



The Influence of Prior Experience on Distance Perception in Fixed and Varying Environments



Norbert Kopčo and Myroslav Fedorenko

Perception and Cognition Lab
P. J. Šafárik University
Košice*, Slovakia



Supported by Slovak Republic grant APVV-23-0054 and by European Union HORIZON-MSCA-2022-SE-01 grant agreement N° 101129903



* pronounced KOH-shih-tse

Introduction

Level-independent distance perception is

possible for nearby sources (< 1 m)

- in **anechoic** space: at locations **away from midline**, using interaural level difference (**ILD**; e.g., Brungart et al., 1998)

- in **reverberation**: in **all directions**, using **reverberation-related** cues like direct-to-reverberant energy ratio (**DRR**) + ILD; e.g., Kopco et al., 2012)

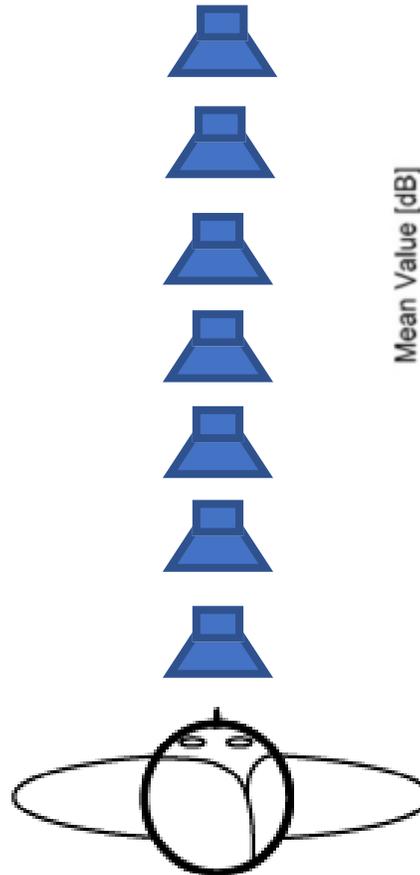
Weights used by listeners to combine

DRR/ILD/other cues depend on

- **context** of currently presented stimuli (Doreswamy et al., 2019),

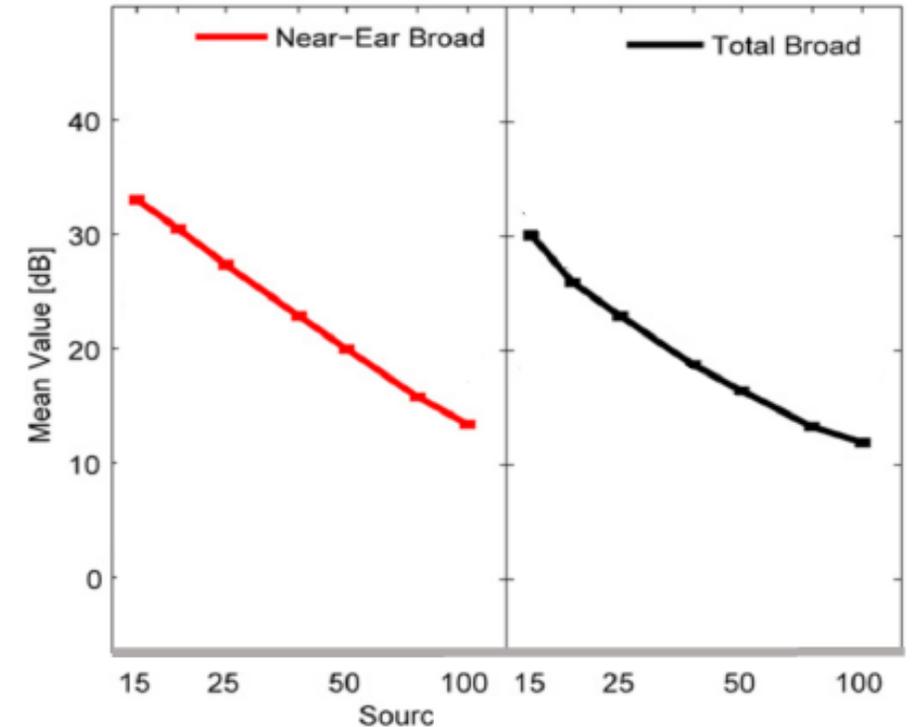
- especially **during initial exposure** to a given room (Hladek et al., 2013).

