Three-dimensional sound localization of nearby sources in echoic rooms

Norbert Kopčo[1], Virginia Best[2], Gabriela Andrejková[1], and Barbara Shinn-Cunningham[3]

[1] P. J. Šafárik University in Košice, Slovakia

[2] Boston University

[3] Carnegie Mellon University

[Work supported by EU HORIZON-MSCA-2022-SE-01 101129903, Slovakia-Austria grant SK-AT-23-0002, and APVV-23-0054]

Introduction

Sound localization is typically examined:
- **separately** in the three spatial dimensions (azimuth, elevation, distance),
- or for **combinations** of **two dimensions** (e.g., azimuth and elevation; Best et al., 2011; azimuth and distance; Ihlefeld and Shinn-Cunninghm, 2011),
- often for far sources (distance > 1 m from head), under headphones, or in anechoic space.

Estimating multiple dimensions simultaneously can be more challenging than for fewer dimensions.

Very few studies looked at 3-D localization, especially for proximal sources:
- Brungart et al. (1999): 3D localization of real anechoic sources near the head (distance < 1 m),
- Santarelli et al. (1999): nearly identical study in a reverberant room,
- both studies focused on response analysis separately in each dimension.

Current study:

- **reanalyze** the Santarelli et al. (1999) data to characterize how localization response **bias** varies with source location simultaneously in all three dimensions.

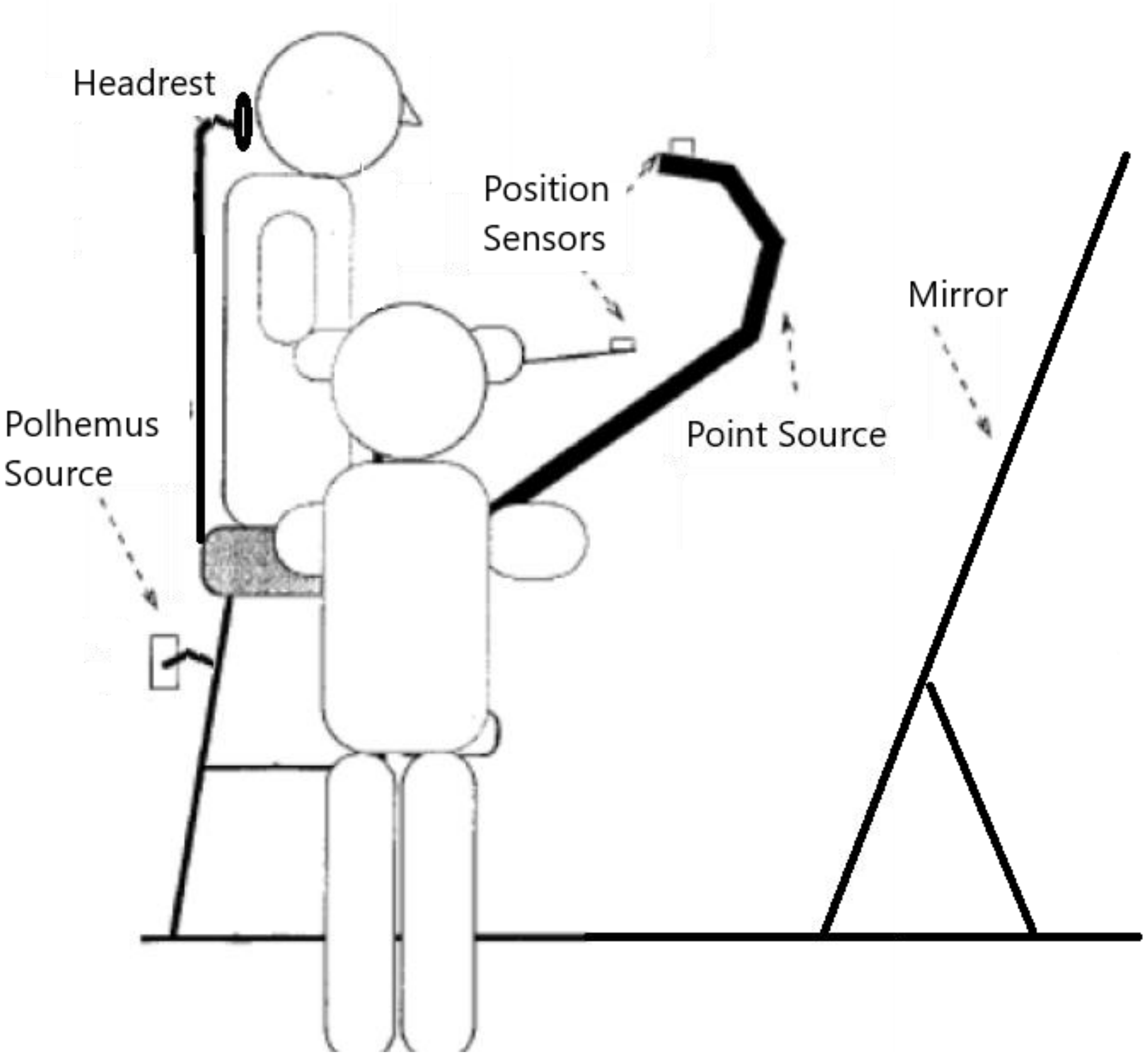
- **Hypothesis**: front-back confusion rate and response biases will be large when all 3 dimensions varied simultaneously.

Methods (Santarelli et al., 1999)

Setup
- 7 subjects (22 – 44 years of age)
- classroom 5 x 4 m, T_{60} ~ 500 ms
- subject in center of room
- freely positionable point source stimulus
- Polhemus Isotraks on point source and response wand
- mirror on plastic easel (for subject to view responses)

Stimuli
- 5 150-ms pink noise bursts separated by 30 ms silence
- random 3D locations within 1 m in right hemisphere
- level equalized (to overcome distance effects) + 15 dB rove

