neural mechanism underlying spatial hearing in primate species, as the marmosets need to localize sounds to orient their head to events of interest and identify their vocalizing conspecifics that are not visible. However, interpretation of neurophysiological data on sound localization requires an understanding of perceptual abilities, and the sound localization behavior of marmosets has not been well studied. The present experiment measured sound localization acuity using an operant conditioning procedure in which marmosets were trained to discriminate changes in sound location in the horizontal (azimuth) or vertical (elevation) dimension. Our results showed that the minimum audible angle (MAA) for horizontal and vertical discrimination was 13.17° and 12.53°, respectively, for 2 to 32 kHz Gaussian noise. Removing the monaural spectral cues tended to increase the horizontal localization acuity (11.31°). Marmosets have larger horizontal MAA (15.54°) in the rear than the front. Removing the high-frequency (> 26 kHz) region of the head-related transfer function (HRTF) affected vertical acuity mildly (15.76°), but removing the first notch (12–26 kHz) region of HRTF substantially reduced the vertical acuity (89.01°). In summary, our findings indicate that marmosets' spatial acuity is on par with other species of similar head size and field of best vision, and they do not appear to use monaural spectral cues for horizontal discrimination but rely heavily on first notch region of HRTF for vertical discrimination.

1. Introduction

The common marmoset (*Callithrix jacchus*) has been a successful model for studying the auditory system in the past two decades (Wang, 2018). The marmoset has been used to study the coding of pitch and complex spectral features in the auditory cortex (Bendor and Wang, 2005; Zhu et al., 2019; Zeng et al., 2021), temporal processing in the auditory cortex (Lu et al., 2001; Kajikawa et al., 2008; Zhou and Wang, 2010; Gao et al., 2016; Liu and Wang, 2022), thalamus (Bartlett and Wang, 2007), and inferior colliculus (Wang et al., 2022), spectral and intensity coding (Sadagopan and Wang, 2008; Song et al., 2022a), and sound source location (Remington and Wang, 2019; Zhou and Wang, 2014, 2012; Lui et al., 2015; Chen et al., 2022). Much is known about auditory cortex connectivity in marmosets (de la Mothe et al., 2012; Reser et al., 2009). Additionally, a number of studies have investigated auditory feedback mechanisms during vocalization (Eliades and Wang, 2008; Hage, 2020) and processing and control of conspecific communication in the prefrontal cortex (Roy et al., 2016; Jovanovic et al., 2022). As a model for hearing loss, marmosets have been used to study the representation of cochlear implant stimulation at the level of the auditory cortex (Johnson et al., 2012, 2016, 2017). Germline expression of a transgenic modification has been achieved in the marmoset (Park et al., 2016; Sasaki et al., 2009), broadening the marmoset's potential as a model for neurologic diseases. There has also been an increasing interest in marmosets as a model for visual cognition (Mitchell et al., 2014; Davis et al., 2020; Song et al., 2022b).

Compared with the anatomy and physiology of the auditory system in marmosets, much less is known of their perceptual abilities. Our laboratory has developed an auditory operant conditioning task for the common marmoset to study auditory perception and for use in behaving neurophysiology (Remington et al., 2012). The task has been employed in the measurement of the audiogram (Osmanski and Wang, 2011), frequency discrimination thresholds (Osmanski et al., 2016), and harmonic resolvability and pitch perception (Osmanski et al., 2013; Song et al., 2016). Here we used this task to measure the spatial hearing acuity of the marmoset, specifically minimum audible angle (MAA).

Abbreviations: MAA, minimum audible angle; HRTF, head-related transfer function.

* Corresponding author.
E-mail address: xiaoqin.wang@jhu.edu (X. Wang).
[1] Co-first author.

https://doi.org/10.1016/j.heares.2023.108722
0378-5955/

Marmoset exhibit rapid head movement (gaze shift) towards sounds of interest using only auditory spatial cues (Slee and Young, 2010; Pandey et al., 2020). When wild or captive marmosets are not visible to each other, they both increase the duration, rate, amplitude, and frequency of their calls to facilitate the sound localization of recipients (Snowdon and Hodun, 1981; Liao et al., 2018). Spatial processing is therefore an important function performed by the marmoset's auditory system, making the marmoset an ideal model species for further studies of spatial coding in the auditory cortex (Remington and Wang, 2019; Zhou and Wang, 2014, 2012).

As in all mammals, sound location perception is determined by three cues: interaural time difference (ITD), interaural level difference (ILD), and spectral shape (Blauert, 1997). These cues are the result of the geometry of the head and ears: the distance between the ear canals determines ITD for localizing low-frequency sounds, the size and shape of the head (and to an extent the neck and shoulders) determines ILD for localizing high-frequency sounds, and the shape of the pinna (or outer ear) modulates the shape of the incoming sound spectrum by introducing orderly variation in the frequency of the first spectral notch occurs with changes in the elevation. This spatially dependent acoustic filter is referred to as the head related transfer function (HRTF) (Blauert, 1997).

Although sound localization in the horizontal plane relies primarily on binaural ITD and ILD cues, the notch frequency and depth of HRTFs also change systematically with azimuth in several species (Wightman and Kistler 1989; Rice et al., 1992; Slee and Young, 2010) (Fig. 1A). Therefore, there is a possibility that in addition to binaural cues, monaural spectral cues may also help marmosets perform a horizontal discrimination task (Butler, 1986; Van Wanrooij and Van Opstal, 2004). However, experimental evidence is lacking. Furthermore, studies in humans and macaques found that sound localization is less accurate at rear locations compared with frontal locations (Oldfield and Parker, 1984; Recanzone and Beckerman, 2004). Whether or not this holds true for marmoset is still unknown.

Sound localization in the vertical plane relies on spectral cues. The HRTF of mammals contains several regions which could be useful for directional hearing (Fig. 1B). For example, marmoset's HRTF includes a first notch region (12 to 26 kHz), in which there exists a prominent spectral notch that varies in frequency and depth systematically with elevation, and a high-frequency region (>24 kHz), which inhabits the upper range of audibility and contains cues that vary less orderly (Rice et al., 1992; Slee and Young, 2010). Interestingly, studies in macaques and cats found that vertical discrimination was degraded by the removal of high frequency region instead of first notch region (Brown et al., 1982). Whether or not first notch region contributes to marmoset vertical sound localization is unknown.

In this study, we first measured the marmoset's horizontal and vertical spatial acuity (MAA) using 2 to 32 kHz Gaussian noise. We also tested 1) whether horizontal MAA in the front and rear was similar, 2) whether monaural spectral cues affected horizontal MAA using random-spectral shape (RSS) stimuli, and 3) whether first notch and high-frequency regions contributed to vertical MAA using Gaussian noise that filtered out the first notch region (4 to 12 kHz) and high-frequency region (4 to 26 kHz), respectively. Our results indicated that the marmoset's horizontal and vertical spatial acuity was approximately on par with other species of a similar head size and field of best vision. Removing the monaural spectral cues even increased the horizontal localization acuity. Consistent with previous studies, we found that marmosets have a larger horizontal MAA on the rear than the front. In contrast with studies in other species, it is the first notch region instead of the high frequency region of HRTF that significantly contributed to the vertical localization acuity.

2. Materials and methods

2.1. Subjects

Subjects were five common marmoset monkeys (2 male, 3 female, 2–5 years old) housed in individual cages in a colony at The Johns Hopkins University School of Medicine. Each animal was given free access to water; food was regulated during testing periods to keep animals at approximately 90% of free feeding weight. Subjects were tested once a day, five to seven days per week between the hours of 0900 and 1800. All experimental procedures were approved by the Institutional Animal Care and Use Committee of the Johns Hopkins University following National Institutes of Health guidelines.

2.2. Testing chamber and acoustic stimuli

Experiments were conducted in a double-walled sound-attenuated chamber (Industrial Acoustics, IAC, New York) with the internal walls, ceiling, and floor lined with ~3 inch acoustic absorption foam (Sonex, Illbruck). Acoustic stimuli were delivered from an array of 5–7 speakers (FT28D, Dome Tweeter, Fostex) mounted 1 m from an animal's head in a horizontal or vertical arc. Horizontal speakers (at 0° elevation) were placed in both frontal and rear locations. Vertical speakers were placed in the median plane in front of subjects (at 0° azimuth). Subjects sat in a wire mesh primate chair mounted onto a single stainless-steel bar such that the animal's head was centered in the room (Remington et al., 2012). Animal's head was free to move. The primate chair was designed to minimize acoustic reflections near the pinna (Fig. 1C). Marmosets moved their pinnae very infrequently, which usually happened when experimenters touched the hairs around the pinnae. Marmosets did not orient their pinnae toward sound sources, and their pinnae position did not change over time in our experimental preparation.

Stimuli were generated in MATLAB (Mathworks, Natick, MA) at a sampling rate of 97.7 kHz using custom software. Digital signals were converted to analog (RX6, Tucker-Davis Technologies, Alachua, FL), then analog signals were attenuated (PA5, Tucker-Davis Technologies), power amplified (Crown Audio, Elkhart, IN), and played through a power multiplexer (PM2R, Tucker-Davis Technologies). Loudspeakers had a relatively flat frequency response curve (± 3–7 dB) and minimal spectral variation across speakers (< 7 dB re mean) across the range of frequencies of the stimuli used; all large (5–7 dB) spectral deviations occurred in narrow bandwidths near the upper limit of speakers' frequency range (above 28 kHz), above the first spectral notch measured in marmoset head related transfer functions (Slee and Young, 2010).

Test stimuli included unfrozen (i.e., trial-unique) Gaussian noises that were band-pass filtered to contain energy between 2 and 32 kHz, 4 and 26 kHz, 4 and 12 kHz and Random Spectral Shape (RSS) stimuli (Barbour and Wang, 2003). RSS stimuli were constructed by summing pure-tones with pseudo-random levels centered around an average sound level. In this report all RSS stimuli were constructed to have energy between 2 and 32 kHz, with 30 tones per octave with levels varying independently in groups of three tones (i.e., 10 bins per octave). The standard deviation of the bin levels was 10 dB. We filtered out frequencies below 2 kHz to accommodate speakers' frequency range and be consistent with our previous neurophysiology experiments (Remington and Wang, 2019; Zhou and Wang, 2014). Stimuli were 200 ms long with 10 ms cosine ramps. Sound level was roved ±10 dB to avoid the use of absolute level as a cue. The mean intensity was −35 dB relative to a maximum output intensity of 95 dB SPL at 0 dB attenuation for a 4 kHz tone.