DRR only



DRR

ILD



ILD + DRR

Current Study

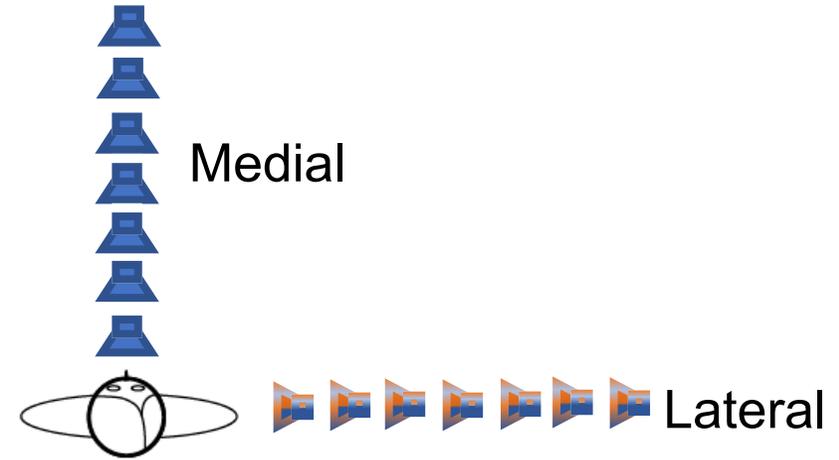
Here we performed an experiment (Kopco & Fedorenko, ARO '24) to examine:

1. How does the **context** of **varying** the environment from **trial to trial** (vs keeping the environment consistent) influence distance perception?
Specifically, when listeners judge distance in 3 different virtual environments on consecutive trials, can they concurrently maintain tuning to **3 separate model rooms**, or do they create **1 combined model**?
2. Does **initial exposure to inconsistent context** affect performance when it later becomes **consistent** (or vice versa)?
3. **Is distance perception** influenced by the **early reflections** when listener is near **the corner** of a room?
4. Is **performance** different for **frontal** vs **lateral** sources (differing in availability of ILD)?

Experiment in Virtual Environment

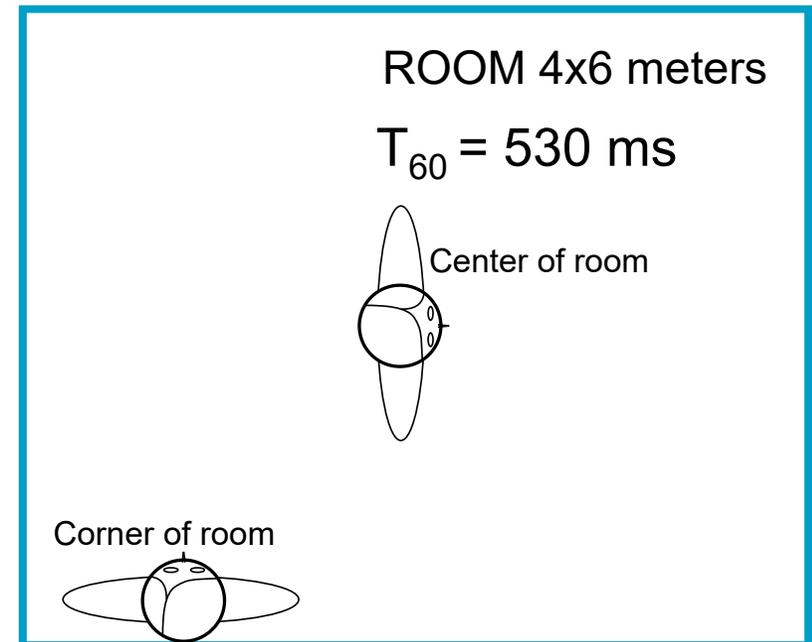
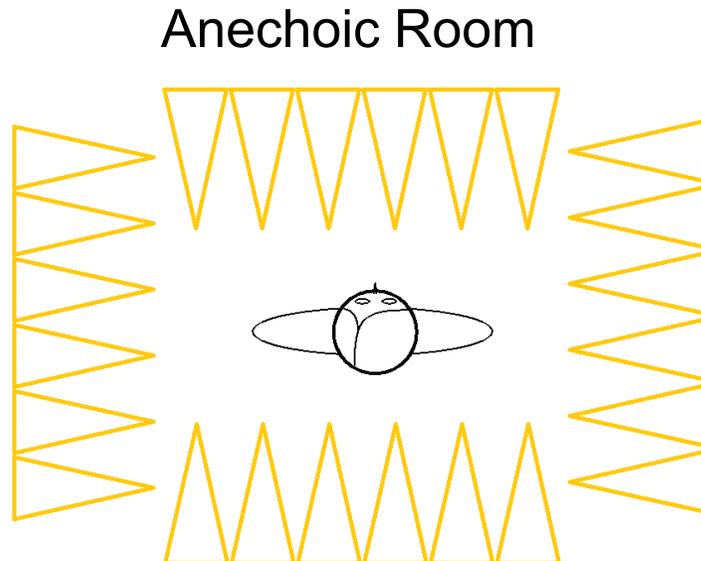
Stimuli:

- five 150-ms-long pink noise bursts (30-ms gaps)
- roved by 15 dB (to eliminate level cue)
- 9 distances (15 to 170 cm, log spaced)
- 2 directions (medial and lateral)



Room conditions:

- 3 virtual environments simulated using **individually** measured **BRIRs**
- **anechoic, center, and corner** of a midsize classroom