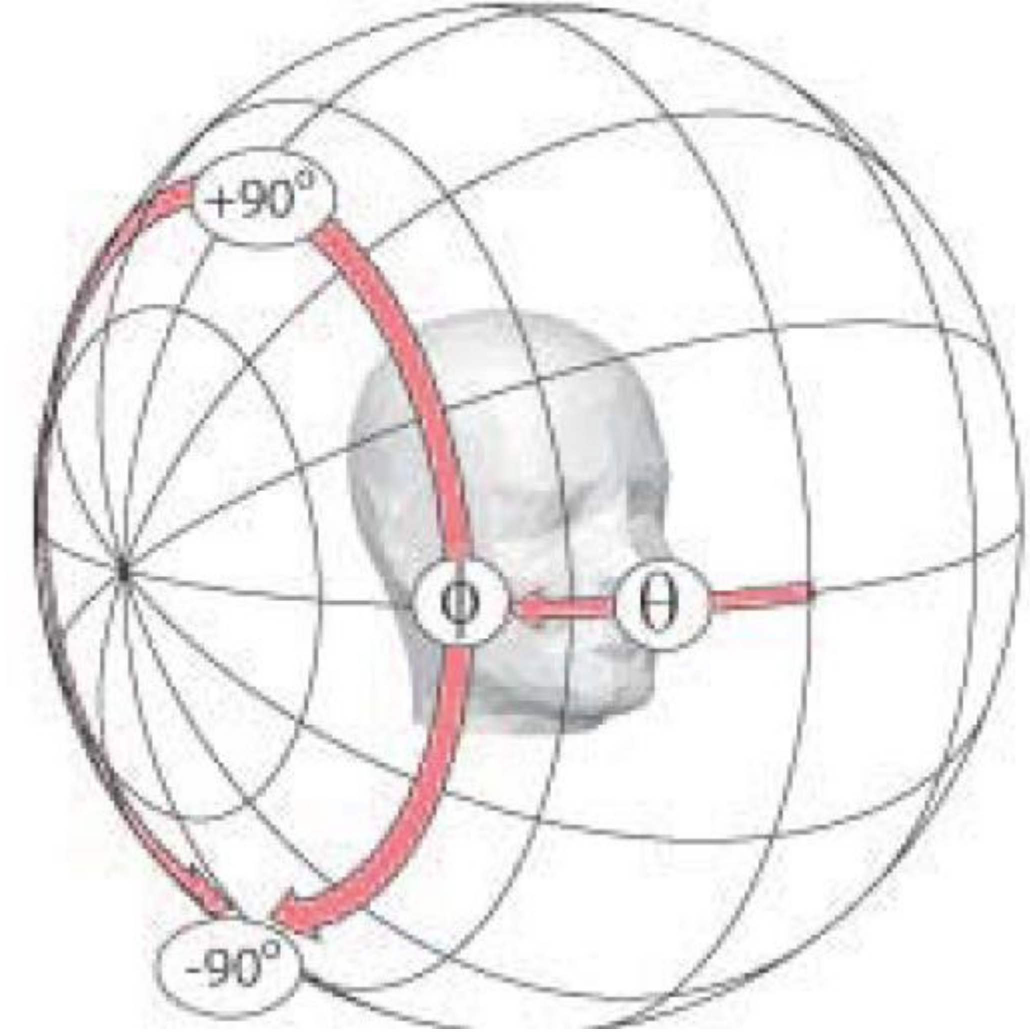
Procedure
- subject's eyes closed, source placed to random location, stimulus presented, source removed, subject responds with eyes open
- 1000 trials in 50-trial blocks, performed over 10 1-hr sessions



Brungart et al. (1999)

Methods: Data Analysis of Directional Data

Instead of using **azimuth and elevation,** use **the interaural polar co-ordinate system** to analyze the 2-D directional responses**:**
- poles aligned with ears,
- **lateral angle** roughly corresponds to iso-ILD/iso-ITD surfaces,
- **polar angle** corresponds to variation in spectral cues.

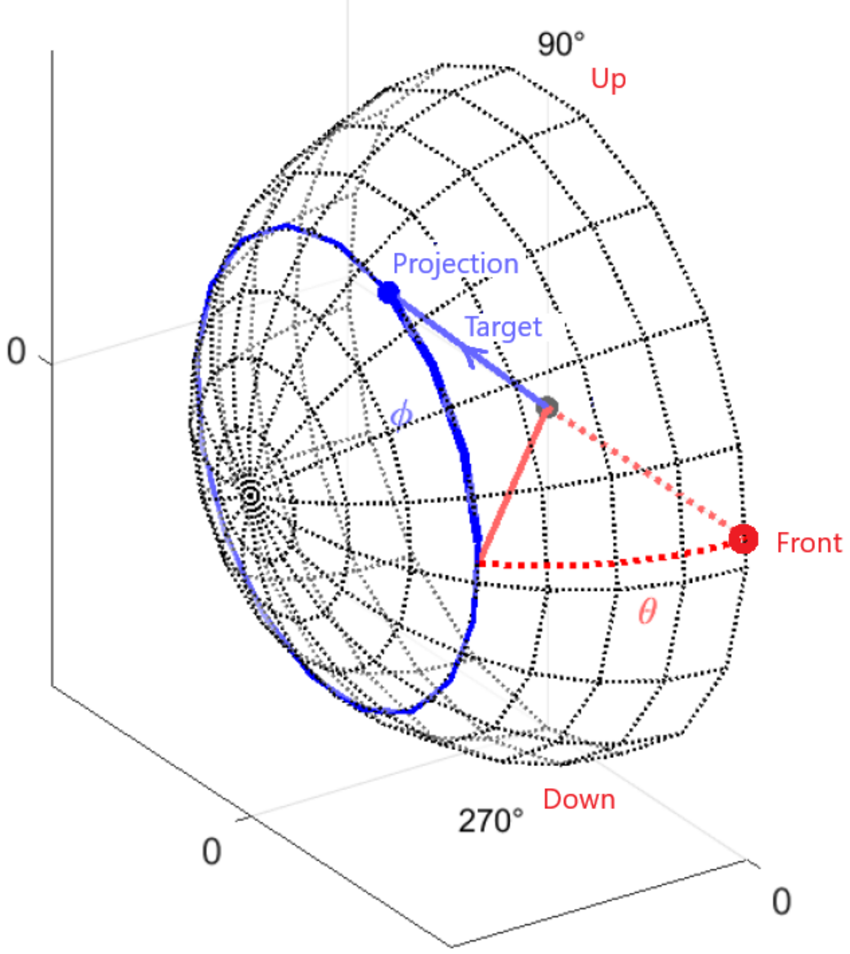


θ = Lateral Angle
ϕ = Polar Angle

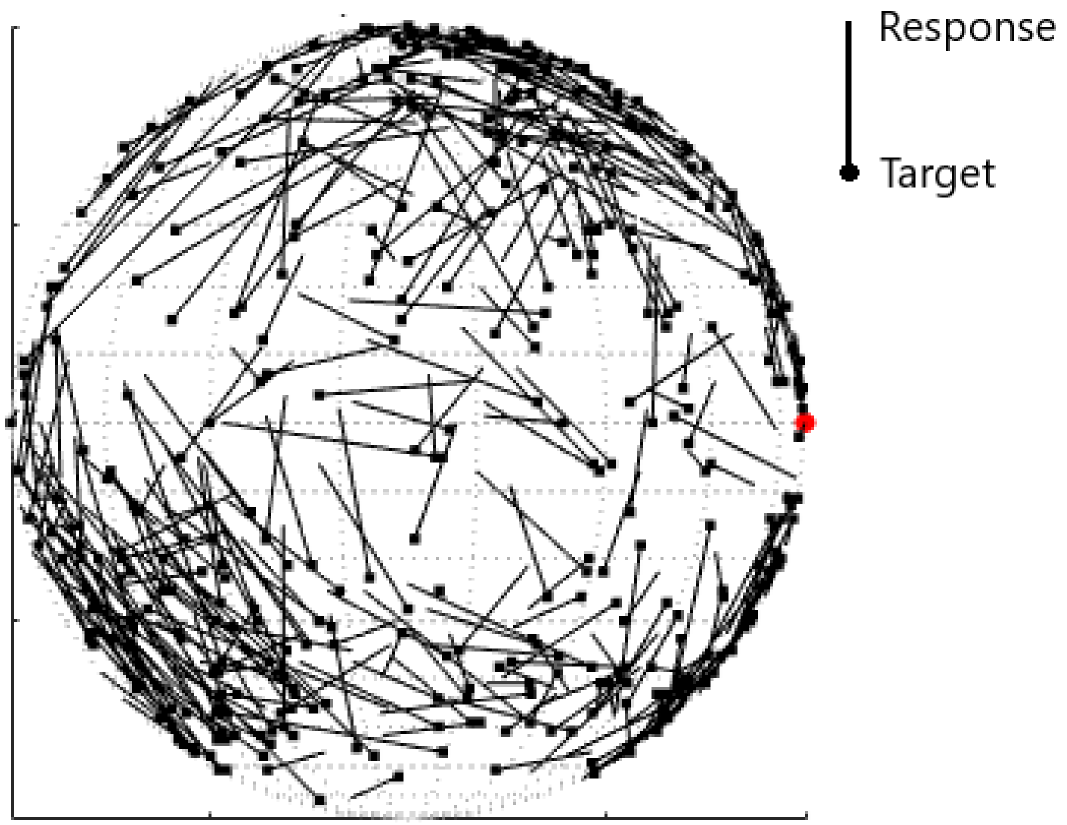
Best al. (2011)