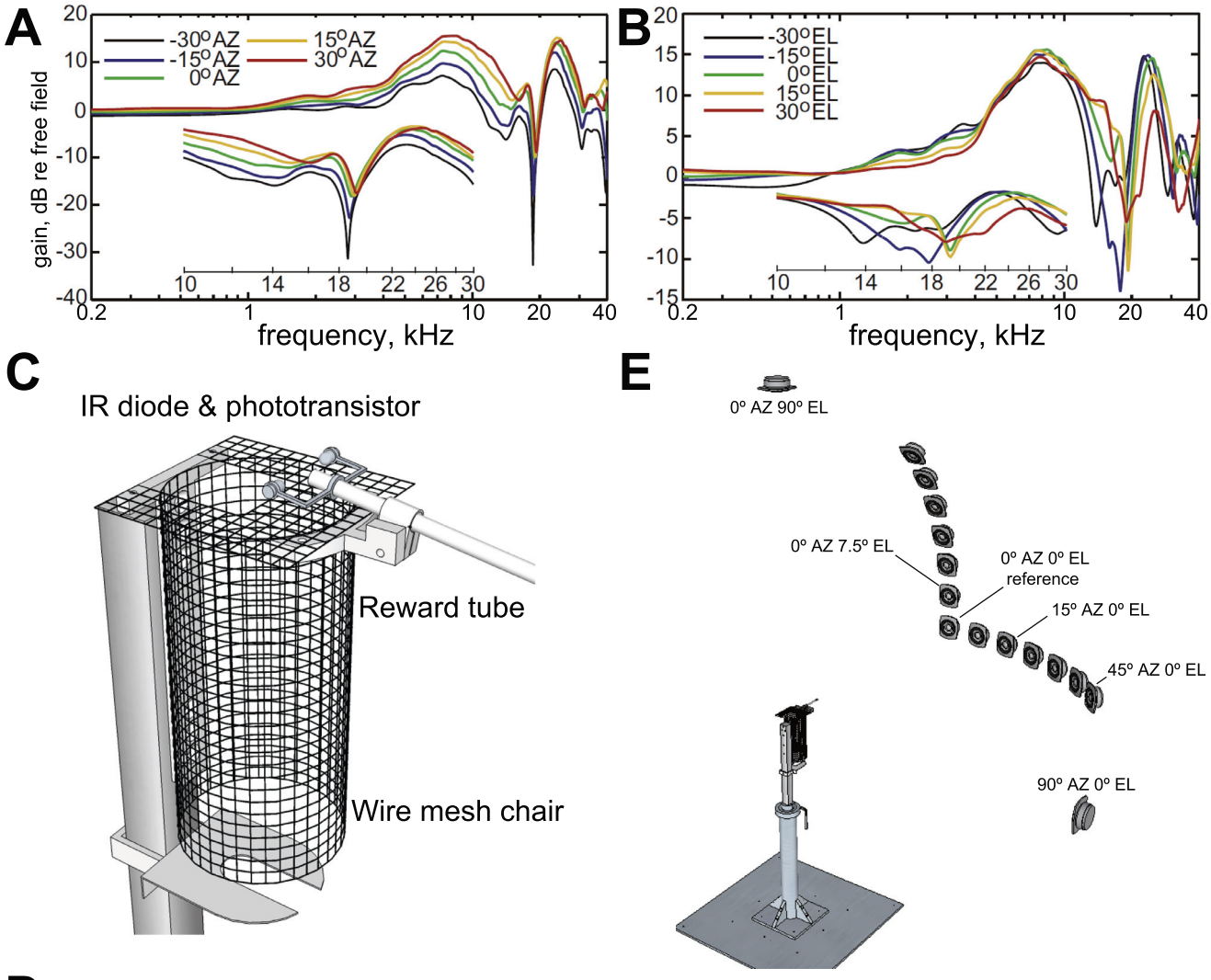

Fig. 1. Marmoset's HRTF, chair, psychophysical testing procedure, and speaker arrangement for measurement of minimum audible angle (MAA)**.** Magnitudes of HRTFs were measured at locations in the frontal field for changes in azimuth (AZ, **A**) and elevation (EL, **B**). Figures are copied from Fig. 8A, B of Slee and Young (2010). **C.** Marmoset chair. Behavior sessions for MAA measurements were conducted while marmosets were seated in a steel wire chair designed to minimize acoustic reflections. Marmosets had to break a photobeam (photodiode & phototransistor) to register a response, which could be accomplished either by licking at the reward tube or moving the entire head forward. **D.** Location discrimination task trial structure. Marmosets listened to sounds played from a reference location, and had to respond when a target location will began alternating with the reference location. If a response was registered within the preset number of alternations, a food reward was given. The next trial began after the animal finished consuming the reward (as measured via the photo beam). A response outside of a target interval resulted in a timeout. **E.** Speaker arrangement for MAA measurements. We measured localization discrimination thresholds in three conditions: frontal azimuth, rear azimuth, and vertically along the median plane. For rear azimuth discrimination, the chair was rotated 180° Head position was monitored and with a closed-circuit camera system and custom image processing software. Marmosets were required to have heads facing forward in order for behavioral trials to proceed. Rear location is not pictured. The reference location was at 0° azimuth and 0° elevation as shown.

2.3. Psychophysical testing procedure

Marmosets were trained to sit in the custom-designed primate chair and perform a Go/No-Go spatial location discrimination task (Brown et al., 1980) using the method of constant stimuli (Gescheider, 1985). Details of behavioral training and behavioral apparatus have been described previously (Osmanski and Wang, 2011; Remington et al., 2012). Behavioral responses involved breaking a photo beam attached to a feeding tube, either by moving the entire head slightly forward into the beam or with licks at the tube (Fig. 1C). Each behavior session was composed of a preset number of trials (typically 80–100), where each trial was composed of a variable duration 'inter-target interval' and a fixed duration' response interval.' Inter-target interval duration was randomized between approximately 3 and 10 s. The response interval was dependent on the number and duration of targets and was typically 5 s in length. During an inter-target interval, a series of sounds were played from a fixed reference location (front or rear, 0° azimuth and 0° elevation) with a 700 ms inter-stimulus interval. The duration of sounds played from reference and target locations were both 200 ms. Behavioral responses during this time result in a mild punishment (such as a timeout or a puff of air at the base of the tail) and a restarting of the trial. If an animal withheld response during the inter-target interval, the sound location began alternating between one of several target locations and the reference location during the response interval. A trial ended when the response interval expired or a response was detected during the response interval. Behavioral responses during this time were reinforced with a small food reward (Fig. 1D). Approximately 30% of trials were "catch trials," which were identical in length to target trials in their timing and structure but in which no targets are delivered (i.e., sounds continued to be delivered from the reference speaker). Thus, during a catch trial the response interval is indistinguishable from the inter-target interval from the animal's perspective. A response during a catch trial response interval is referred to as a false alarm (or false positive). Animals which withheld a response successfully on catch trials (correct reject) did not receive a reward.

Head position was tracked using an infrared camera and a custom algorithm written for MATLAB using the image processing toolbox. Prior to each stimulus delivery, the algorithm checked to see if the eyes could be located in an experimenter-defined region. Therefore, animals had to face forward in order for behavioral trials to proceed. If an animal closed its eyes, the algorithm would fail to find the eyes and the program would pause. In addition to assuring that animals maintained open eyes, a specific head position can be considered as indication that an animal is engaged in the task.

To quantify behavior performance for the measurement of MAA in marmosets, we constructed psychometric functions over a fixed set of distances (i.e., target locations) from the reference location. Angles between the reference location and target locations were 7.5°, 15°, 22.5°, 30°, 37.5°, 45°, 90°, and 180° (Fig. 1E). We did not test 30° and 37.5° in the rear locations. The target locations were chosen for each MAA measurement based on animals' performance early in testing such that hit rates were sufficiently high to motivate animals to perform the task. In some cases, the hardest targets were omitted for certain animals. Final thresholds were calculated from psychometric functions averaged over 2 to 5 sessions during which thresholds were no higher than the mean plus 5° and during which time thresholds did not appear to be systematically decreasing (Table 1).

Thresholds were determined to be the point at which the linearly interpolated psychometric function equaled 50%. To adjust for false positives, a corrected hit rate was used:

$$HRc = \frac{HR - FAR}{1 - FAR}$$

where {HR} is the raw hit rate, and {FAR} is the false alarm (or false positive) rate (Gescheider, 1985). Sessions with a false alarm rate (FA/SM<0.3, Table 2) greater than 30% were not included in threshold calculations (less than 15%). Not all animals were tested in each condition (n/a, Table 2).

In addition to interpolation of the psychometric function, we have also fitted the psychometric function to get the threshold (MAA). The psychometric function ψ is defined by four parameters: α, threshold or MAA; β, slope; γ; guess rate or lower asymptote; λ, lapse rate or upper asymptote.

$$\psi(x;\alpha,\beta,\lambda,\gamma) = \gamma + (1-\gamma-\lambda)F(x:\ \alpha,\ \beta)$$

x denotes the stimulus intensity (distance between two speakers), and F is the sigmoid function. Here, we chose the reversed Gumbel function, which has a longer tail:

$$F(x:\ \alpha,\ \beta) = \exp\left(\log(0.5)e^{C\frac{x-\alpha}{\beta}}\right)$$

$$C = \log(-\log(0.95)) - \log(-\log(0.05))$$

Because of the limited trials collected in the psychophysics experiments, the parameters of a psychometric function are not fully constrained by the empirical data. Therefore, the uncertainty of estimated parameters always remains. Typically, the uncertainty is expressed in the form of confidence intervals around the point estimates (Schutt et al., 2016). Here, we use Bayesian inference to estimate the posterior distribution of the MAA with a numerical integration method (Schutt et al., 2016). We chose the psignifit 4 MATLAB toolbox (https://github.com/wichmann-lab/psignifit/wiki) to estimate the MAAs (Fig. 2 right, Figs. 3, 5 right and 6).

In order to compare two models (i.e., two stimuli conditions), we used two methods: transformed likelihood ratio (TLR) test and Bayesian inference. Model comparisons (Figs. 4 and 7) were performed using the Palamedes MATLAB toolbox (https://www.palamedestoolbox.org/index.html).

The likelihood function associated with threshold/MAA α and slope β is:

$$L(\alpha,\beta|y) = \prod_{k=1}^{N} p(y_k|x_k;\alpha,\beta)$$

$p(y_k|x_k;\alpha,\beta)$ is the probability of observance of response y (lick) on trial k given stimulus angle x, and N is the number of total trials across stimuli. Since we assumed each trial is independent, the likelihood of all trial is the product of the likelihood of each trial.

The goal of maximum likelihood estimation is to choose the parameters α, β that maximize (arg max, argument of the maximum) the likelihood function $L(\alpha,\beta|y)$.

$$\widehat{\alpha,\beta} = \arg\max_{\alpha,\ \beta} L(\alpha,\beta|y)$$

$\widehat{\alpha,\beta}$ represent the best choice of values for the two parameters.

To compare the difference between paired models, we define 'lesser model' for the more restrictive model and 'fuller model' for the less restrictive model. In other words, the lesser model is nested under the fuller model. For example, either the threshold or slope is constricted in the lesser model but both parameters are unconstructed in the fuller model. For each paired model, we can calculate the statistical 'p-value' using a likelihood ratio test. The likelihood ratio is the ratio of the likelihood under the lesser model to the likelihood under the fuller model.

$$\Delta = \frac{L\left(\widehat{\alpha_L,\beta_L}|y\right)}{L\left(\widehat{\alpha_F,\beta_F}|y\right)}$$

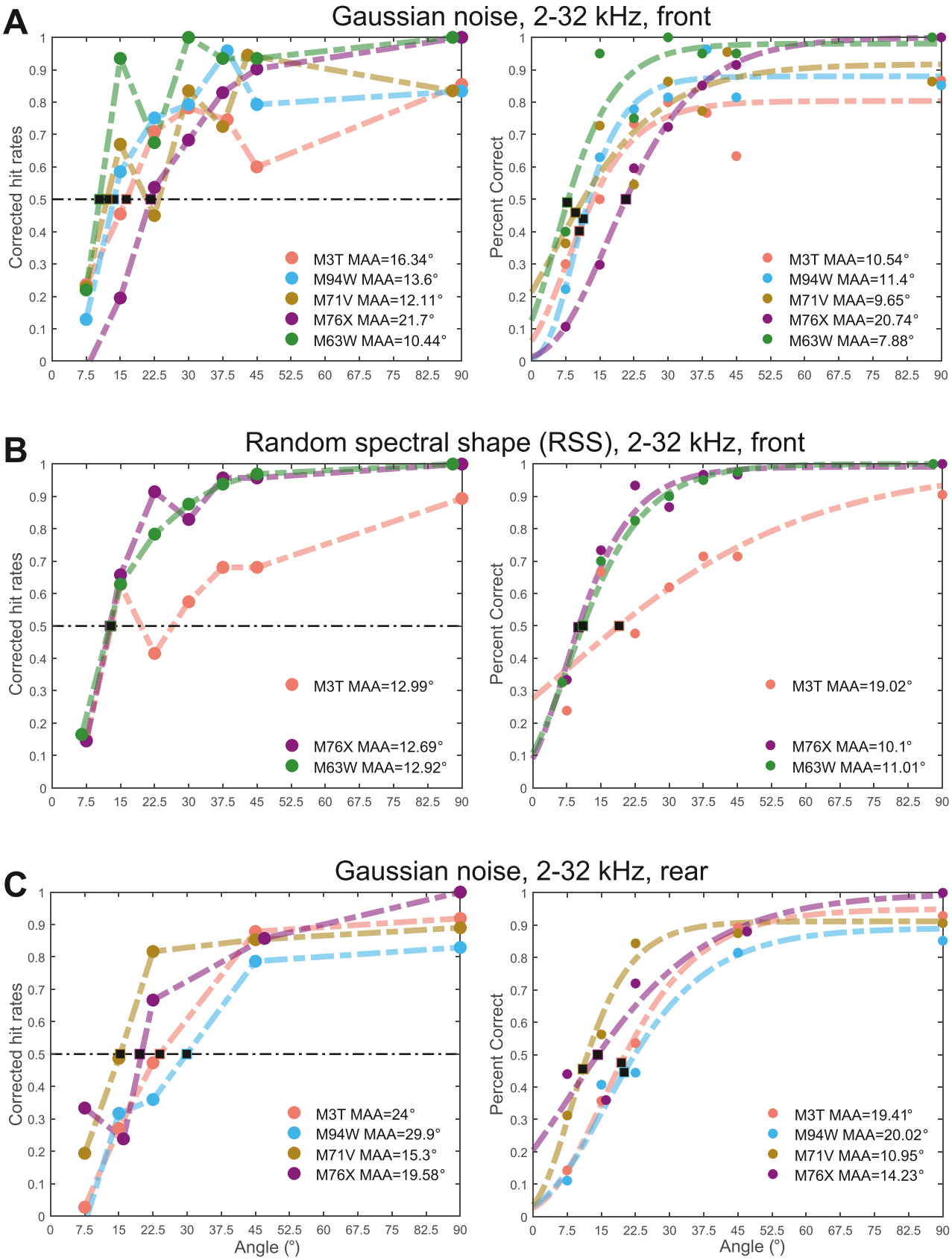

Fig. 2. Psychometric functions for horizontal location discrimination for different stimuli. Different marmosets are represented by different colors. MAAs are represented by black squares. In the left column, data points represent averaged corrected hit rates to different angles between the reference and target locations. Black dashed lines represent the threshold. In the right column, data points represent hit rate, and lines represent the fitted psychometric function. **A.** Gaussian noise, band passed between 2 and 32 kHz. **B.** RSS stimuli, containing energy between 2 and 32 kHz. For RSS, the frequency profile varied between stimuli, making it difficult to discriminate stimuli on the basis of monaural spectral cues. **C**. Psychometric functions for rear locations, stimuli were 2–32 kHz Gaussian noise. Note that 30° and 37.5° were not tested in those animals.