Experiment in Virtual Environment

One trial

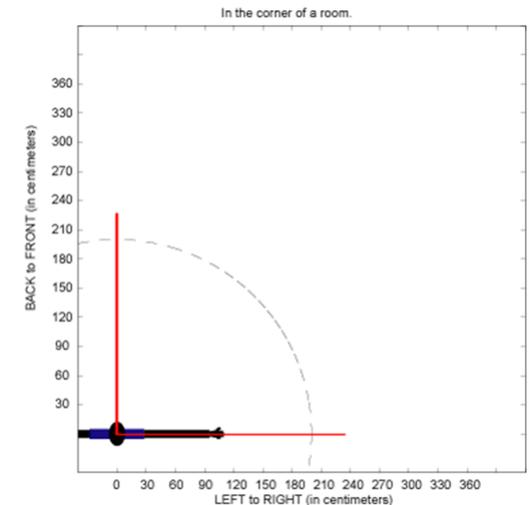
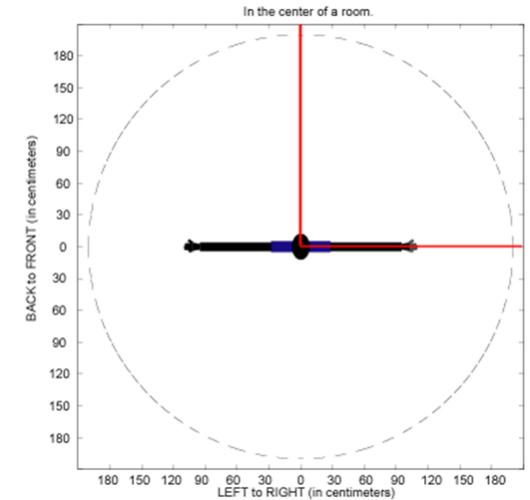
- subject **informed** about **room condition**
- simulated source **presented** over headphones
- subject **indicated heard position** by a mouse click on screen

Each subject performed two sessions: **FIXED** and **MIXED (contexts)**

- session consisting of **6 blocks**, each containing **8 runs** (4 med, 4 lat)
- each **run** had **45 trials**,
with sources coming from **one direction**, only **varying in distance**