Methods: 3-D visualization of Direction

Targets and responses projected onto surface of a unit sphere (observer at origin), separately for near and far locations (split at 50 cm).



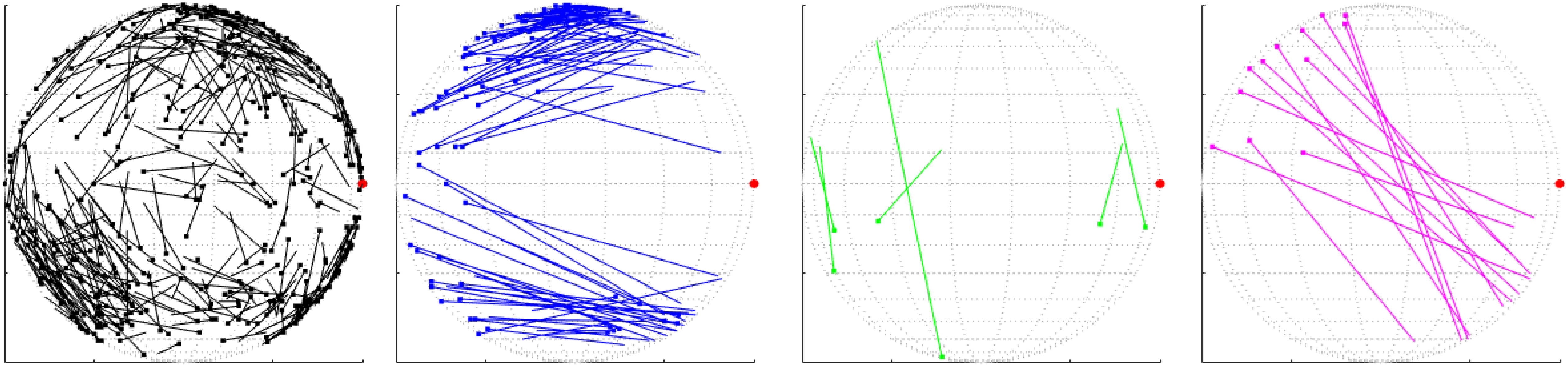
VIEW FROM (az = 45°, el = 45°).



Sample data of one subject (N=372) – SIDE VIEW

Methods: Front-Back, Up-Down, Diagonal Confusions

Confusions defined as responses in opposite hemisphere (*re.* target) for which both stimulus and response are more than 10° away from the frontal and/or horizontal plane.



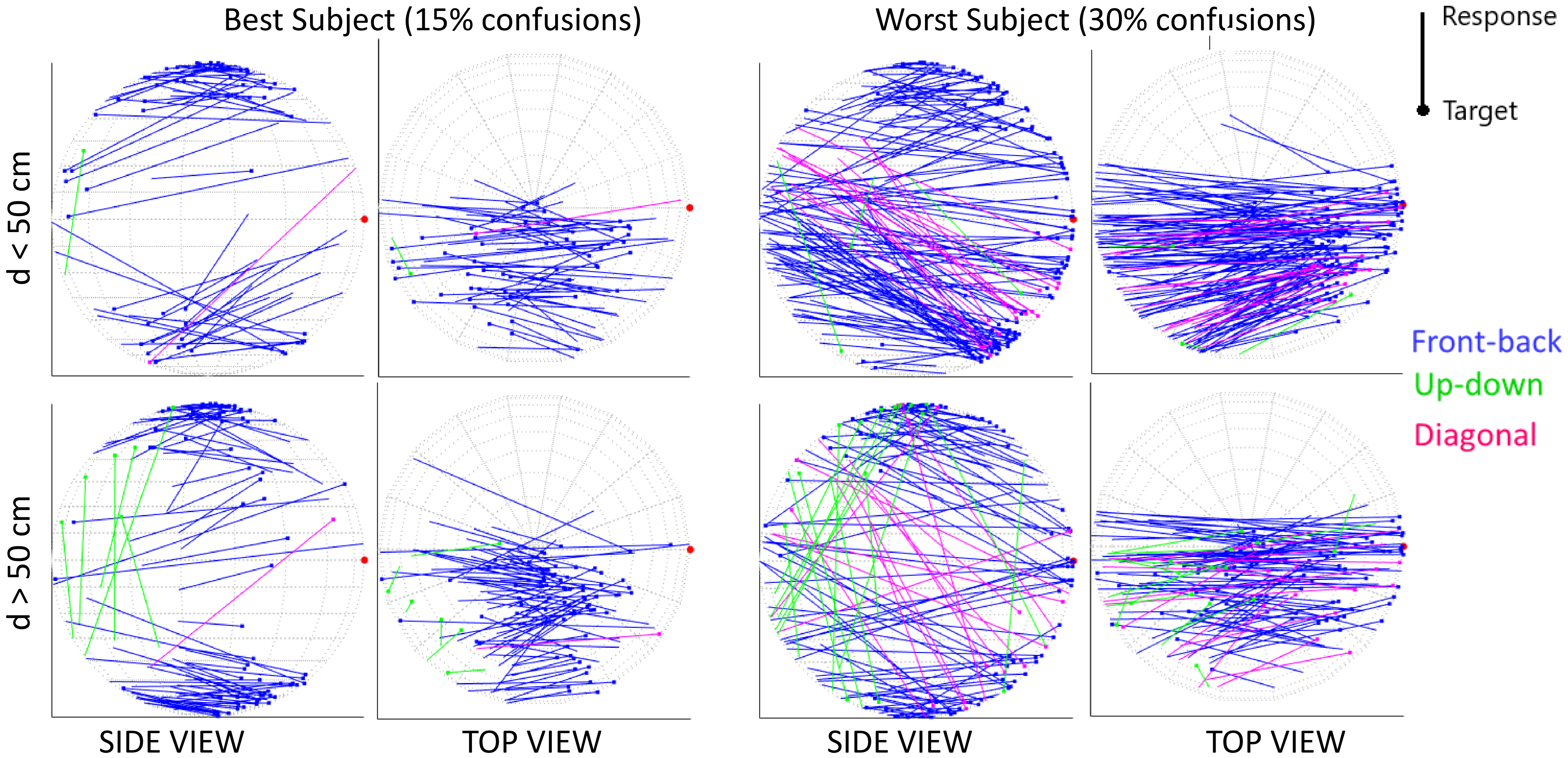
Correct Data (N=372)

