Effects of Context and Preceding Exposure on Distance Perception in Varying and Fixed Virtual Environments

Norbert Kopčo and Myroslav Fedorenko†

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

* pronounced KOH-shih-tse

† also at M. Dragomanov Ukrainian State University, Kyiv, Ukraine

Supported by Science Grant Agency of the Slovak Rep. VEGA 1/0350/22 and by European Union HORIZON-MSCA-2022-SE-01 grant agreement N° 101129903

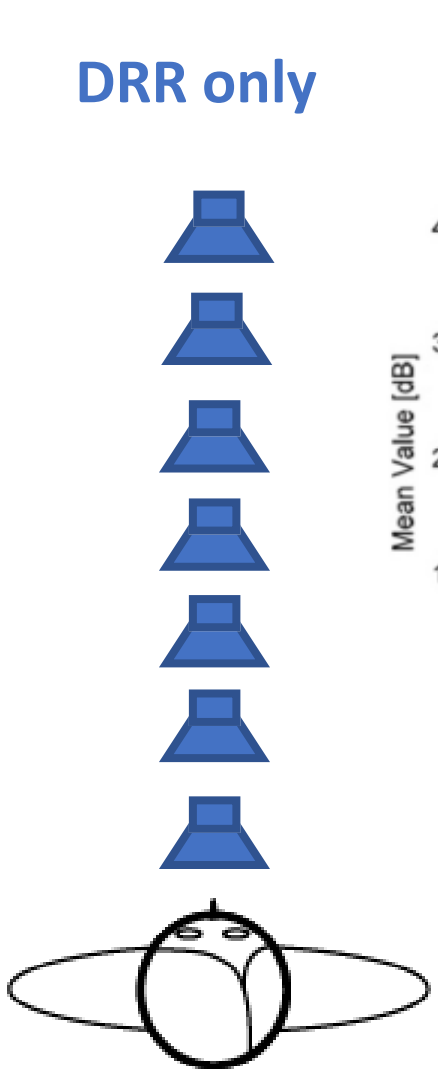
Introduction

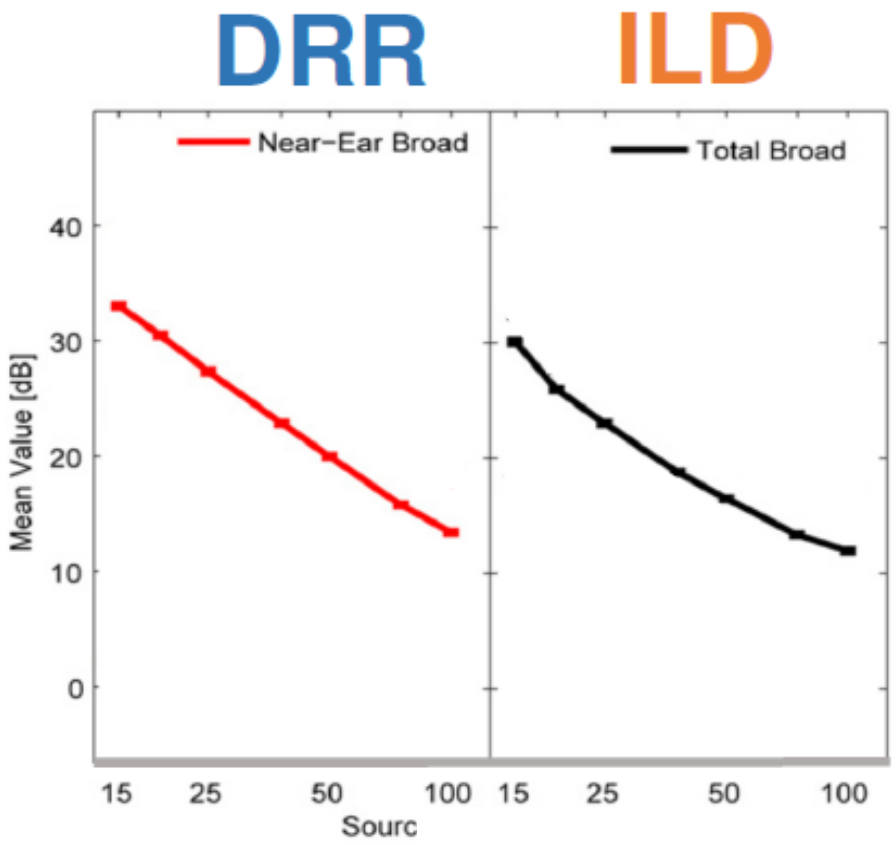
Level-independent distance perception possible for nearby sources (< 1 m)

- 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 previously presented stimuli (Doreswamy et al., 2019).