Table 1
Session and training information for Figs. 2–7. Untested animal and stimulus are labeled with "n/a". Note that some animals have been trained to discriminate both spatial and nonspatial stimuli before testing the six stimuli used in this experiment. There is usually only one session per day, but the animal is not trained every day.

Subject	Condition/Sound stimuli	# sessions averaged for computing MAA	# sessions already trained	months days since 1st session
M3T	1, Gaussian noise, 2–32 kHz, front	3	439	4mo 16d
M94W	1, Gaussian noise, 2–32 kHz, front	4	172	8mo 18d
M71V	1, Gaussian noise, 2–32 kHz, front	3	84	2mo 19d
M76X	1, Gaussian noise, 2–32 kHz, front	5	98	14mo 20d
M63W	1, Gaussian noise, 2–32 kHz, front	4	75	14mo 15d
M3T	2, Random spectral shape, 2–32 kHz, front	3	441	4mo 16d
M94W	2, Random spectral shape, 2–32 kHz, front	n/a	n/a	n/a
M71V	2, Random spectral shape, 2–32 kHz, front	n/a	n/a	n/a
M76X	2, Random spectral shape, 2–32 kHz, front	3	126	18mo 29d
M63W	2, Random spectral shape, 2–32 kHz, front	4	115	18mo 21d
M3T	3, Gaussian noise, 2–32 kHz, rear	4	527	6mo 25d
M94W	3, Gaussian noise, 2–32 kHz, rear	3	216	9mo 22d
M71V	3, Gaussian noise, 2–32 kHz, rear	4	121	3mo 22d
M76X	3, Gaussian noise, 2–32 kHz, rear	3	108	15mo 12d
M63W	3, Gaussian noise, 2–32 kHz, rear	n/a	n/a	n/a
M3T	4, Gaussian noise, 2–32 kHz, elevation	3	718	10mo 22d
M94W	4, Gaussian noise, 2–32 kHz, elevation	3	298	14mo 6d
M71V	4, Gaussian noise, 2–32 kHz, elevation	3	181	8mo
M76X	4, Gaussian noise, 2–32 kHz, elevation	n/a	n/a	n/a
M63W	4, Gaussian noise, 2–32 kHz, elevation	n/a	n/a	n/a
M3T	5, Gaussian noise, 4–26 kHz, elevation	3	1219	13mo 1d
M94W	5, Gaussian noise, 4–26 kHz, elevation	3	339	16mo 3d
M71V	5, Gaussian noise, 4–26 kHz, elevation	3	209	9mo 25d
M76X	5, Gaussian noise, 4–26 kHz, elevation	3	81	13mo 24d
M63W	5, Gaussian noise, 4–26 kHz, elevation	3	56	11mo 5d
M3T	6, Gaussian noise, 4–12 kHz, elevation	2	1459	14mo 23d
M94W	6, Gaussian noise, 4–12 kHz, elevation	4	371	17mo 11d
M71V	6, Gaussian noise, 4–12 kHz, elevation	4	251	11mo 12d
M76X	6, Gaussian noise, 4–12 kHz, elevation	n/a	n/a	n/a
M63W	6, Gaussian noise, 4–12 kHz, elevation	n/a	n/a	n/a

In practice, the likelihood ratio is log transformed in case the ratio is very small.

$$TLR = -2 \times \log_e \Delta$$

Furthermore, the TLR from simulations is asymptotically distributed as a chi-square (χ^2) distribution, and the degree of freedom in chi-square distribution is the difference between the number of free parameters of two models. In the Figs. 4A, 7A, the degree of freedom is two for the left panels and is one for the middle and right panels. We simulated the TLR 1000 times (10,000 simulations generate similar results). The TLR from experiments is fixed and is used to compute the p-values. Specifically, the p-value of simulations measure the ratio of the number of simulations with TLR larger than experimental TLR to 1000. The p-value of chi-square equals the upper tail of chi-square cumulative distribution. In MATLAB, it equals to 1 - chi2cdf (experimental TLR, degree of freedom).

In the Palamedes toolbox, the Bayesian estimation is based on the Markov Chain Monte Carlo (MCMC) (Prins and Kingdom, 2018) instead of numerical integration method that is used in the psignifit toolbox. To compare the difference among two conditions, we reparametrized the MAA (α) parameter such that we could directly evaluate the difference between the MAAs. We fitted a fuller model that allows the MAA to differ and a lesser model that have same MAA (difference is zero). Therefore, the two MAA parameters have been changed to two new parameters, one corresponding to the sum of the MAA, one corresponding to the difference of the MAA. This program did this with a model matrix $M = [1\ 1;\ 1\ -1]$. Each row defines a new parameter as a linear combination of the MAAs.

The first row [1 1] defines a new parameter as the sum of the MAAs: sum/effect 1 = (1) * (MAA in condition 1) + (1) * (MAA in condition 2). The second row [1 −1] defines a second parameter as the difference between the MAAs: difference/effect 2 = (1) * (MAA in condition 1) + (−1) * (MAA in condition 2).

The program will generate a posterior distribution of effects 1 and 2. If the difference between the two conditions is significant, then the 95% confidence interval should exclude zero. One advantage of Bayesian inference method over likelihood ratio test method is that it does not require two compared conditions to have the same number of observations (i.e., angles), so that it could be used to compare the parameters between any two conditions.

Raw data from this study is available in Table 2. Data analysis codes are modified from the previous mentioned toolboxes. Codes for generating Figs. 2–7 are customized. We have uploaded all data and codes (https://github.com/ccg1988/marmoset_MAA).

3. Results

3.1. Horizontal acuity

The ability of marmosets to discriminate sound source azimuth of Gaussian noise stimuli is illustrated in Fig. 2A. The five animals tested showed generally good agreement in thresholds, with a median MAA of 13.6° Most of the variation among the subjects occurred at 15° separation; a majority of animals could reliably discriminate 22.5° separation, yet none could do so for 7.5° separation (Fig. 2A, left). In addition to using the hit rate that has been corrected with the false alarm rate, we have also used the uncorrected hit rate and fitted the psychometric function with the Bayesian estimation method (Fig. 2A, right; Fig. 3A). Bayesian inference requires a prior that represents beliefs about the MAA (gray dashed lines, Fig. 3A–C). Here, we assumed that MAA was within the range of 0° to 90° with decreasing probability at 7.5° and 90° Although we never knew the real MAA in each animal, observing the data could reduce our uncertainty and give us a

Table 2
Raw data for Figs. 2–7. Units of the first row stands for the degree. Values of columns 3 to 10 and 19 are the total hits (true positive) and false alarms (false positive) of multiple sessions, respectively. Values of columns 11 to 18 and 20 are the total target trials and catch trials of multiple sessions, respectively. Sub, subject; Con, condition; Ta, target; FA, false alarm; SM, sham/catch trials. n/a, not tested animal nor stimulus. Conditions 1–3, raw data of Figs. 2–4; conditions 4–6, raw data of Figs. 5–7. Only a session with a false alarm rate of less than 0.3 (i.e., FA/SM<0.3) is included.

Sub	Con	Hit 7.5	Hit 15	Hit 22.5	Hit 30	Hit 37.5	Hit 45	Hit 90	Hit 180	Ta 7.5	Ta 15	Ta 22.5	Ta 30	Ta 37.5	Ta 45	Ta 90	Ta 180	FA	SM
M3T	1	9	15	22	24	23	19	26	24	30	30	30	30	30	30	30	30	5	60
M94W	1	6	17	21	22	26	22	23	26	27	27	27	27	27	27	27	27	6	56
M71V	1	8	16	12	19	17	21	19	19	22	22	22	22	22	22	22	22	8	46
M76X	1	5	14	28	34	40	43	47	n/a	47	47	47	47	47	47	47	n/a	12	94
M63W	1	8	19	15	20	19	19	20	n/a	20	20	20	20	20	20	20	n/a	9	39
M3T	2	5	14	10	13	15	15	19	11	21	21	21	21	21	21	21	21	5	48
M94W	2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M71V	2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M76X	2	10	22	28	26	29	29	30	n/a	30	30	30	30	30	30	30	n/a	13	59
M63W	2	13	28	33	36	38	39	40	n/a	40	40	40	40	40	40	40	n/a	23	120
M3T	3	4	10	15	n/a	n/a	25	26	n/a	28	28	28	n/a	n/a	28	28	n/a	10	84
M94W	3	3	11	12	n/a	n/a	22	23	n/a	27	27	27	n/a	n/a	27	27	n/a	11	83
M71V	3	10	18	27	n/a	n/a	28	29	n/a	32	32	32	n/a	n/a	32	32	n/a	14	95
M76X	3	11	9	18	n/a	n/a	22	25	n/a	25	25	25	n/a	n/a	25	25	n/a	8	50
M63W	3	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M3T	4	8	10	20	20	22	22	21	23	24	24	24	24	24	24	24	24	34	142
M94W	4	n/a	12	22	19	31	29	30	32	n/a	32	32	32	32	32	32	32	16	95
M71V	4	9	17	22	22	21	23	24	24	24	24	24	24	24	24	24	24	23	144
M76X	4	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M63W	4	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M3T	5	11	6	10	20	21	21	24	25	25	25	25	25	25	25	25	25	29	149
M94W	5	9	10	24	24	25	22	27	28	29	29	29	29	29	29	29	29	19	116
M71V	5	3	8	12	11	12	9	14	13	16	16	16	16	16	16	16	16	18	103
M76X	5	6	14	14	26	22	25	26	25	29	29	29	29	29	29	29	29	3	58
M63W	5	n/a	6	11	20	22	23	23	21	n/a	24	24	24	24	24	24	24	5	49
M3T	6	n/a	n/a	n/a	n/a	4	5	9	11	n/a	n/a	n/a	n/a	13	13	13	13	4	28
M94W	6	n/a	n/a	n/a	n/a	5	1	11	24	n/a	n/a	n/a	n/a	26	26	26	26	5	48
M71V	6	n/a	n/a	n/a	n/a	5	8	8	18	n/a	n/a	n/a	n/a	22	22	22	22	6	71
M76X	6	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M63W	6	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