Front-Back Errors (N=90)

Up-Down Errors (N=6)

Diagonal Errors (N=12)

Results: Front-Back, Up-Down, Diagonal Confusions



Results: Front-Back, Up-Down, Diagonal Confusions

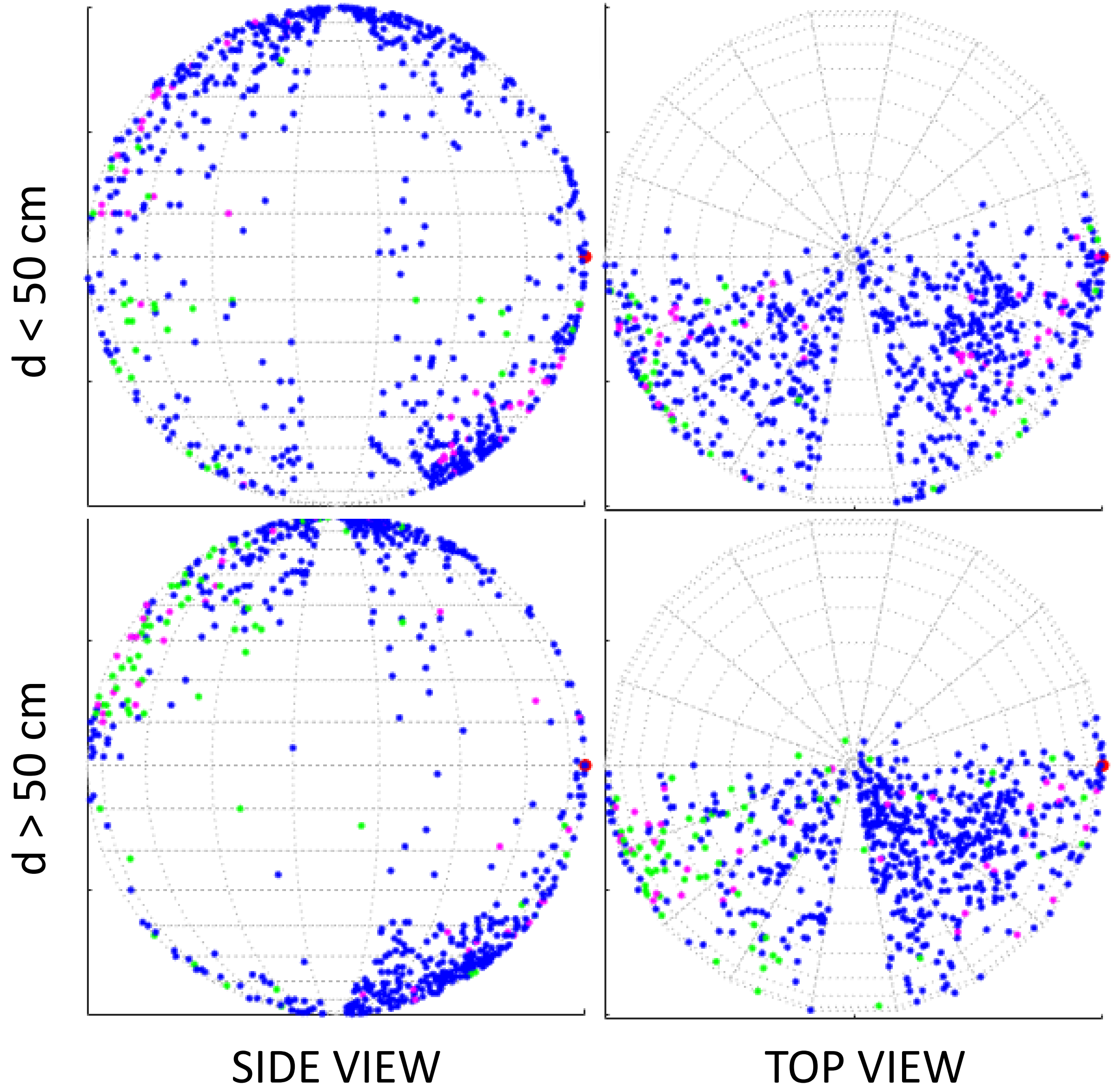
Figure only shows the target location, with color indicating type of confusion / response quadrant.

On average, **20%** of responses were confusions.

Front-back (87% of confusions):
- concentrated away from horizontal plane, especially near bottom in front of frontal plane.

Up-down (6%):
- mostly at back (more for far sources)
- lower back quadrant for nearby sources, but
- upper back for far sources.

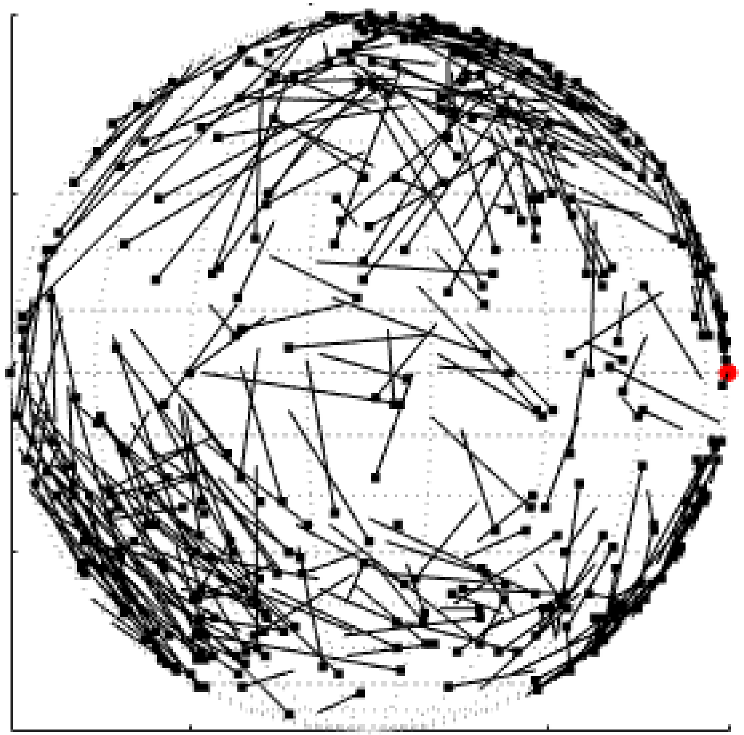
Diagonal (6%):
- mostly along the main diagonal
- some subjects mostly upper back, some lower front