- **FIXED** context sessions: simulated **room fixed** within a **block**
- **MIXED** context sessions: **room selected randomly** on each trial

Two subject groups

- **initFixed** group (4 subjects): **FIXED** session followed by **MIXED**
- **initMixed** group (4 subjects): **MIXED** session followed by **FIXED**



Results: Stimulus-Response Corr. Coef. r

Performance (analyzed as r on log scale)

tends to be better for:

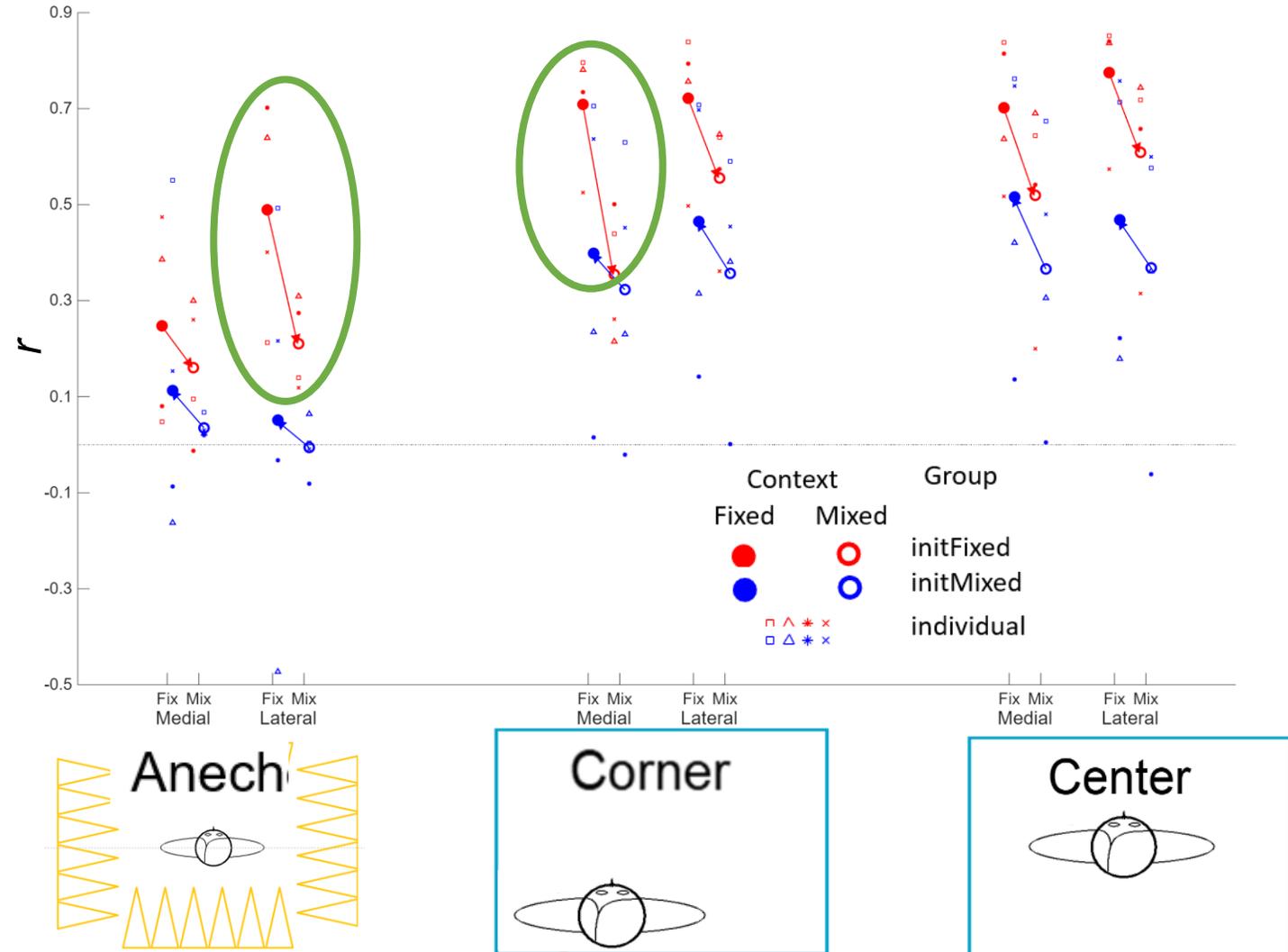
- **initFixed** group,
- Fixed context (●, ●),
- lateral direction,
- room (CE>CO>>AN)