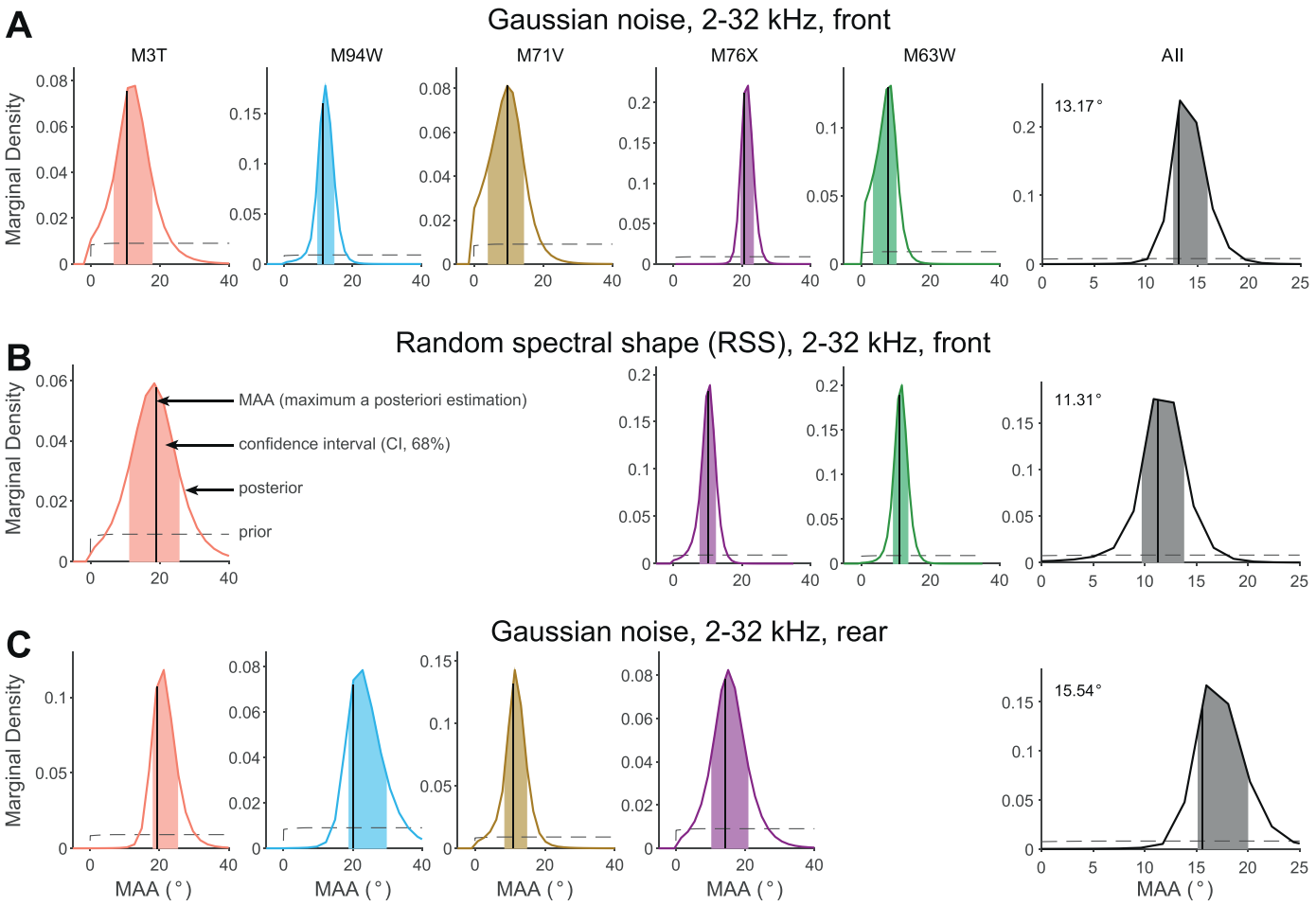

Fig. 3. Prior, posterior, confidence interval, and MAAs among three stimuli conditions for horizontal location discrimination. Different marmosets are represented by different colors in columns 1–5, and column 6 represents data from all marmosets. MAAs in the right column of Fig. 2 are estimated here (vertical line) use the posterior of Bayesian estimation. A-C, each row represents one stimulus condition as shown in Fig. 2. Figures are generated using psignifit toolbox using the data from Table 2. This program assumes that the threshold/MAA is within the range of the data (7.5° to 90°) and with decreasing probability up to half the range (41.25°) above or below the measured data (−33.75° to 131.25°) (setBorders function). It will automatically narrow the borders when the likelihood is near to zero to save the computing power (moveBorders function). In our experiment, MAA is close to lower border but far away from higher border. Therefore, the lower border is always −33.75° and higher border is always less than 90° Here, we modified the program so that the lower border is fixed at 0° (getStandardPriors function). This modification has minor or no effect over MAA estimation in Figs. 2 and 5.

posterior distribution and confidence interval (CI) about the real MAA. The performance of Bayesian estimation could be examined with three metrics: 1) the distance between individual observations (color spots) and fitted psychometric functions (color dashed line) (Fig. 2A); 2) the width and peak amplitude of the posterior distribution (Fig. 3A); 3) whether the posterior and CI fell inside the stimulus range (Fig. 3A). For example, M76X and M94W were well fitted because the fitted curves almost touched all seven dots (Fig. 2A), the posterior distributions were narrow and had large peaks, and their posteriors and CIs were within the stimulus range (Fig. 3A). In contrast, M63W was under fitted since its 68% CI fell outside the smallest test angle (7.5°, Fig. 3A). Therefore, the estimated MAA (7.86°, Fig. 2A) was biased towards 7.5° The median MAA among five animals was 10.51° (Fig. 2A). To get a more accurate and less variable estimate of MAA, we summarized the data (2nd to 6th row, Table 2) from five animals and got the group MAA which was 13.17° (Fig. 3A). We will use the group MAA estimated from all animals as the threshold in each stimulus condition.

To reduce the possibility that marmosets could be using spectral cues for azimuth discrimination, we tested three marmosets' ability to discriminate the location of random spectral shape (RSS) stimuli which varied spectrally on each stimulus presentation and made it difficult to compare successive stimuli on the basis of frequency spectrum. We hypothesized that if subjects were relying on spectral cues to perform azimuth discrimination, thresholds should be higher when discriminating RSS stimuli. Psychometric functions for RSS stimuli are shown in Figs. 2B and 3B. The median MAA for corrected and uncorrected hit rates was 12.92° and 10.08°, respectively, and MAA estimated from three animals was 11.31° Interestingly, all three MAAs were smaller than those in the previous condition.

To statistically compare the difference between RSS front and Gaussian noise front conditions, we first used the transformed likelihood ratio (TLR, see Methods) test. Calculation of TLR includes six steps: 1) compute the probability of observance of response (hit or not) in each trial and stimulus for two conditions, 2) multiply all probabilities to get likelihoods that are associated with threshold (MAA) and slope in two conditions, 3) estimate the thresholds and slopes that maximize the likelihoods without constrictions (fuller model), and 4) estimate the thresholds and slopes that maximize the likelihoods with constrictions on both threshold and slope, either threshold or slope (lesser model), 5) compute the likelihood ratio which equals the maximum likelihood with unconstrained threshold and slope divided by the maximum likelihood with constrained threshold, slope, or both, 6) transform the likelihood ratio to the logarithm of the likelihood ratio (TLR). A larger TLR in-

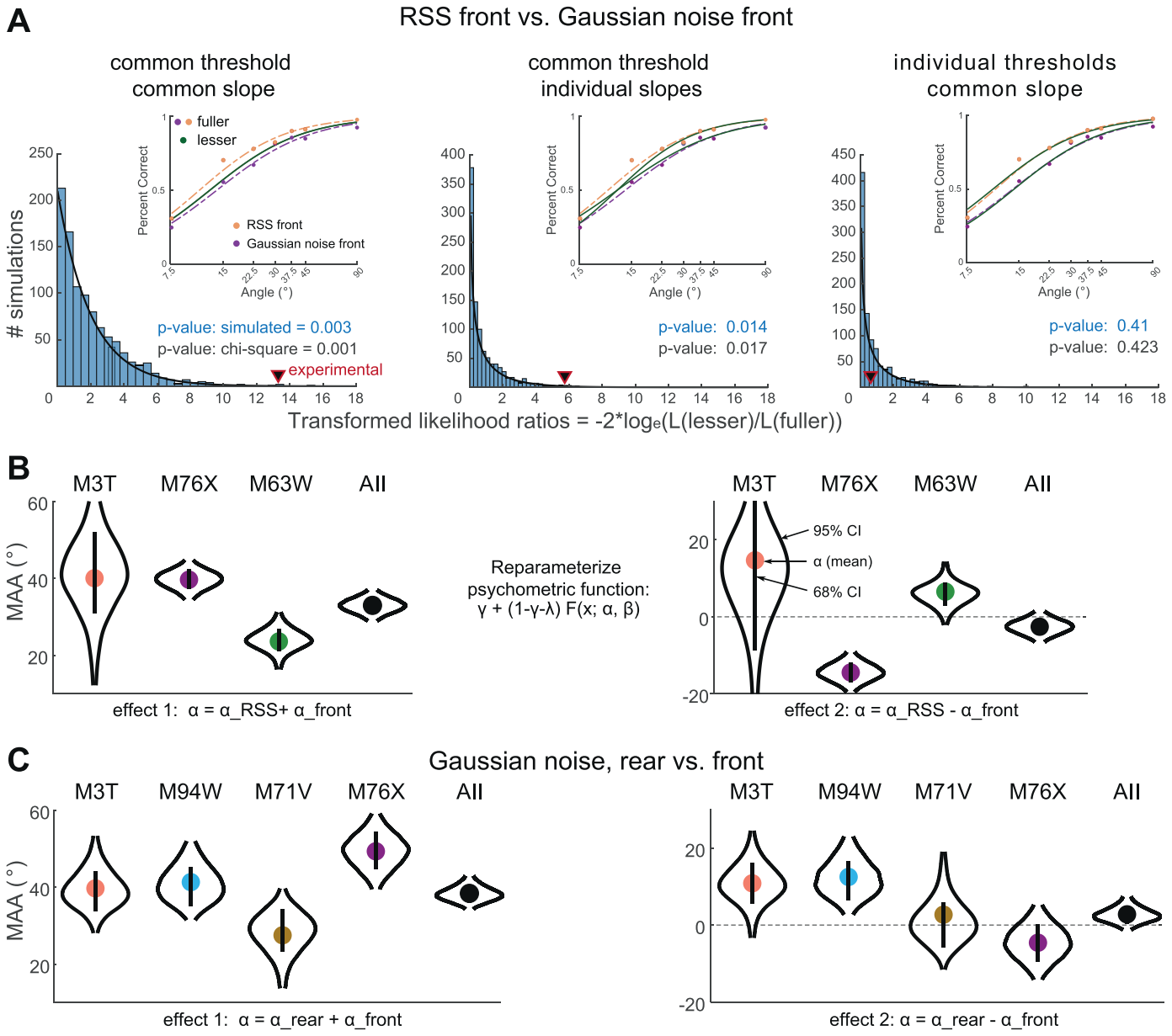

Fig. 4. Comparisons between three stimuli conditions for horizontal location discrimination using transformed likelihood ratio (TLR, using maximum-likelihood estimation) and Bayesian inference. **A**, Model comparisons based on the TLR of maximum-likelihood estimation using data from all marmosets (insets). The fuller models (violet and brown lines and dots) are same across three panels, whereas the lesser/nested models (green curves) are different. The thresholds (i.e., MAA) and slopes are unconstrained in the fuller models. In contrast, the threshold and slope are both constrained in the left panel, only threshold is constrained in the middle panel, and only slope is constrained in the right panel. Blue histogram shows the distribution of 1000 simulated TLR values, black triangle shows the experimental TLR value, and black line shows the scaled chi-square distribution with 2 (left) or 1 (middle and right) degree of freedom. The p-value of simulations measure the ratio of the number of simulations with TLR larger than experimental values to the number of simulations (1000). In the left, only 3 simulations are larger than experimental values, but there are 410 simulations in the right panel. The p-value of chi-square equals the upper tail of chi-square cumulative distribution. **B**, Comparing the posterior between RSS front stimuli with Gaussian noise front stimuli based on the CI of Bayesian inference. The threshold parameter in the psychometric function has been reparametrized using the sum (effect 1) and difference (effect 2) of the threshold of individual conditions. Effect 2 $>$ 0 indicates MAA of RSS stimuli is larger than MAA of Gaussian noise stimuli. For M76X, the 95% CI is less than 0, suggesting MAA of RSS is smaller than MAA of Gaussian noise. Different marmosets are represented by different colors. **C**, similar to **B** but comparing the Gaussian noise rear with front stimuli. α, threshold or MAA; β, slope; γ; guess rate; λ, lapse rate. All figures were generated using psignifit toolbox using the data from Table 2.

dicated a poor fit of the less model to the fuller model. The TLR was asymptotically distributed as chi-square with degrees of freedom equaled to the difference of the number of free parameters between fuller and lesser model, which was used to compute the chi-square p-value. We also used the Monte Carlo simulations to generate the distribution of 1000 TLRs from the lesser model. The simulated p-value equaled to the ratio of the number of simulations with TLR larger than experimental TLR to the number of simulations (1000). In other words, this value was an estimate of the probability that an experiment characterized by the lesser model would generate data resulting in as high a TLR as the experimental TLR. By convention, if it was less than 0.05, then we concluded that the lesser model does not describe the experiment well.