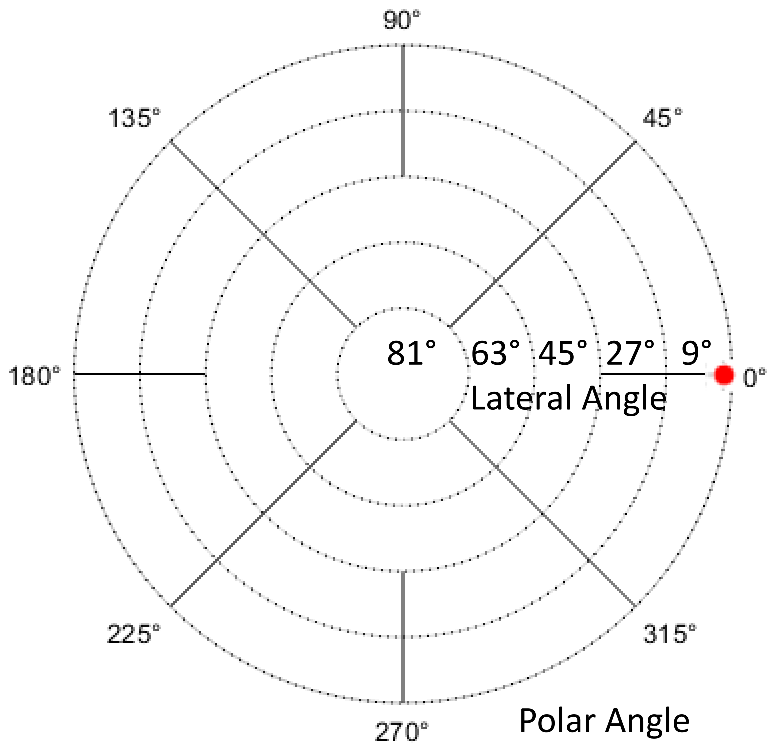
Methods: 3-D visualization of Direction and Dist.

Clean data (after confusions removed) binned into **25** directional **bins** by
- **lateral angle** (5 bins centered at θ=[9, 27, 45, 63, 81]°)
- **polar angle** (8 bins at ϕ=[0, 45, … , 315]° for θ=[9, 27]°, 4 bins at ϕ=[0, 90, 180, 270]° for θ=[45, 63]°, 1 bin for θ=[81]°)

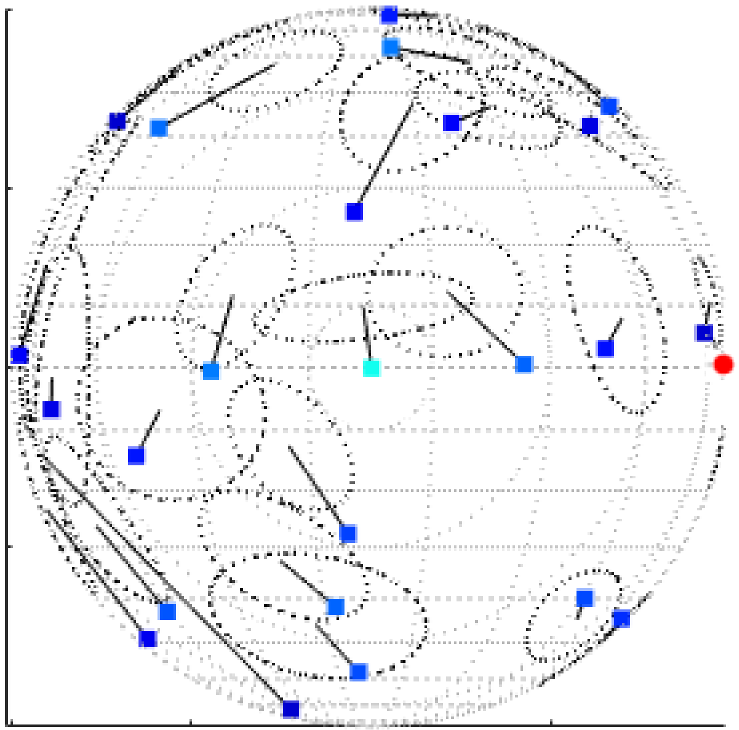
After binning, the mean stimulus and response directions determined in cartesian coordinates (SPAK toolbox, Leong & Carlile 1998), using Kent distribution ovals to visualize spherical spread of biases across responses/subjects.



Clean Raw Data (SIDE VIEW)



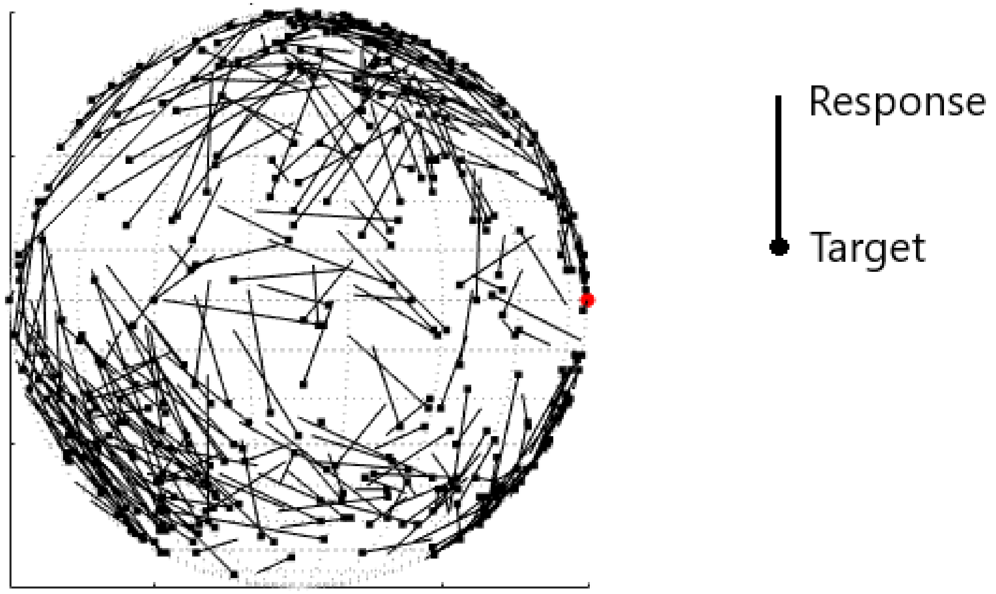
Bins on a Flattened Surface of Hemisphere