However, **initial context** matters:

- **starting with Mixed** (○) results in **bad** performance that **improves slightly** when switched to Fixed (●)
- **starting with Fixed** (●) results in **good** performance that **deteriorates dramatically** after switch to Mixed (○)

Why is the impact of **Fixed-to-Mixed** transition so **large** in **MED CO** and **LAT AN** conditions in **initFixed** group?

Analyze st.dev., stim-resp slope, & bias.



Results: Response Standard Deviation

St. dev. averaged across distance:

- **worse** in **AN** than in **CO/CE** rooms
- **worse** in **Mixed** (\circ , \circ) context

Direction x Room x Group x Context interaction:

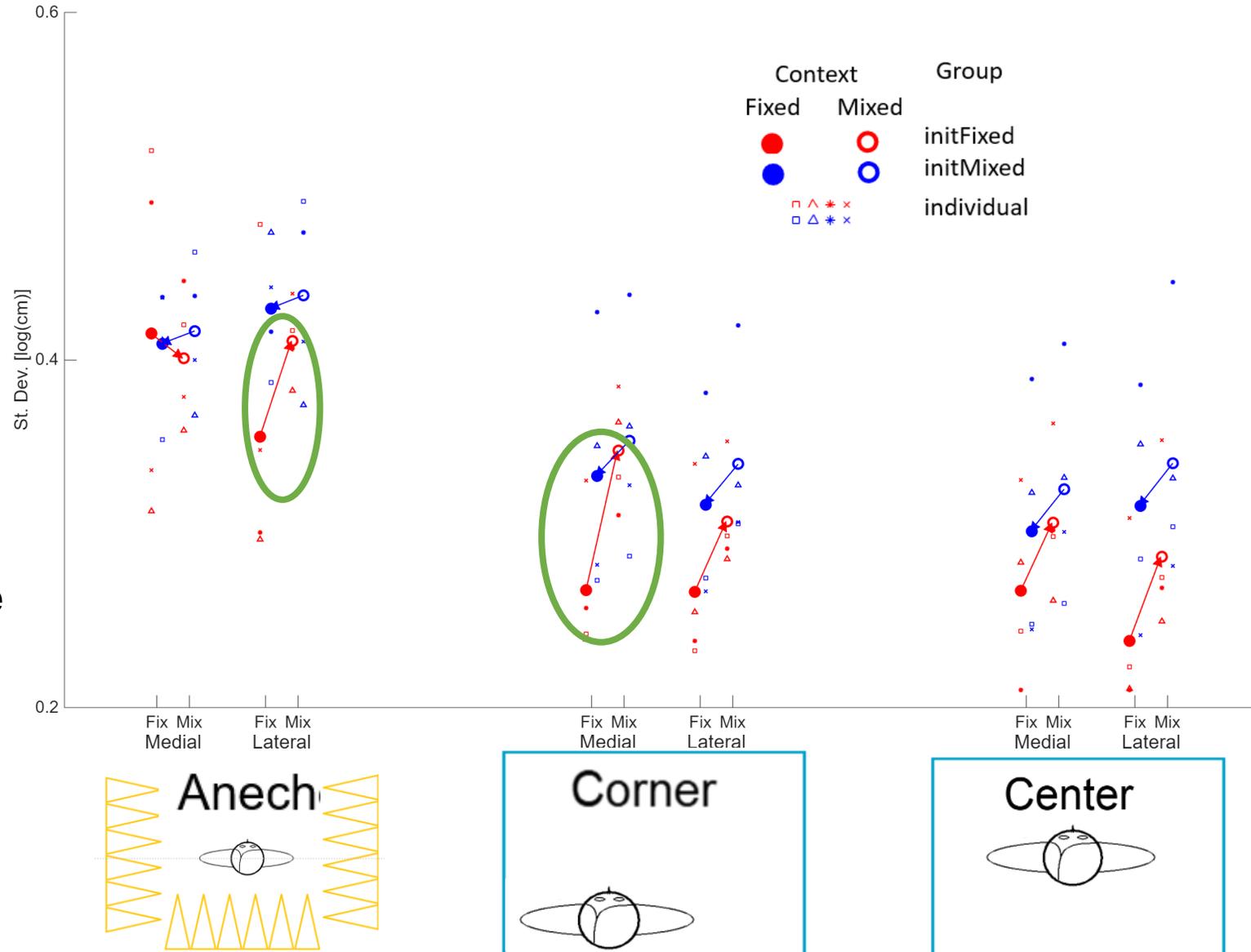
initMixed group