Fig. 4A showed the distribution of 1000 simulated TLRs (blue histogram), one experimental TLR (black triangle), chi-square distribution (black fitted curve), simulated and chi-square p-values, two fuller models (violet and brown curves and dots), and two lesser models (green curves). When comparing the distributions of TLR among three comparisons, it was obvious that the lesser model was poorly fitted when constraining both threshold and

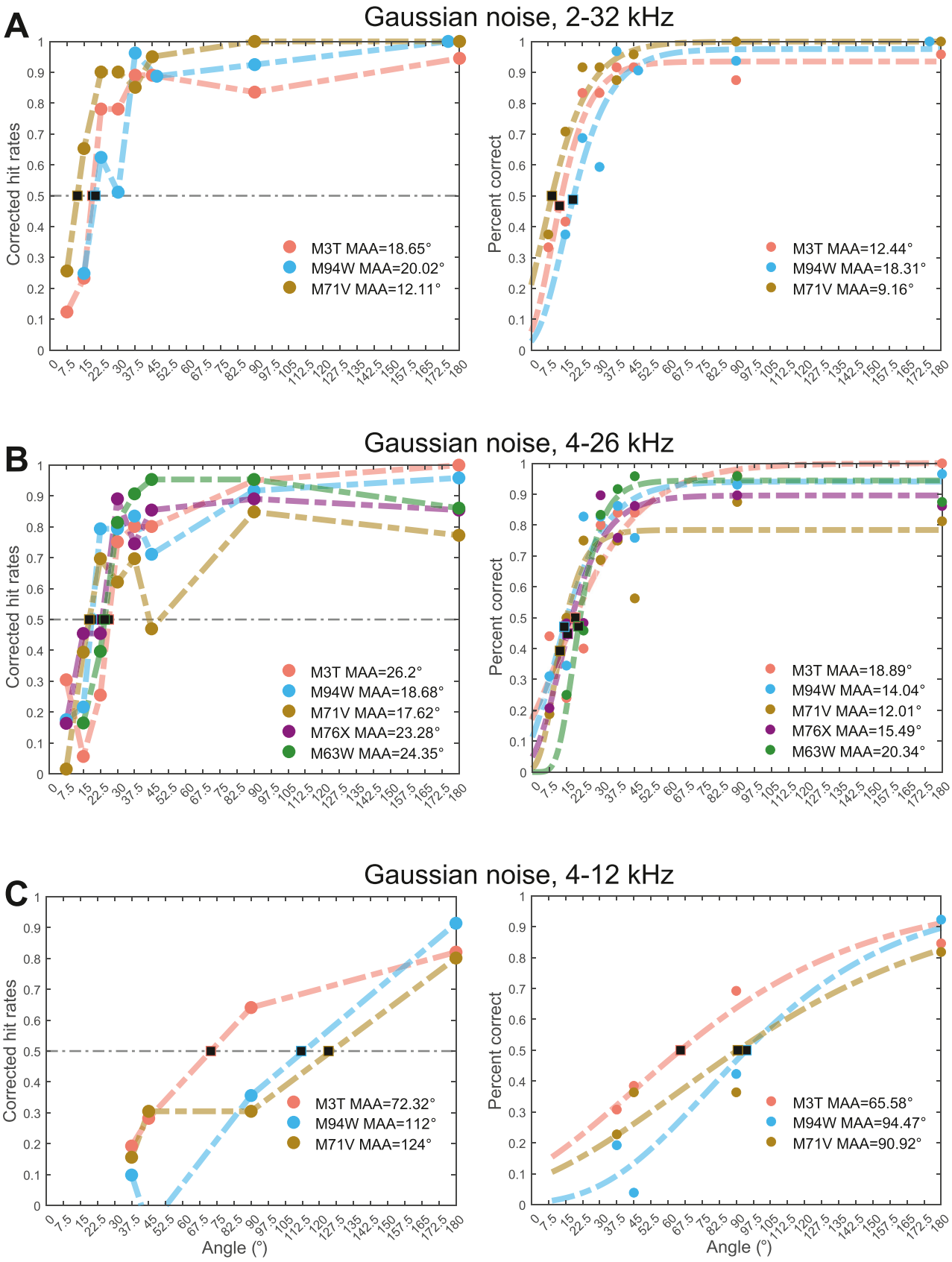

Fig. 5. Psychometric functions for vertical locations discrimination for different stimuli. Format same as Fig. 3. **A.** Gaussian noise, band pass filtered between 2 and 32 kHz. Note that 7.5° was not tested in M94W. **B.** Gaussian noise, filtered between 4 and 26 kHz to exclude acoustic information above the first spectral notch region (12 to 26 kHz) measured previously in marmoset HRTFs. Note that 7.5° was not tested in M94W. **C.** Gaussian noise, filtered between 4 and 12 kHz to exclude acoustic information in the range of the first spectral notch region. M94W had a negative corrected hit rate at 45° This was because the hit rate (HR) was lower than the false alarm rate (FAR) at 45° Note that only sessions with a false alarm rate lower than 30% were included.

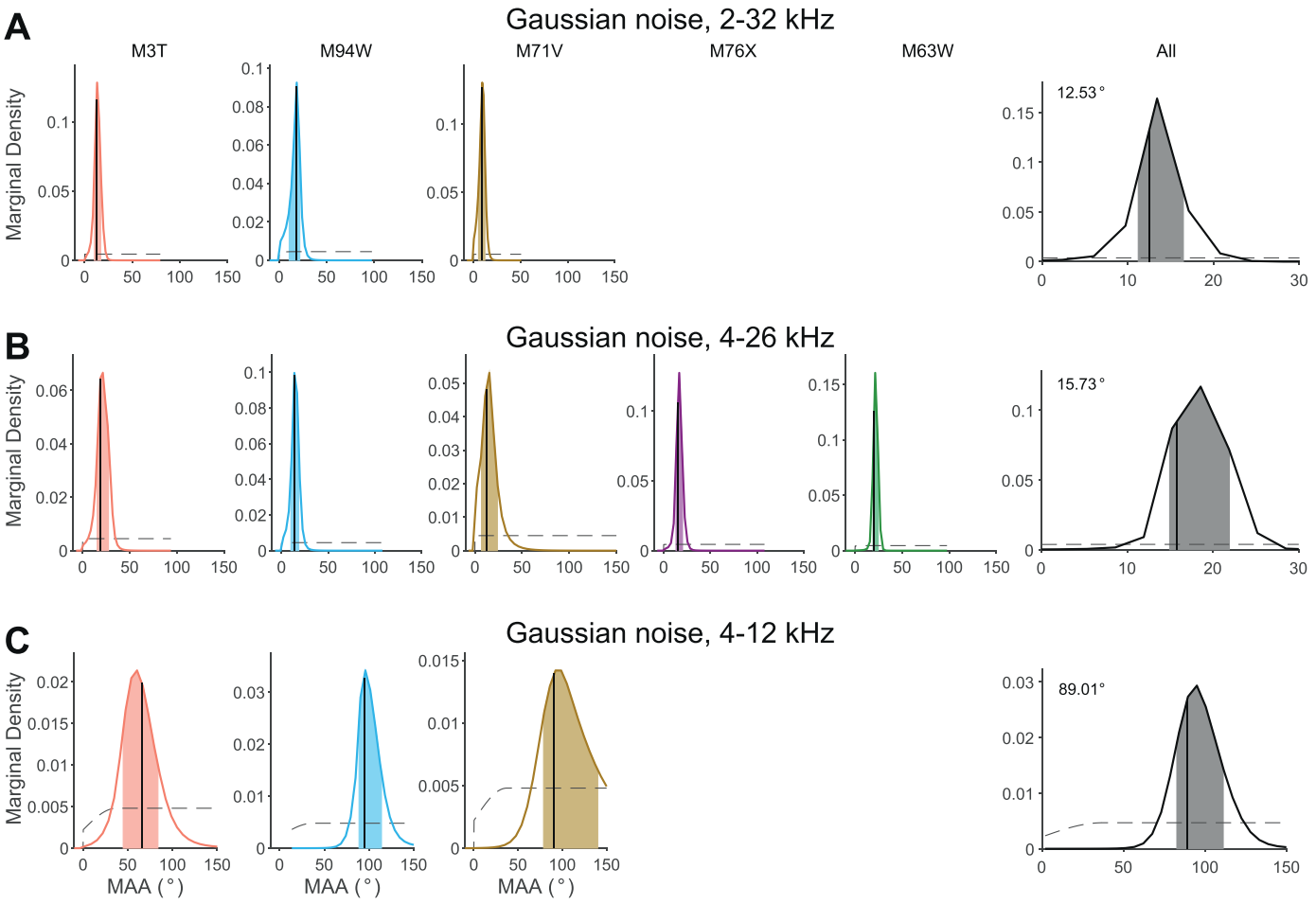

Fig. 6. Prior, posterior, confidence interval, and MAAs among three stimuli conditions for vertical location discrimination. The format is the same as Fig. 3, and the stimuli are the same as Fig. 5.

slope (left vs. middle/right). The simulated p-values also showed that only 3 out of 1000 TLR was larger than experimental TLR when constraining both threshold and slope (left panel). Therefore, we could reject the null hypothesis that the two conditions have the same threshold and slope. This raised the following question: was the threshold different, the slope different, or both? To answer this question, we constrained the threshold or slope while leaving the other parameter unconstrained (middle and right panels). When the threshold was constrained, the lesser model had little overlap with the fuller model (middle panel, inset), and the simulated TLRs were significantly smaller than the experimental TLR ($p = 0.014$). In contrast, when the slope was constrained, the lesser model largely overlapped with the fuller model (right panel, inset), and the simulated and experimental TLRs were similar ($p = 0.41$). In summary, our likelihood ratio test showed that RSS stimuli have significantly smaller MAA than Gaussian noise stimuli.

We have also compared the MAA from two conditions using the previous mentioned Bayesian inference method. Our interest here is not about the specific value of two MAAs, but about whether there is a difference between the MAAs. Therefore, instead of directly comparing the MAA between two conditions, we transformed two MAAs into two effects: effect 1 or fuller model allowed the MAAs to differ and effect 2 or lesser model that have same MAA (difference was zero). In other words, the two MAAs have been changed to two effects, one corresponding to the sum of the MAA, one corresponding to the difference of the MAA. Similar to Fig. 3 which showed the CI of estimated MAA, Fig. 4B showed the CI of effect 1 and 2. If the 95% CI of effect 2 excluded zero, then we could conclude that MAAs between two conditions were significantly different. Furthermore, we also calculated the ratio of MAA in effect 2 to the MAA in effect 1, which indicates the difference ratio. We used difference ratio to measure the relative distance of MAAs between two conditions. A 1° difference between 7° and 8° MAAs is more prominent than between 20° to 21° MAAs. For example, in the M76X, the 95% CI of effect 2 was −19.36° to −9.03° and the difference ratio was −0.36. In the M63W, the 95% CI of effect 2 was −1.67° to 13.78° and the difference ratio was 0.27. When comparing all animals between two conditions, the 95% CI of effect 2 was −9.09° to 0.39°, thus RSS stimuli tend to have smaller MAA with −0.11 difference ratio than Gaussian noise stimuli. Together, monaural spectral cues tend to impair but not improve the azimuth discrimination in marmosets.

Data from previous studies suggest that localization is less accurate at rear locations compared with frontal locations (Oldfield and Parker, 1984; Recanzone and Beckerman, 2004). We tested four marmosets' ability to discriminate azimuth at rear locations using Gaussian noise stimuli. These results are shown in Figs. 2C and 3C. The median MAA for corrected and uncorrected hit rates was 21.79° and 16.8°, respectively, and MAA estimated from four animals was 15.54° Interestingly, all three MAAs were larger than the same stimuli tested in the front location. When comparing the data from all animals, MAA in the rear tend to have larger MAA than in the front (Fig. 4C, −1.09° to 7.52° 95% CI, 0.08 difference ratio).