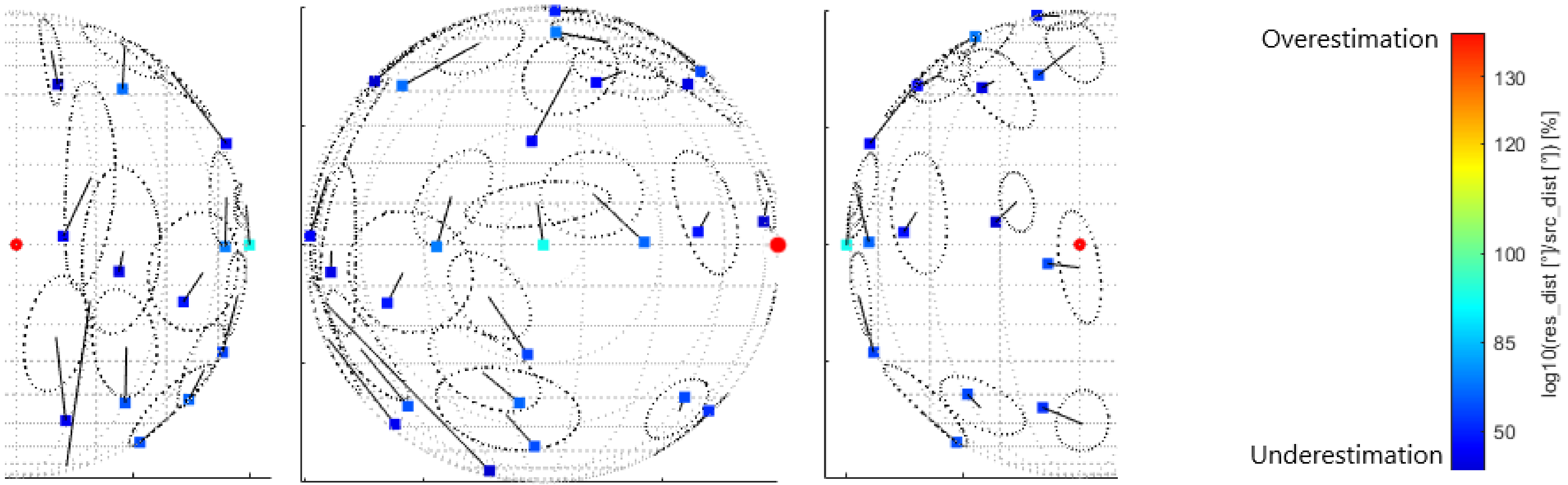


Mean + Distribution of Resps

Methods: 3-D visualization of Direction and Dist.

Mean relative response **distance** biases, determined on a log scale, indicated by color of the target circle.





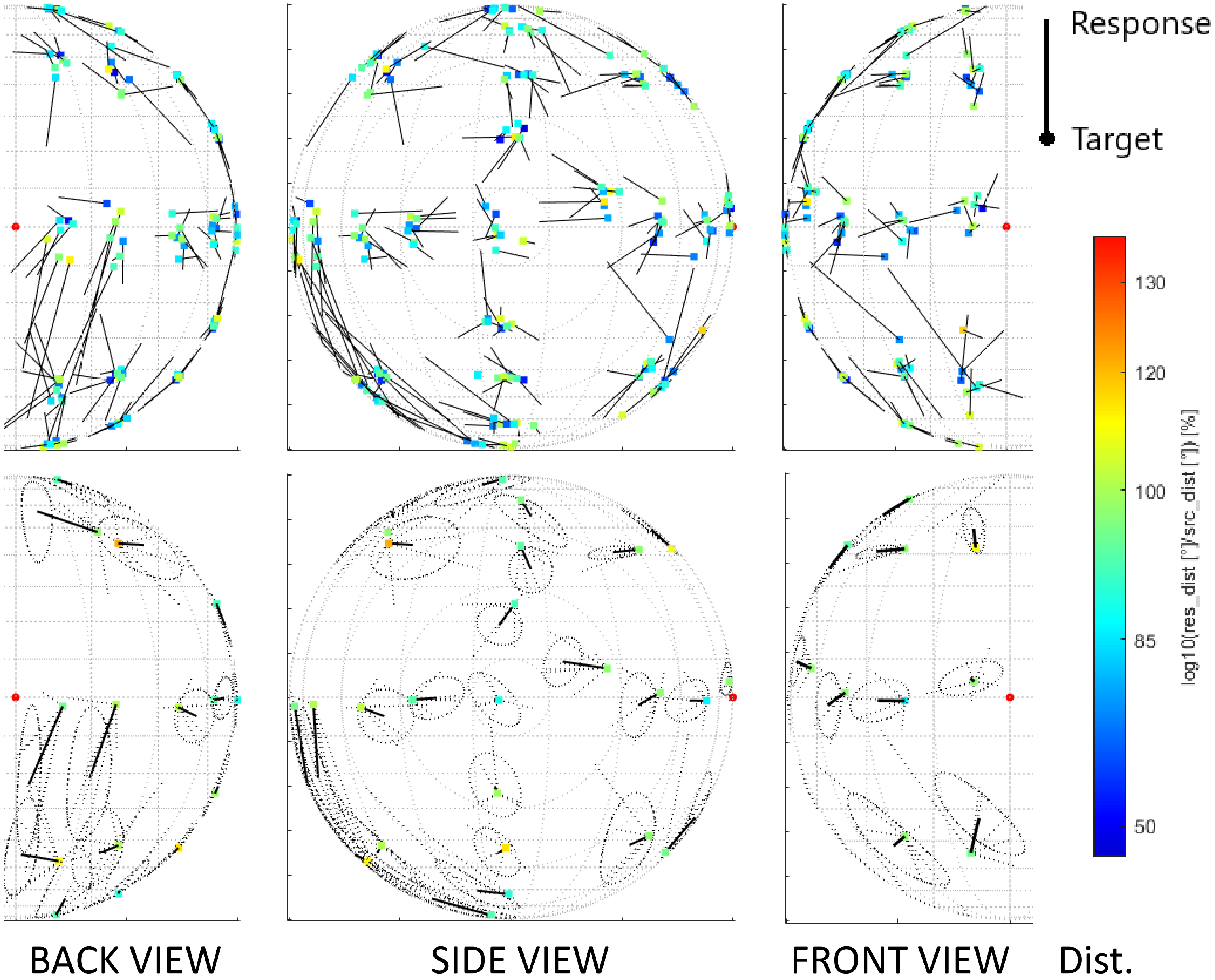
BACK VIEW SIDE VIEW FRONT VIEW Distance Bias

Results: Biases in all 3 dimensions (w/o confusions)

Example data for d > 50 cm

Top row: individual subject data in 25 directional bins.

Bottom row: data for each subject shifted to across-subject mean target location.