- **small** direction-independent increase in **Mixed** context (\circ)

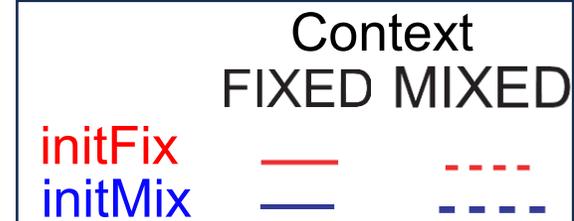
initFixed group

- **large** direction-dependent increase in Mixed context (\circ), especially for **CO MED** and **AN LAT**

Response variability increases with mixed context in the same conditions in which r decreased.



Results: Mean Response



Resp. distance grows mostly linearly

with **stim. distance**:

- **weakly** in **AN**,
- **strongly** in **CO/CE**,
- **more** for **lateral** than **medial** direction

Responses in Mixed context

shifted re. Fixed:

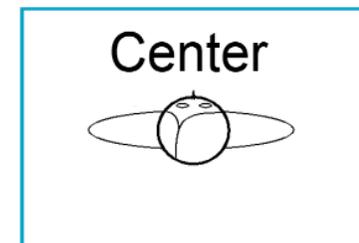
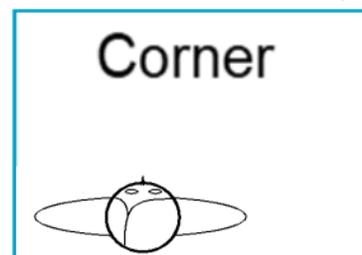
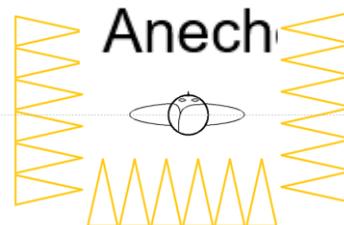
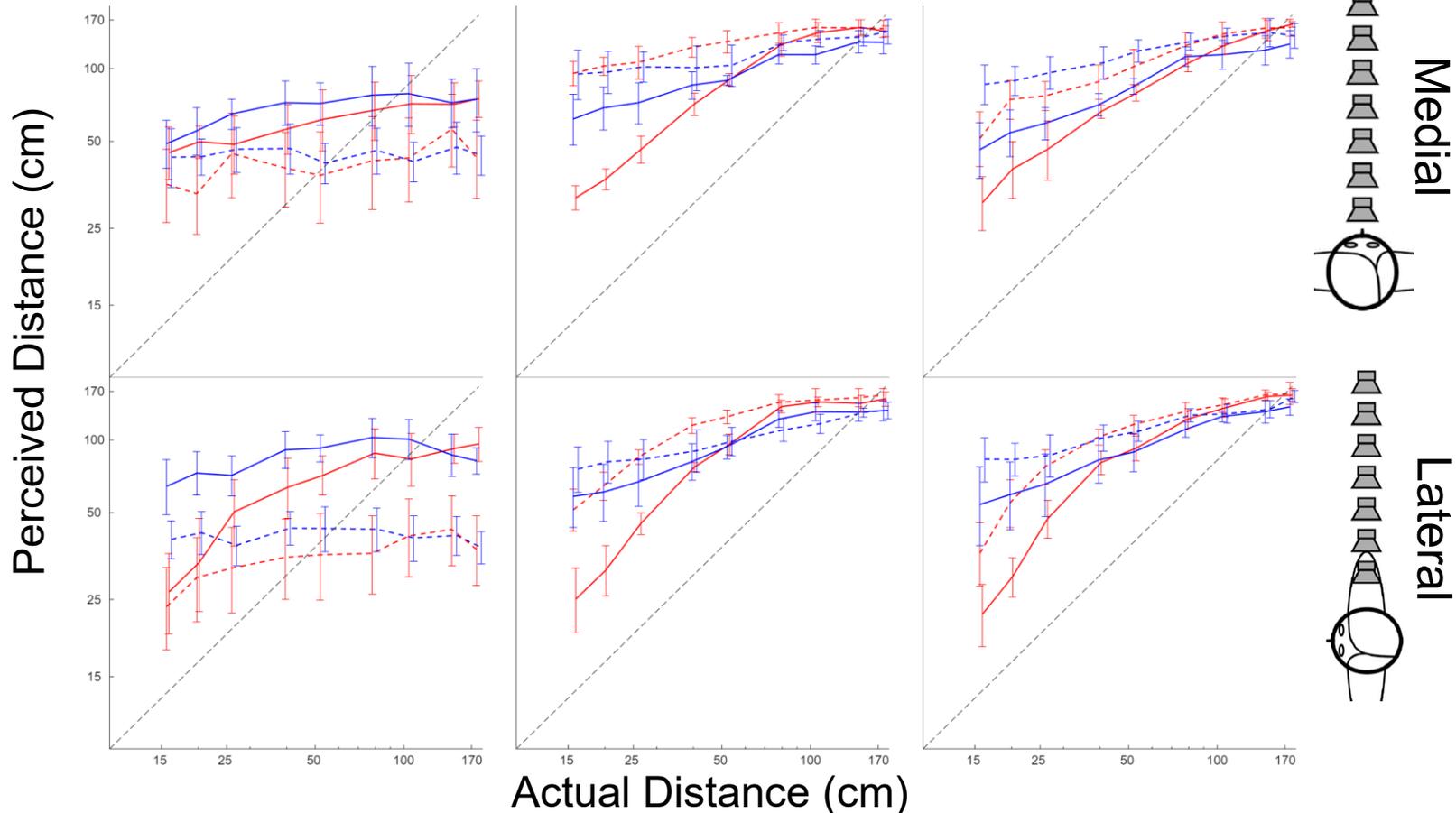
- **closer** in **AN**,
- **away** in **CE/CO**.