In summary, the RSS stimuli have the smallest MAA (11.31°), and Gaussian noise stimuli at the rear have the largest MAA

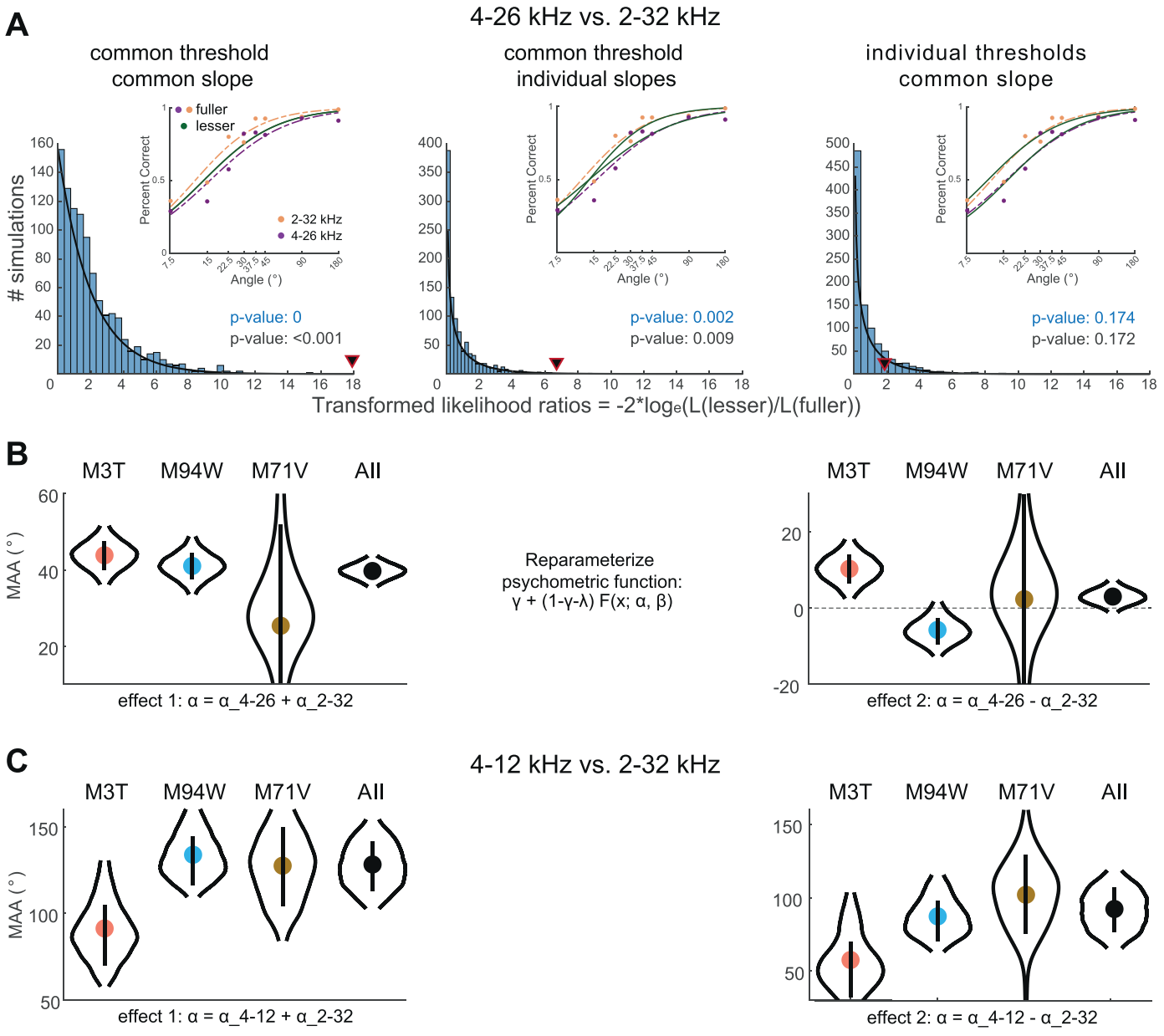

Fig. 7. Comparisons between three stimuli conditions for vertical location discrimination using likelihood ratio test and Bayesian inference. The format is the same as Fig. 4, and the stimuli are the same as Figs. 5, 6.

(15.54°), with Gaussian noise stimuli at the front having the intermediate MAA (13.17°).

3.2. Vertical acuity

While horizontal spatial information is contained in the binaural differences between the two ears, additional information is required to compute sound source location in a 2-dimensional space. To do this, the auditory system makes use of spectral cues generated primarily by the pinnae. These cues in marmosets are located at frequencies higher than 12 kHz (Fig. 1B) (Slee and Young, 2010). We measured vertical location acuity first using the same Gaussian noise stimuli used to test horizontal acuity. Psychometric functions for three animals are shown in Fig. 5A. The median MAA for the corrected hit rate and the uncorrected hit rate was 18.65° and 12.43°, respectively. MAA estimated from three animals was 12.53° (Fig. 6A). Although the first two values were larger than the horizontal acuity measured with the same stimuli (13.6° and 10.51°), the last value was comparable (13.17°). Together, marmosets have similar horizontal and vertical spatial acuity.

To reduce the possibility that marmosets were using high frequency cues, vertical discrimination acuity was measured in five animals using Gaussian noises filtered between 4 and 26 kHz. Psychometric functions for five marmosets discriminating elevation using the mid-frequency Gaussian noise stimuli (4–26 kHz) are shown in Fig. 5B. The median MAA for the corrected hit rate and the uncorrected hit rate was 23.28° and 15.49°, respectively. MAA estimated from five animals was 15.76° (Fig. 6B). All three MAAs were larger than the previous condition tested with 2–32 kHz stimuli.

Similar to Figs. 4A, 7A showed the distribution of simulated and experimental TLRs, chi-square distribution, two p-values, two fuller models, and two lesser models. When constrained both threshold and slope, the TLRs from 1000 simulations were all smaller than the TLR from experiment ($p = 0$, left panel). By constraining either the threshold or slope, we found that the condition difference was

mainly contributed by the threshold ($p = 0.002$, middle panel), although the slope may also play a role ($p = 0.174$, right panel). In summary, our likelihood ratio test showed that 2–32 kHz stimuli have significantly smaller MAA than 4–26 kHz stimuli.

We further compared the two conditions by reparametrizing MAA as the sum and difference of individual MAAs (Fig. 7B). When comparing the data from all animals, stimuli without high frequency cues tended to have larger MAA than stimuli with high frequency cue (−1.04° to 7.18° 95% CI, 0.08 difference ratio). Therefore, the high frequency cues contribute to marmosets' vertical sound localization acuity.

In marmosets, this first notch was shown to vary between 12 and 24 kHz (Slee and Young, 2010). We additionally tested three animals using sounds which were filtered to only include stimulus information below the location of the first notch region (Gaussian noise, 4–12 kHz). Animals tested with this stimulus exhibited extreme difficulty in discriminating vertical location (Figs. 5C and 6C). One animal had a threshold of 72°, while the other two animals could not discriminate any of the locations in the frontal hemifield (112° and 124°). All three animals, however, could reliably discriminate front and rear locations (i.e., MAA < 180°), probably due to the boosted sound amplitude between 7 and 12 kHz at the front location (Slee and Young, 2010). The median MAA for the corrected hit rate and the uncorrected hit rate was 112° and 90.7°, respectively. MAA estimated from three animals was 89.01° Compared with 2–32 kHz stimuli, MAA tested with 4–12 kHz stimuli were significantly larger in all cases (Fig. 7C).

In summary, the 2–32 kHz stimuli have the smallest MAA (12.53°), stimuli without high-frequency region have the intermediate MAA (15.76°), and stimuli without first notch region have the largest MAA (89.01°).

4. Discussion

4.1. Horizontal and vertical sound location discrimination

The MAA for horizontal localization measured in marmosets is higher than many other mammals. Several species have been shown to discriminate sound locations separated by less than 10° in azimuth, including cats (Heffner and Heffner, 1988), macaques (Brown et al., 1980), and opossums (Ravizza and Masterton, 1972). Despite significant training, 10° seemed to be unachievable for azimuth discrimination by marmosets in the current study. Animals are not over-motivated (high hit rate and false-alarm rate), as all sessions have at most 44% hit rate at the smallest test angle, and less than 15% of sessions have false alarm rates larger than 30% and they are excluded (Table 2). Animals are also not under-motivated (low hit rate and false positive rate), as all sessions have at least 80% hit rate at the largest test angle and at least 5% false alarm rate (Table 2). Therefore, animals are in the optimal motivational range with a high hit rate and low false alarm rate (Groblewski et al., 2020; Moore and Kuchibhotla, 2022).

Higher MAA in marmosets than in many other mammals is also unlikely due to the use of sounds filtered above 2 kHz. The upper limit for ITD cues is negatively correlated with the animal's physiological ITD range and is limited by the animal's auditory nerve phase locking ability (Keating et al., 2014). In humans, the upper limit for ITD cues is 1.3 kHz (Klumpp and Eady, 1956), but it is 4.5 kHz in ferrets, which have a similar ITD range to marmosets (Keating et al., 2014). In cats and squirrel monkeys, the fine structure phase locking in the auditory nerve has been observed at frequencies up to 4 kHz (Johnson, 1980; Rose et al., 1967). Because marmosets' ILD cues are relatively small for sound frequencies below 5 kHz (Slee and Young, 2010), marmosets likely use ITD cues to localize sound with the frequency between 2 and 5 kHz in this study. Consistent with this hypothesis, cats and macaques have both been shown to exhibit high horizontal acuity using test stimuli of comparable spectra (Brown et al., 1980; Huang and May, 1996). Also, there is some evidence that several small mammals (tree shrew, rat, gerbil) can use ITD cues at frequencies significantly higher than 2 kHz (Heffner and Masterton, 1980; Masterton et al., 1975). As ITD cues are known to be dependent on the physical distance between the two ears, it is perhaps not surprising that marmosets, a species with small head sizes, do not appear to be expert localizers like cats and macaques. A meta-analysis of data from Multiple species shows a relatively good correlation between head size (as defined by maximum interaural time difference) and horizontal MAA (Fig. 8A) (Brown and May 2005). Placing the marmoset into this dataset shows that the performance measured in the present study is almost exactly what would be predicted based on head size. For future work, it is of great interest to know the frequency range of ITD (and ILD) cues marmoset used for horizontal sound localization.

Although sound localization in the horizontal plane relies primarily on binaural cues, HRTF shape has also been observed to change with azimuth in several species (Rice et al., 1992; Wightman and Kistler, 1989), and monaural spectral cues can be used to perform a horizontal discrimination task (Butler, 1986; Van Wanrooij and Van Opstal, 2004). As this task measured discrimination ability rather than absolute localization accuracy, the usefulness of these spectral cues could be increased. Surprisingly, our results showed that MAA in the RSS condition was lower than MAA measured in the Gaussian noise condition, suggesting that marmosets used primarily binaural cues to discriminate horizontal locations and monaural spectral cues interfere with horizontal localization. One explanation is that spectral cues from two ears backup binaural cues in case only one ear is left, and spatial information conveyed by spectral cues may not be coaligned with binaural cues. Further work is needed to test this hypothesis.

A comparison of marmoset vertical acuity with other tested animals is shown in Fig. 8B. Our results showed that marmosets' MAA for vertical localization were similar to MAA in the horizontal localization. This finding is consistent with several species previously tested, such as cats (Martin and Webster 1987) and macaques (Brown et al., 1982). Vertical acuity in cats (Martin and Webster, 1987) and macaques (Brown et al., 1982) has been shown to be much better (roughly equal to horizontal acuity) than in marmosets. The exceptional vertical discrimination in these species, however, is degraded by the removal of high frequency region (> 18 kHz, Huang and May, 1996; > 2 kHz, Brown et al., 1982). However, we found that high frequency region (> 26 kHz) has limited contribution to vertical acuity. In contrast, the first notch region (12 to 26 kHz) plays a major role in vertical sound discrimination. In marmosets, the frequency and depth of the first notch region change systematically with both azimuth and elevation (Fig. 1A, B). It will be very interesting to examine the 2-dimensional spatial receptive fields of neurons with their best frequency tuned to the first notch region.