Results: Directional Biases

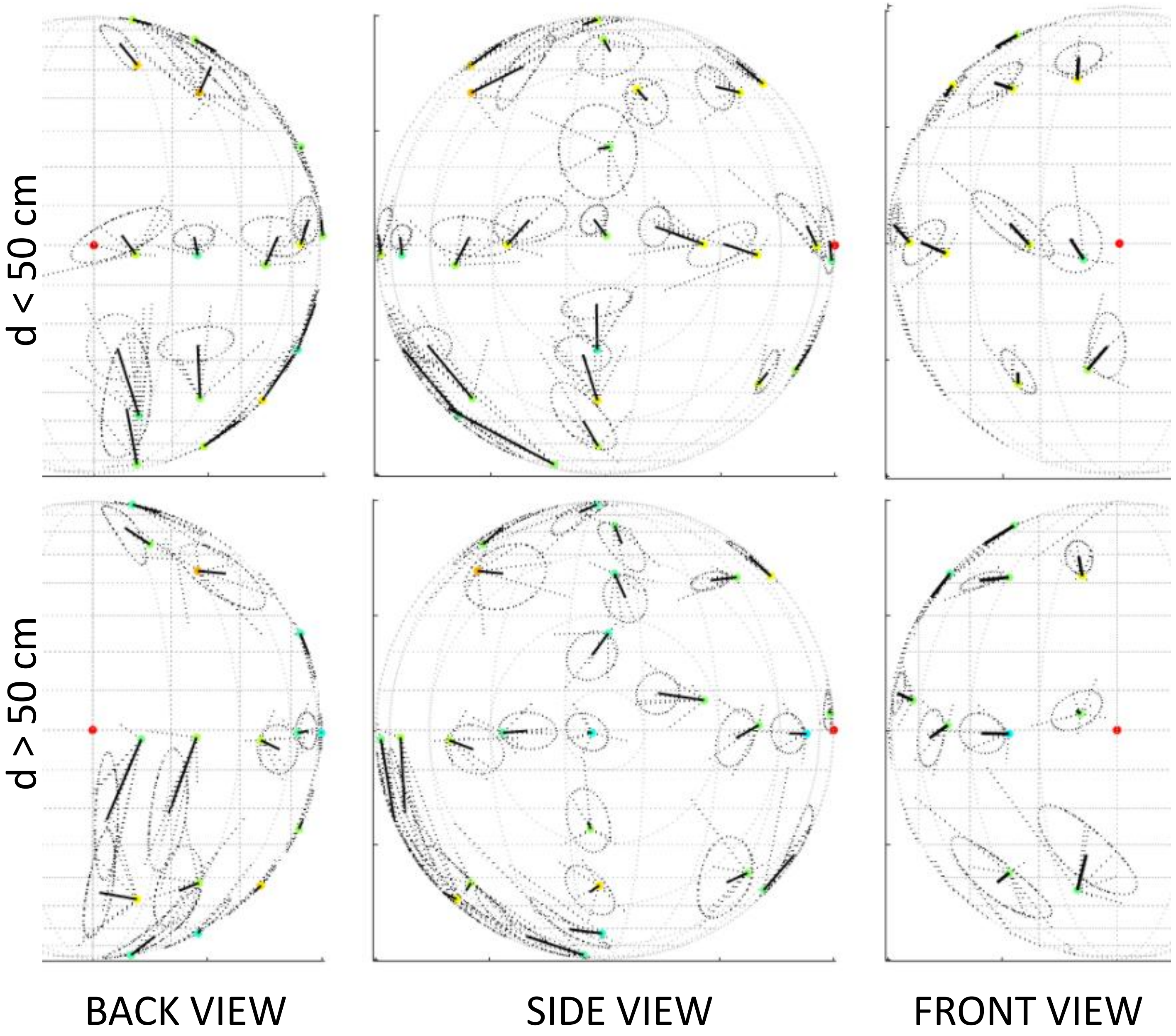
Largest biases for sources behind and below the subject's head:

- responses biased downwards and medially for far sources (20-22°),
- response bias reversed to upwards bias for near sources (17-30°).

Similar results to up-down errors.

For near sources only, a general upwards bias trend.

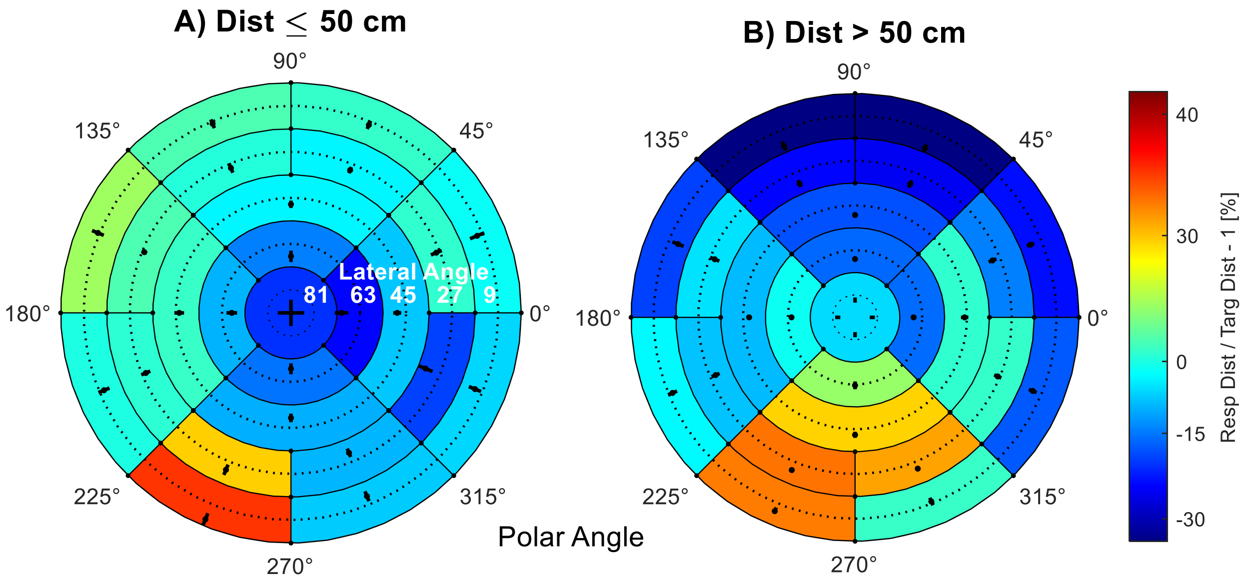
For both near and far sources close to the horizontal plane and away from the medial vertical plane, a lateral bias was observed (2-14°).