Nearby targets

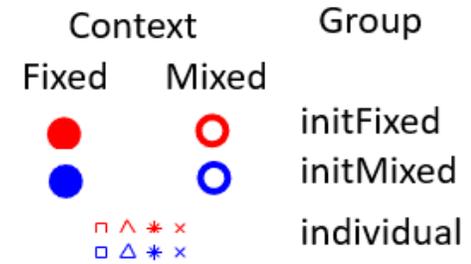
overestimated.

Distant targets

underestimated in **AN**,
mostly overestimated in **CE/CO**
(likely **edge effect** due to limited response range).



Results: Slopes



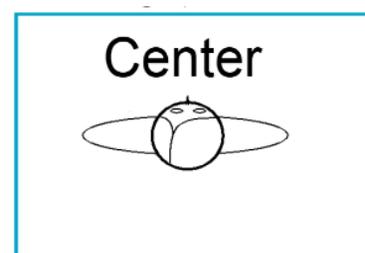
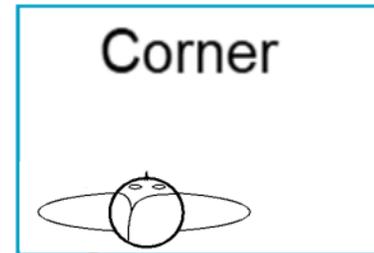
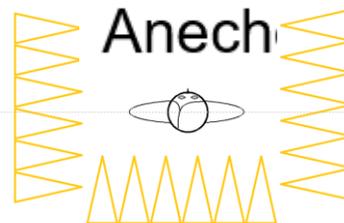
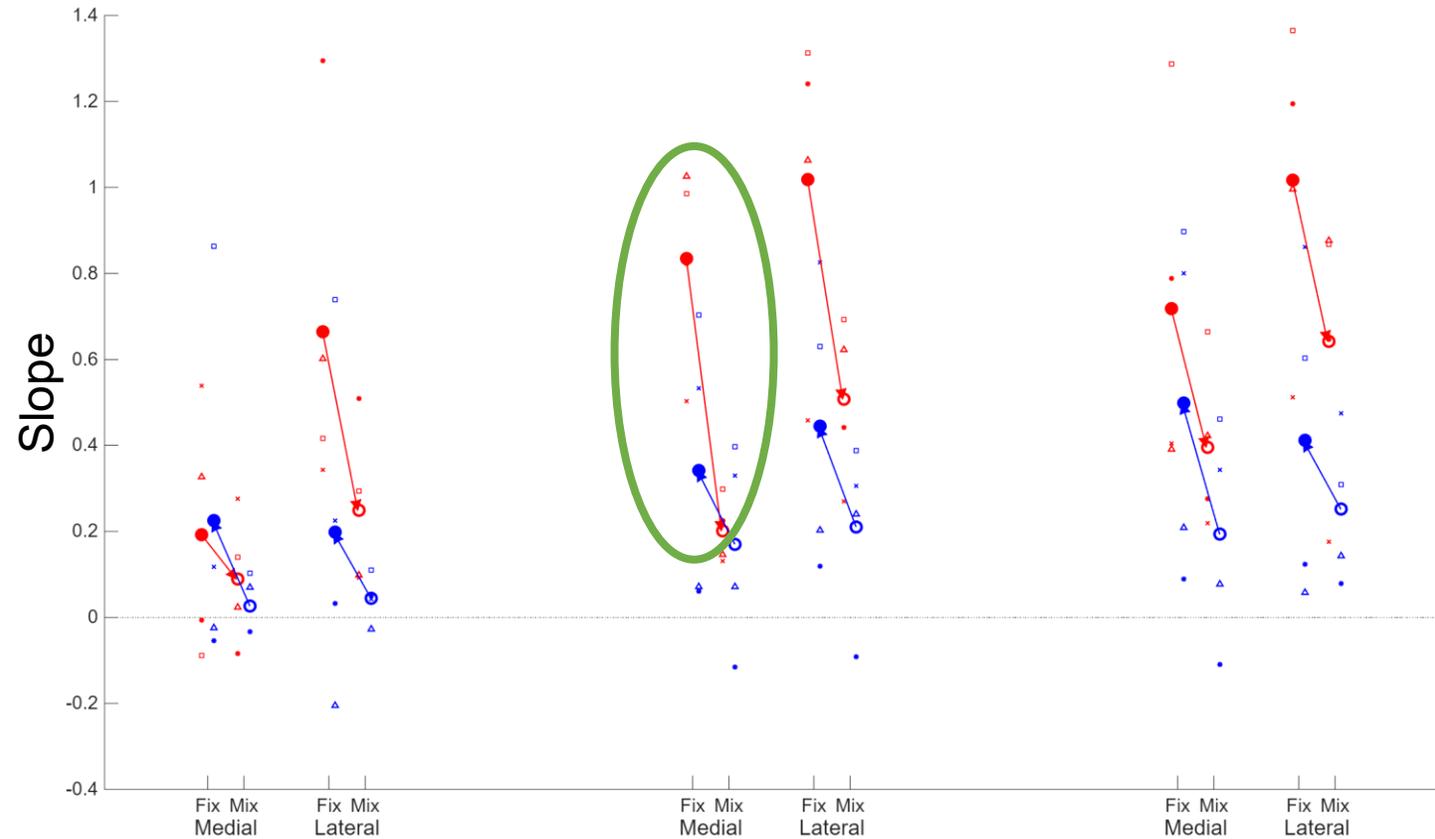
(Estimated **without 100 - 170 cm dists** because of edge effect).

AN:

- worst for **AN MED** (no cues)
- good in **AN LAT (ILD)** but only for **initFix** group Fixed context (●)

In all **CE/CO** conditions a similar pattern:

- **slope is steeper** in **Fixed** context
- **initMix** group
- slightly
- **initFix** group
- strongly,
- depending on direction,
- strongest in **initFix** MED CO cond.



Change in Stim-Resp slope can partially explain the changes in r with Fixed-to-Mixed switch.

Results: Bias (avg'd across dist)