4.2. MAA and field of best vision

Many explanations have been proposed to explain the variations in MAA among mammalian species, including whether the animal's head is large or small, a predator or prey, diurnal or nocturnal, or has large or small binocular visual fields (Heffner and Heffner, 2016). As mentioned earlier, MAA is negatively correlated with head size (Fig. 8A), but MAA of horse (25°) and cattle (30°) showing that having a large functional head size did not necessarily produce good sound-localization acuity. Heffner and Heffner (1992) discovered that sound-localization acuity was closely and positively correlated with the size of an animal's field of best vision (Fig. 8C). This observation led to the

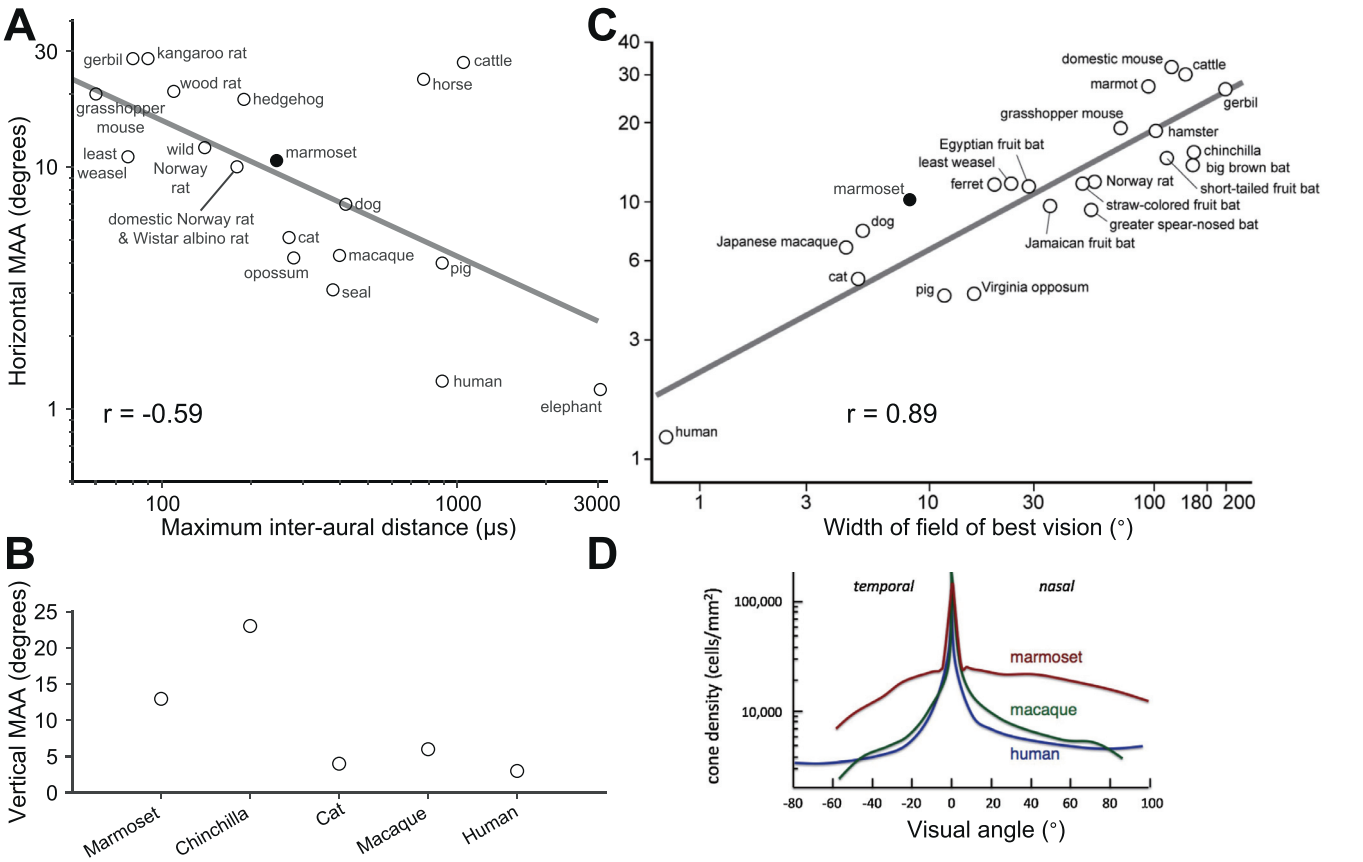

Fig. 8. Comparative sound localization acuity. **A**. Horizontal MAA as a function of head size in 18 mammals, plus the present value for marmosets (250 μs). Marmosets are roughly in line with the trend gathered from previous studies (diagonal line). Figure was reproduced and modified from Brown and May (2005). **B**. Vertical sound localization discrimination thresholds in five mammals. **C**. Horizontal MAA as a function of width of the field of best vision in 22 mammals, plus the present value for marmosets (7.8°, Troilo et al., 1993). Figure was modified from Heffner and Heffner, 2016. Width of field of best vision was defined as the width of the horizontal visual field, in degrees, that subtends the portion of the retina containing ganglion cell densities greater than or equal to 75% of the maximum density. **D**. Cone density as a function of retinal position for three primate species. Figure was copied from Mitchell and Leopold, 2015. Note the r values come from original figures, and we did not recalculate them.

proposal that the primary function of sound localization is to acoustically guide the eyes to the sound source (Heffner and Heffner, 1992). Placing the marmoset into this dataset shows that its MAA is roughly matched but deviates by nearly 4° from the fitted line, suggesting it has a relatively worse localization acuity when considering its field of best vision. One explanation is that the marmoset has a broad visual field, so it does not require higher sound localization acuity to guide the gaze. Indeed, Fig. 8D shows that the marmoset has notably higher cone density in the retinal periphery. The other explanation is marmoset's living environment contains noise and obstacles, which may perturb sound localization cues and affect visual feedback correction of sound localization. In mammals, because marmoset has higher visual acuity than most species and has fovea, which is unique to primates, it would be interesting to study how the visual spatial cues modulate and integrate with auditory spatial cues in marmosets.

4.3. Implication of using marmosets as a model in auditory neuroscience

The marmosets have emerged as a successful model in auditory neuroscience (Wang, 2018). One good example is the studies of harmonic and pitch with both behavior and neurophysiology techniques. Marmosets not only can discriminate complex harmonic tones and perceive pitch (Osmanski et al., 2013; Song et al., 2016) but also have harmonic template neurons and pitch centers in their auditory cortex (Feng and Wang, 2017; Bendor and Wang, 2005). Furthermore, the hearing range of marmoset is also similar to humans (Osmanski and Wang, 2011). Those characters make marmoset a good model for nonspatial auditory research.

In this study, we found that the MAA of the marmoset is similar to the rat and almost ten times larger than humans (Fig. 8A). Does this indicate that the marmoset is only suitable for nonspatial auditory research? The answer is no. One reason is that the MAA is only one metric of sound localization ability and may not be optimal for marmoset. In cats and macaques, their sound localization accuracy improved significantly when animals could use head orienting to report the absolute sound locations (Tollin et al., 2005; Populin, 2006). This may also hold true in the marmosets, which exhibit rapid head movement toward interesting sounds (Slee and Young, 2010; Pandey et al., 2020). The other reason is that even if marmosets are not a sound localization "expert" like barn owls, cats, and macaques, no spatial research was and will be excluded for this reason. For example, neurons in the marmoset caudal auditory cortex have sharper spatial tuning than neurons from the rostral area, which is consistent with cats, macaques, and humans (Zhou and Wang, 2012; Remington and Wang, 2019). Arguably, the marmoset is a good model for spatial hearing research due to its easily accessible and smooth auditory cortex than cats and macaques. This property allows us to leverage the optical imaging method to investigate cortical spatial processing with high spatial resolution, for example, mapping the auditory space.

Declaration of Competing Interest

No conflicts of interest are declared by the authors.

CRediT authorship contribution statement

Chenggang Chen: Conceptualization, Methodology, Software, Data curation, Visualization, Writing – review & editing. **Evan D. Remington:** Conceptualization, Methodology, Software, Investigation, Data curation, Visualization, Writing – original draft. **Xiaoqin Wang:** Conceptualization, Funding acquisition, Resources, Writing – review & editing, Supervision.

Data Availability

We have uploaded all data and codes (https://github.com/ccg1988/marmoset_MAA).

Acknowledgements

We thank J. Estes and N. Sotuyo for assistance with animal care, C. Baldwin, J. Green, M. Gu, and T. Saavedra for generous assistance in performing some of the experiments, and M. Osmanski for his comments on the manuscript.

Funding

This work was supported by National Institute of Deafness and Other Communications Disorders grants DC003180 and DC005808 (X.W.) and The Kavli Foundation grant GD2651 (C.C.).

References

Barbour, D.L., Wang, X., 2003. Contrast tuning in auditory cortex. Science 299, 1073–1075. doi:10.1126/science.1080425.

Bartlett, E.L., Wang, X., 2007. Neural representations of temporally modulated signals in the auditory thalamus of awake primates. J. Neurophysiol. 97, 1005–1017. doi:10.1152/jn.00593.2006.

Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165. doi:10.1038/nature03867.

Blauert, J., 1997. Spatial Hearing: The Psychophysics of Human Sound Localization. MIT Press Rev. Edition.

Brown, C.H., Beecher, M.D., Moody, D.B., Stebbins, W.C., 1980. Localization of noise bands by Old World monkeys. J. Acoust. Soc. Am. 68, 127–132. doi:10.1121/1.384638.

Brown, C.H., May, B.J., 2005. Comparative Mammalian Sound Localization. Springer.

Brown, C.H., Schessler, T., Moody, D., Stebbins, W., 1982. Vertical and horizontal sound localization in primates. J. Acoust. Soc. Am. 72, 1804–1811. doi:10.1121/1.388653.

Butler, R.A., 1986. The bandwidth effect on monaural and binaural localization. Hear. Res. 21, 67–73. doi:10.1016/0378-5955(86)90047-x.

Chen, C., Xu, S., Wang, Y., and Wang, X.. (2022). Location-specific facili in primate auditory cortex. Preprint at 10.1101/2022.06.19.496736.

Davis, Z.W., Muller, L., Martinez-Trujillo, J., Sejnowski, T., Reynolds, J.H., 2020. Spontaneous travelling cortical waves gate perception in behaving primates. Nature 587, 432–436. doi:10.1038/s41586-020-2802-y.

de la Mothe, L.A., Blumell, S., Kajikawa, Y., Hackett, T.A., 2012. Cortical connections of auditory cortex in marmoset monkeys: lateral belt and parabelt regions. Anat. Rec. 295, 800–821. doi:10.1002/ar.22451, (Hoboken).

Eliades, S.J., Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453, 1102–1106. doi:10.1038/nature06910.

Feng, L., Wang, X., 2017. Harmonic template neurons in primate auditory cortex underlying complex sound processing. Proc. Natl. Acad. Sci. USA 114, E840–E848.

Gao, L., Kostlan, K., Wang, Y., Wang, X., 2016. Distinct subthreshold mechanisms underlying rate-coding principles in primate auditory cortex. Neuron 91, 905–919. doi:10.1016/j.neuron.2016.07.004.

Gescheider, G.A. (1985). Psychophysics : method, theory, and application, 2nd ed. (L. Erlbaum Associates).

Groblewski, P.A., Ollerenshaw, D.R., Kiggins, J.T., Garrett, M.E., Mochizuki, C., Casal, L., Cross, S., Mace, K., Swapp, J., Manavi, S., Williams, D., Mihalas, S., Olsen, S.R., 2020. Characterization of learning, motivation, and visual perception in five transgenic mouse lines expressing GCaMP in distinct cell populations. Front Behav. Neurosci. 14, 104. doi:10.3389/fnbeh.2020.00104.

Hage, S.R., 2020. The role of auditory feedback on vocal pattern generation in marmoset monkeys. Curr. Opin. Neurobiol. 60, 92–98. doi:10.1016/j.conb.2019.10.011.

Heffner, H.E., Heffner, R.S., 2016. The evolution of mammalian sound localization. Acoust. Today 12 (1).

Heffner, H., Masterton, B., 1980. H in Glires: domestic rabbit, cotton rat, feral house mouse, and kangaroo rat. J. Acoust. Soc. Am. 68, 127–132. doi:10.1121/1.384638.