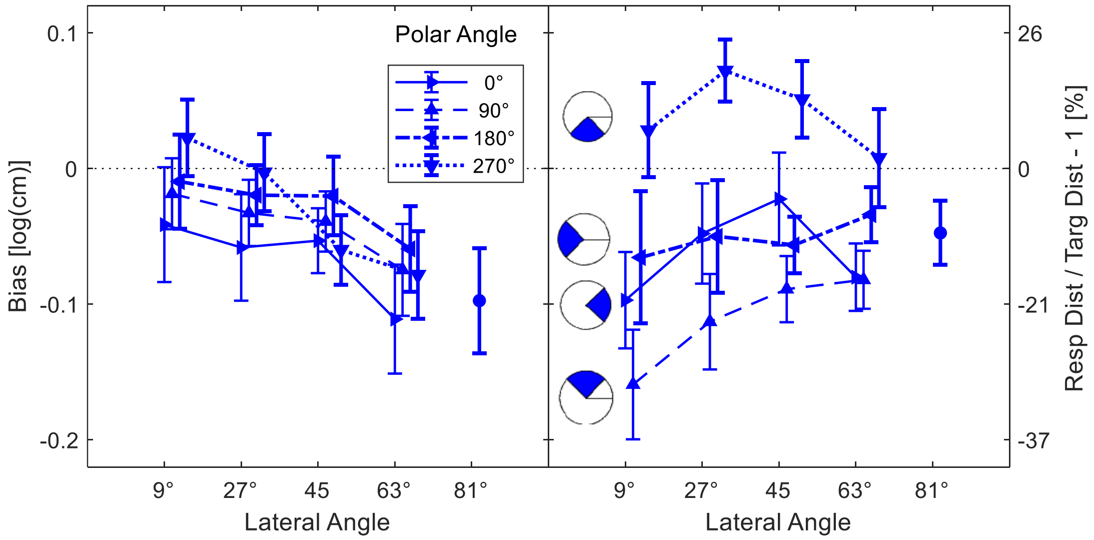
Methods: Distance Biases

Analysis

- FB/UD/Diagonal confusions included
- data plotted in 2 formats:
    - upper panels: polar plot of color-coded distance bias in 25 directional bins



    - lower panels: distance bias as a function of lateral angle, parametrized by polar angle



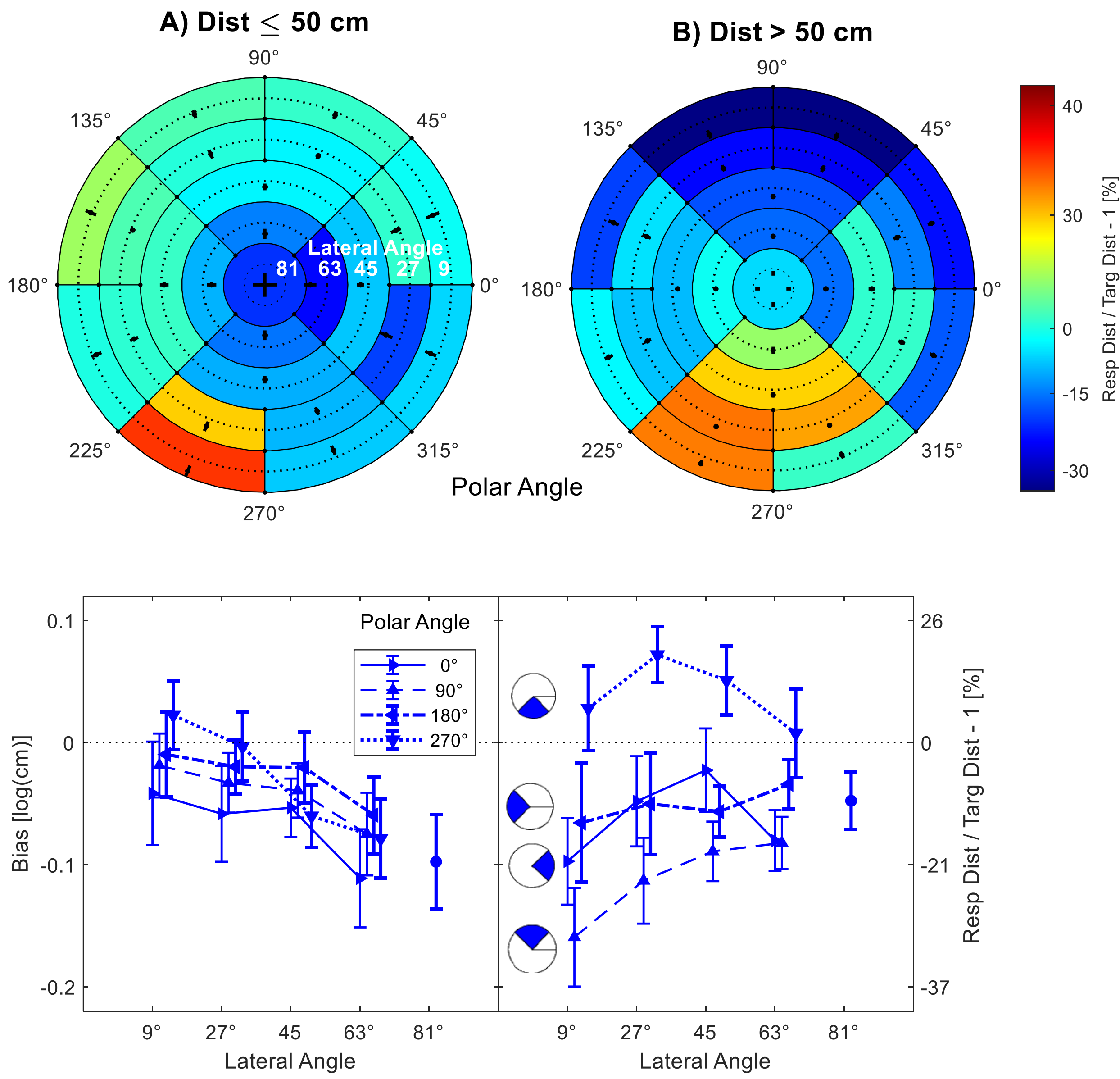
Results: Distance Biases

Upper panels:

- slight underestimation near horizontal plane (0 to -20%),
- underestimation above subject up to -30% for far sources (dark blue patch),
- overestimation below subject up to 40%, especially behind the subject (red and orange patches).

Lower panels:

- same trends shown.



Summary and Discussion

We examined **3D localization** of nearby sources (d < 1 m) **distributed in 3D** in **real reverberation**.

High Front-Back confusion rate (17%), and lower but considerable Up-Down and Diagonal confusion rates (1.5% each).

- Uncertainty about location mostly affects the ability to map spectral cues to polar angle. Not clear why that dominantly causes FB confusions.
- Brungart et al. (1999), using a more liberal criterion, reported 10% front-back rate (and no other confusions) in anechoic room. Reverberation might cause frequency-to-frequency spectral variation, increasing the uncertainty in spectral cue-to-polar angle mapping.

Directional responses showed a complex pattern of biases:

- largest biases (20° to 30°) for sources behind and below listener, reversing with distance
- general upwards bias trend for nearby sources,
- lateral bias for sources in horizontal plane.

Some of the effects might be due to response method. But most point to distortions in auditory space representation (Brimijoin, 2018).

Summary and Discussion

Distance biases:

- general underestimation (10%), not consistent with previous distance-only studies (reporting overestimation for $d < 1$ m, Zahorik et al., 2005)
- strong underestimation (30%) for sources above listener and overestimation (more than 20%) for sources below. Space might be perceptually warped / distorted in the vertical dimension as sources rarely come from these directions.

Next steps:

- compare to anechoic data of Brungart et al. (1999)
- directional data: analyze the data separately for lateral and polar errors and/or establish significant effects in 2D (or 3D!)
- Analyze response variance.

Take-home message:

This study provides normative characterization of how good (bad) localization performance is in the "best-case" real reverberant room scenario.

The results might be useful as a reference for development of VR/XR applications in which near field is simulated.

THANK YOU!