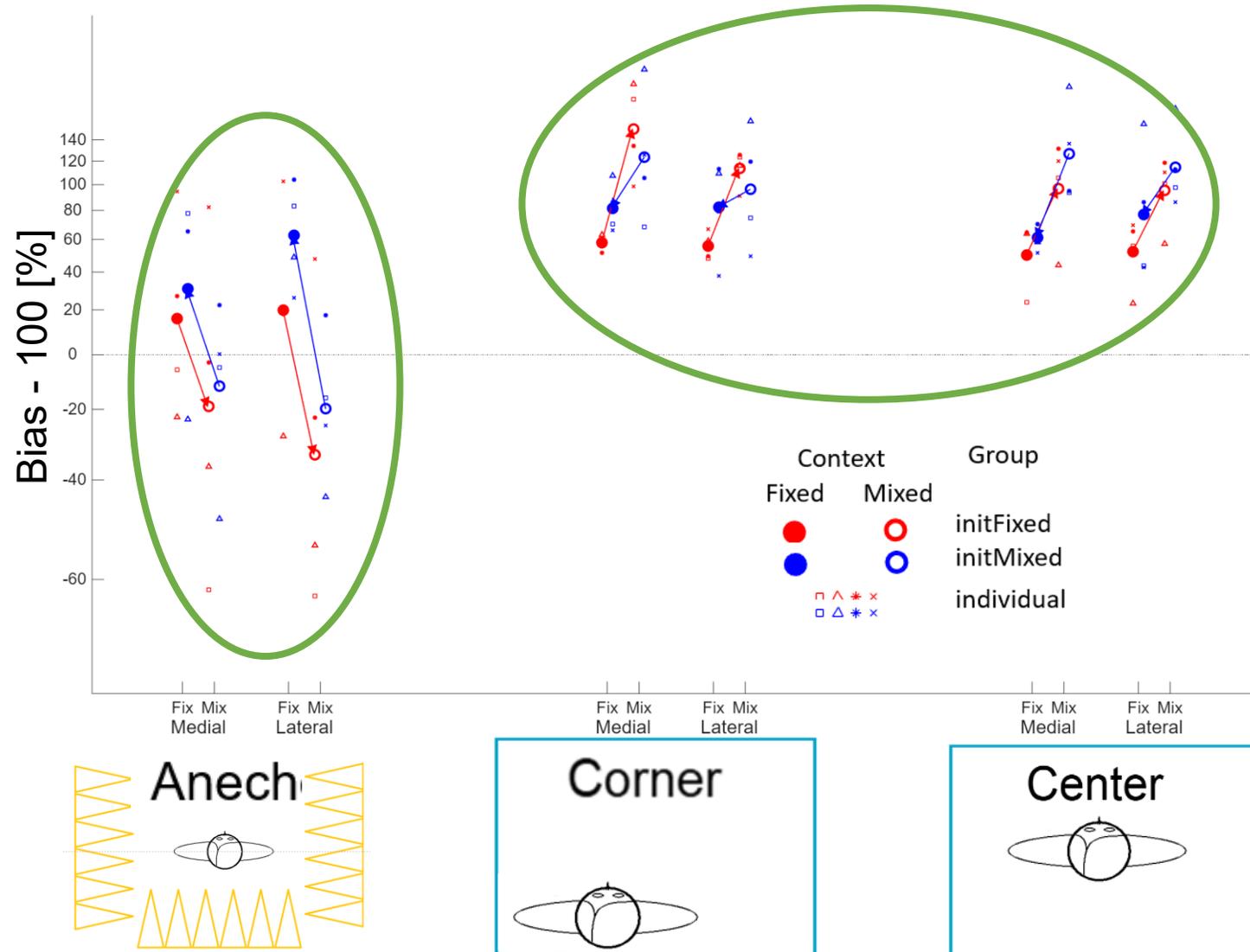
In **fixed context** (●, ●):

- bias approximately constant across rooms and directions.

In **mixed context** (○, ○):

- for **both directions**:
 - in **AN**, responses shifted **closer**,
 - in **CO/CE**, resps shifted **away**.

These biases do not influence r .



Summary and Discussion

Group **starting in Fixed** context:

- in the **initial, Fixed context:**

level-independent distance perception (measured by r) is better

- in reverberant than AN rooms, with CE slightly better than CO →

Reverberation improves performance, but early reflections are detrimental.

- for lateral than medial sources →

Both reverb-related and ILD cues are used by listeners.

- **after switching** from Fixed to Mixed context:

- performance (r) worsened in all conditions,

but the effect differed by room and direction:

- largest decrease in CO Med, both due to increased st.dev. and shallower stim-resp slopes →

**When relying on reverb-related cues, early reflections are the most “damaging”
(but not when ILD is still available in CO Lat).**

- large decrease in AN Lat (and increase in st.dev.) →

Listeners use ILD cue less.

- in CE/CO, performance still **better** in lateral than medial direction →

But, ILD cues likely still used to some extent.

Summary and Discussion (cont.)

Group starting in Mixed context:

- overall worse performance (in terms of r , st.d. and slopes) than group starting Fixed,
 - improvement in Fixed context very small in all conditions,
 - no difference between lateral and medial performance
- If starting in a Mixed context, listeners use reverb-related cues but not ILD, and they do not benefit from being exposed to consistent-room context, or from ILD even when all the cues are available in the Fixed session (performed as 2nd).

Biases for both **initMixed** and **initFixed** groups (independent of r):

Switching from Fixed to Mixed context from trial to trial induces underestimation in AN and overestimation in CE/CO.

- Listeners cannot separately process distance cue information from different rooms on trial-by-trial basis.
- Listeners likely create a single DRR-to-distance mapping/model in Mixed context, since:
 - in AN ~ very large DRR → percepts biased closer,
 - in CE/CO ~ smaller DRR → percepts biased away.

Conclusions

How the distance cues are combined and weighted for level-independent distance perception depends on:

- the **current and initial context**,
- **which cues** (ILD/reverb-related) are **available**,
- **presence of early reflections**.

Listeners cannot maintain concurrent models of multiple environments when they change rapidly
→ **strategy to boost externalization in hearing aids or VR.**

Mechanisms of **statistical and implicit learning**, as well as **adaptation to reverberation, like in speech perception**, likely play a role.

Thank you!

<http://pcl.upjs.sk>

Acknowledgement:

Barb Shinn-Cunningham and Matt Schoolmaster contributed to data collection and initial analysis.

(Schoolmaster, Kopčo, & Shinn-Cunningham, J Acoust Soc Am 113, 2285, 2003)