Heffner, R.S., Heffner, H.E., 1988. Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure. Hear. Res. 36, 221–232. doi:10.1016/0378-5955(88)90064-0.

Heffner, R.S., Heffner, H.E., 1992. Visual factors in sound localization in mammals. J. Comp. Neurol. 317, 219–232. doi:10.1002/cne.903170302.

Huang, A.Y., May, B.J., 1996. Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. J. Acoust. Soc. Am. 100, 1070–1080. doi:10.1121/1.416293.

Johnson, D., 1980. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J. Acoust. Soc. Am. 68, 127–132. doi:10.1121/1.384638.

Johnson, L.A.,, Della Santina, C.C., Wang, X., 2012. Temporal bone characterization and cochlear implant feasibility in the common marmoset (Callithrix jacchus). Hear. Res. 290. doi:10.1016/j.heares.2012.05.002.

Johnson, L.A., Della Santina, C.C., Wang, X., 2016. Selective neuronal activation by cochlear implant stimulation in auditory cortex of awake primate. J. Neurosci. 36, 12468–12484. doi:10.1523/JNEUROSCI.1699-16.2016.

Johnson, L.A., Della Santina, C.C., Wang, X., 2017. Representations of time-varying cochlear implant stimulation in auditory cortex of awake marmosets (callithrix jacchus). J. Neurosci. 37, 7008–7022. doi:10.1523/JNEUROSCI.0093-17.2017.

Jovanovic, V., Fishbein, A.R., de la Mothe, L., Lee, K.F., Miller, C.T., 2022. Behavioral context affects social signal representations within single primate prefrontal cortex neurons. Neuron 110, 1318–1326. doi:10.1016/j.neuron.2022.01.020, e1314.

Keating, P., Nodal, F.R., King, A.J., 2014. Behavioural sensitivity to binaural spatial cues in ferrets: evidence for plasticity in the duplex theory of sound localization. Eur. J. Neurosci. 39 (2), 197–206.

Kajikawa, Y., de la Mothe, L.A., Blumell, S., Ste -D'Angelo, S.J., D'Angelo, W., Camalier, C.R., Hackett, T.A, 2008. Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex. Hear. Res. 239, 107–125. doi:10.1016/j.heares.2008.01.015.

Klumpp, R.G., Eady, H.R., 1956. Some measurements of interaural time difference thresholds. J. Acoust. Soc. Am. 28, 859–860. doi:10.1121/1.3652859.

Liao, D.A., Zhang, Y.S., Cai, L.X., Ghazanfara, A.A., 2018. Internal states and extrinsic factors both determine monkey vocal e ction. Proc. Natl. Acad. Sci. U. S. A. 115 (15).

Liu, X.P., Wang, X., 2022. Distinct neuronal types contribute to hybrid temporal encoding strategies in primate auditory cortex. PLoS Biol. 20, e3001642. doi:10.1371/journal.pbio.3001642.

Lu, T., Liang, L., Wang, X., 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat. Neurosci. 4, 1131–1138. doi:10.1038/nn737.

Lui, L.L., Mokri, Y., Reser, D.H., Rosa, M.G., Rajan, R., 2015. Responses of neurons in the marmoset primary auditory cortex to interaural level differences: comparison of pure tones and vocalizations. Front. Neurosci. 9, 132. doi:10.3389/fnins.2015.00132.

Martin, R.L., Webster, W.R., 1987. The auditory spatial acuity of the domestic cat in the interaural horizontal and median vertical planes. Hear. Res. 30, 239–252. doi:10.1016/0378-5955(87)90140-7.

Masterton, B., Thompson, G.C., Bechtold, J.K., RoBards, M.J., 1975. Neuroanatomical basis of binaural phase-difference analysis for sound localization: a comparative study. J. Comp. Physiol. Psychol. 89, 379–386. doi:10.1037/h0077034.

Mitchell, J.F., Leopold, D.A., 2015. The marmoset monkey as a model for visual neuroscience. Neurosci. Res. 20–46.

Mitchell, J.F., Reynolds, J.H., Miller 2014. Active vision in marmosets: a model system for visual neuroscience. J. Neurosci. 34, 1183–1194. doi:10.1523/JNEUROSCI.3899-13.2014.

Moore, S., Kuchibhotla, K.V., 2022. Slow or sudden: re-interpreting the learning curve for modern systems neurosci nce IBRO Neurosci. Rep. 13, 9–14.

Oldfield, S.R., Parker, S.P., 1984. Acuity nd localisation: a topography of auditory space. I. Normal hearing conditions. Perception 13, 581–600. doi:10.1068/p130581.

Osmanski, M.S., Song, X., Guo, Y., Wang, X., 2016. Frequency discrimination in the common marmoset (Callithrix jacchus). Hear. Res. 341, 1–8. doi:10.1016/j.heares.2016.07.006.

Osmanski, M.S., Song, X., Wang, X., 2013. The role of harmonic resolvability in pitch perception in a vocal nonhuman primate, the common marmoset (Callithrix jacchus). J. Neurosci. 33, 9161–9168. doi:10.1523/JNEUROSCI.0066-13.2013.

Osmanski, M.S., Wang, X., 2011. Measurement of absolute auditory thresholds in the common marmoset (Callithrix jacchus). Hear. Res. 277, 127–133. doi:10.1016/j.heares.2011.02.001.

Pandey, S., Simhadri, S., Zhou, Y., 2020. Rapid head movements in common marmoset monkeys. iScience 23 (2), 100837. doi:10.1016/j.isci.2020.100837.

Park, J.E., Zhang, X.F., Choi, S.H., Okahara, J., Sasaki, E., Silva, A.C., 2016. Generation of transgenic marmosets expressing genetically encoded calcium indicators. Sci. Rep. 6, 34931. doi:10.1038/srep34931.

Populin, L.C., 2006. Monkey sound localization: head-restrained versus head-unrestrained orienting. J. Neurosci. 26 (38), 9820–9832. doi:10.1523/JNEUROSCI.3061-06.2006.

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. doi:10.3389/fpsyg.2018.01250.

Ravizza, R.J., Masterton, B., 1972. Contribution of neocortex to sound localization in opossum (Didelphis virginiana). J. Neurophysiol. 35, 344–356. doi:10.1152/jn.1972.35.3.344.

Recanzone, G.H., Beckerman, N.S., 2004. Effects of intensity and location on sound location discrimination in macaque monkeys. Hear. Res. 198, 116–124. doi:10.1016/j.heares.2004.07.017 .

Remington, E.D., Osmanski, M.S., Wang, X., 2012. An operant conditioning method for studying auditory behaviors in marmoset monkeys. PLoS ONE 7, e47895. doi:10.1371/journal.pone.0047895.

Remington, E.D., Wang, X., 2019. Neural representations of the full spatial field in auditory cortex of awake marmoset (*Callithrix jacchus*). Cereb. Cortex 29, 1199–1216. doi:10.1093/cercor/bhy025.

Reser, D.H., Burman, K.J., Richardson, K.E., Spitzer, M.W., Rosa, M.G., 2009. Connections of the marmoset rostrotemporal auditory area: express pathways for analysis of affective content in hearing. Eur. J. Neurosci. 30, 578–592. doi:10.1111/j.1460-9568.2009.06846.x.

Rice, J.J., May, B.J., Spirou, G.A., Young, E.D., 1992. Pinna-based spectral cues for sound localization in cat. Hear. Res. 58, 132–152. doi:10.1016/0378-5955(92)90123-5.

Rose, J.E., Brugge, J.F., Anderson, D.J., Hind, J.E., 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J. Neurophysiol. 30, 769–793. doi:10.1152/jn.1967.30.4.769.

Roy, S., Zhao, L., Wang, X., 2016. Distinct neural Activities in premotor cortex during natural vocal behaviors in a new world primate, the common marmoset (*Callithrix jacchus*). J. Neurosci. 36, 12168–12179. doi:10.1523/JNEUROSCI.1646-16.2016.

Sadagopan, S., Wang, X., 2008. Level invariant representation of sounds by populations of neurons in primary auditory cortex. J. Neurosci. 28, 3415–3426. doi:10.1523/JNEUROSCI.2743-07.2008.

Sasaki, E., Suemizu, H., Shimada, A., Hanazawa, K., Oiwa, R., Kamioka, M., Tomioka, I., Sotomaru, Y., Hirakawa, R., Eto, T., et al., 2009. Generation of transgenic non-human primates with germline transmission. Nature 459, 523–527. doi:10.1038/nature08090.

Schutt, H.H., Harmeling, S., Macke, J.H., Wichmann, F.A., 2016. Painfree and accurate Bayesian estimation of psychometric functions for (p[illegible]ally) overdispersed data. Vision Res. 122, 105–123.

Slee, S.J., Young, E.D., 2010. Sound localization cues in the marmoset monkey. Hear. Res. 260, 96–108. doi:10.1016/j.heares.2009.12.001.

Snowdon, C.T., Hodun, A., 1981. Acoustic adaptations in pygmy marmoset contact calls: locational cues vary with distances between conspecifics. Behav. Ecol. Sociobiol. 9 (4), 295–300 (Print).

Song, X., Guo, Y., Chen, C., Wang, X., 2022a. A silent two-photon imaging system for studying in vivo auditory neuronal functions. Light Sci. Appl. 11, 96. doi:10.1038/s41377-022-00783-y.

Song, X., Guo, Y., Li, H., Chen, C., Lee, J.H., Zhang, Y., Schmidt, Z., Wang, X., 2022b. Mesoscopic landscape of cortical functions revealed by through-skull wide-field optical imaging in marmoset monkeys. Nat. Commun. 13, 2238. doi:10.1038/s41467-022-29864-7.

Song, X., Osmanski, M.S., Guo, Y., Wang, X., 2016. Complex pitch perception mechanisms are shared by humans and a New World monkey. Proc. Natl. Acad. Sci. U. S. A. 113, 781–786. doi:10.1073/pnas.1516120113.

Tollin, D.J., Populin, L.C., Moore, J.M., Ruhland, J.L., Yin, T.C., 2005. Sound-localization performance in the cat: the effect of restraining the head. J. Neurophysiol. 93 (3), 1223–1234. doi:10.1152/jn.00747.2004.

Van Wanrooij, M.M., Van Opstal, A.J., 2004. Contribution of head shadow and pinna cues to chronic monaural sound localization. J. Neurosci. 24, 4163–4171. doi:10.1523/JNEUROSCI.0048-04.2004.

Wang, X., 2018. Cortical coding of auditory features. Annu. Rev. Neurosci. 41, 527–552. doi:10.1146/annurev-neuro-072116-031302.

Wang, X., Zhang, Y., Bai, S., Qi, R., Sun, H., Li, R., Zhu, L., Cao, X., Jia, G., Li, X., Gao, L., 2022. Corticofugal modulation of temporal and rate representations in the inferior colliculus of the awake marmoset. Cereb. Cortex doi:10.1093/cercor/bhab467.

Wightman, F.L., Kistler, D.J., 1989. Headphone simulation of free-field listening. I: stimulus synthesis. J. Acoust. Soc. Am. 85, 858–867. doi:10.1121/1.397557.

Zeng, H.H., Huang, J.F., Li, J.R., Shen, Z., Gong, N., Wen, Y.Q., Wang, L., Poo, M.M., 2021. Distinct neuron populations for simple and compound calls in the primary auditory cortex of awake marmosets. Natl. Sci. Rev. 8, nwab126. doi:10.1093/nsr/nwab126.

Zhou, Y., Wang, X., 2010. Cortical processing of dynamic sound envelope transitions. J. Neurosci. 30, 16741–16754. doi:10.1523/JNEUROSCI.2016-10.2010.

Zhou, Y., Wang, X., 2012. Level dependence of spatial processing in the primate auditory cortex. J. Neurophysiol. 108, 810–826. doi:10.1152/jn.00500.2011.

Zhou, Y., Wang, X., 2014. Spatially extended forward suppression in primate auditory cortex. Eur. J. Neurosci. 39, 919–933. doi:10.1111/ejn.12460.

Zhu, S., Allitt, B., Samuel, A., Lui, L., Rosa, M.G.P., Rajan, R., 2019. Sensitivity to vocalization pitch in the caudal auditory cortex of the marmoset: comparison of core and belt areas. Front. Syst. Neurosci. 13, 5. doi:10.3389/fnsys.2019.00005.