ILD + DRR

Introduction

In **reverberation** (but not in anechoic space), distance perception improves **spontaneously, without feedback or** any **training**, just by listeners actively performing the task in sessions with duration of **several hours** (even if split over multiple days).
(Shinn-Cunningham, 2000; Santarelli, 2000)

This **spontaneous learning** in a fixed room can strongly depend on **availability of cues** (e.g., level vs. DRR), especially **during initial exposure** to a given room (Hladek et al., 2013).

Current Study

In **virtual** and **mixed reality**, the presented **environments** can **change rapidly**.
How does **consistency of simulated environment** affect **distance perception** and the **spontaneous learning** processes?

1. How does **varying** the environment from **trial to trial** (vs keeping the environment consistent) influence distance perception?
   E.g. when listeners perform the task in 3 different virtual environments:
   Will they be able to concurrently maintain/tune to **3 separate model rooms**, or will they create **1 combined model**?

2. Does **initial exposure to in/consistent rooms** affect performance in both consistent and inconsistent contexts?
   E.g., if starting in consistent rooms means that listeners will learn characteristics of each room, will it transfer to better performance in inconsistent rooms in a later session?

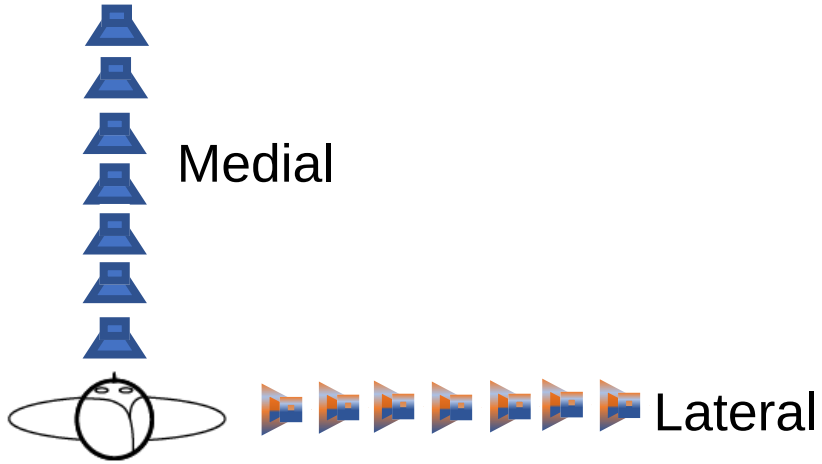
3. Is distance perception and spontaneous learning influenced by the **early reflections** when listener is near **the corner** of a room?

4. Is **spontaneous learning** of room-specific distance cues **inhibited** by room **inconsistency**?

Experiment in Virtual Environment

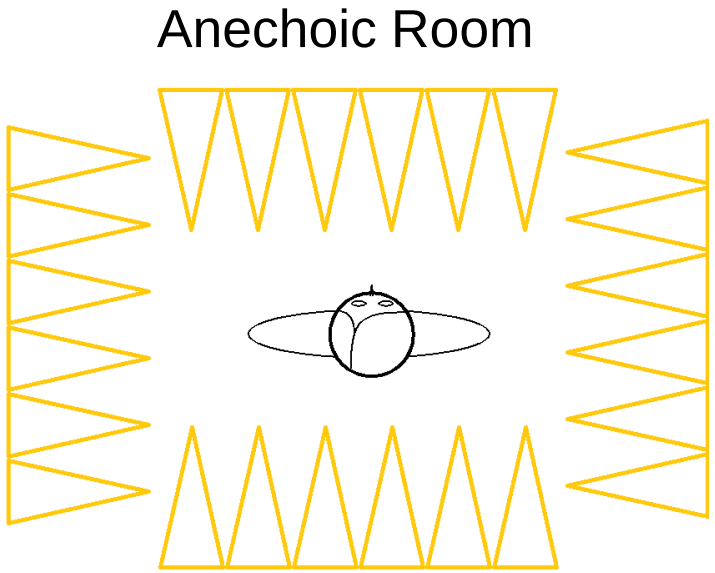
Stimuli:

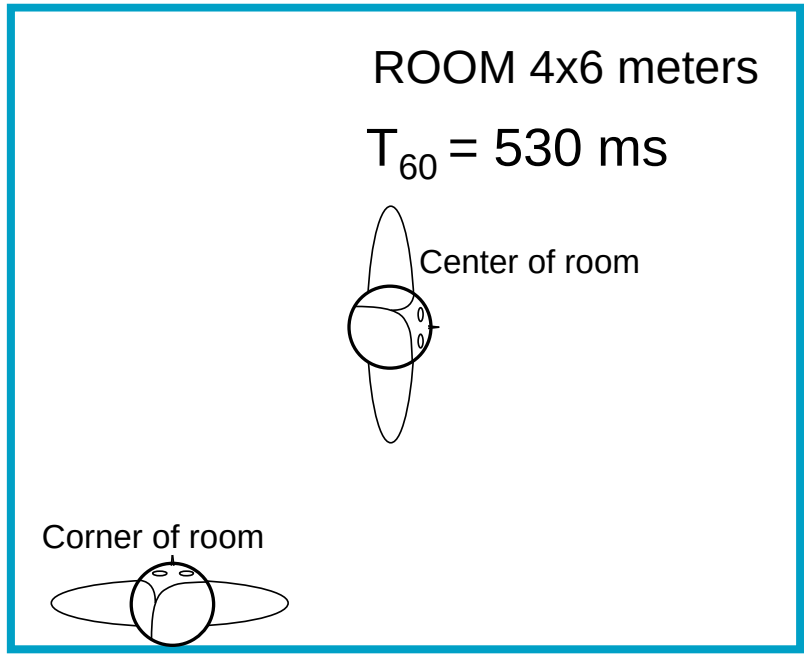
- five 150-ms-long pink noise bursts (30-ms gaps)
- roved by 15 dB (to eliminate level cue)
- 9 distances (15 to 170 cm, log spaced)
- 2 directions (medial and lateral)



Room conditions:

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



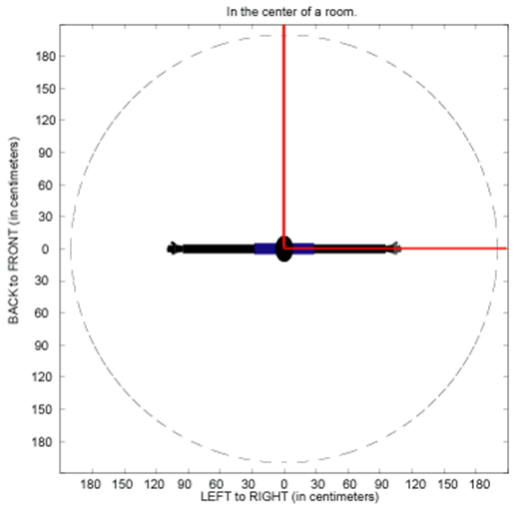


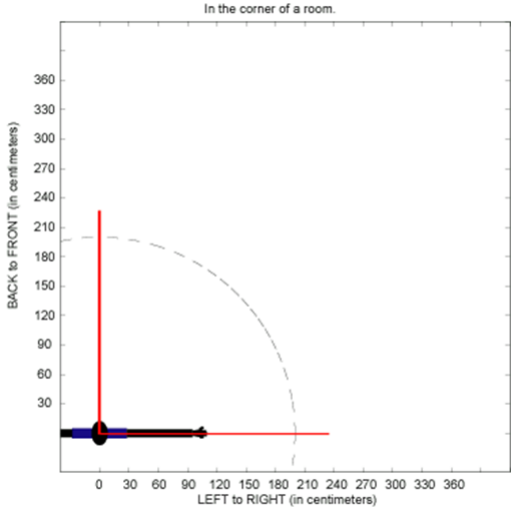
Experiment in Virtual Environment

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 (**contexts**): **FIXED** and **MIXED**
- session consisting of 6 **blocks**, each containing 8 **runs**
- each **run** had 45 trials which **held direction fixed**, only varying distance
- **FIXED** sessions: simulated room fixed within a **block**
- **MIXED** sessions: simulated 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: Fixed Room Context

InitFixed Group

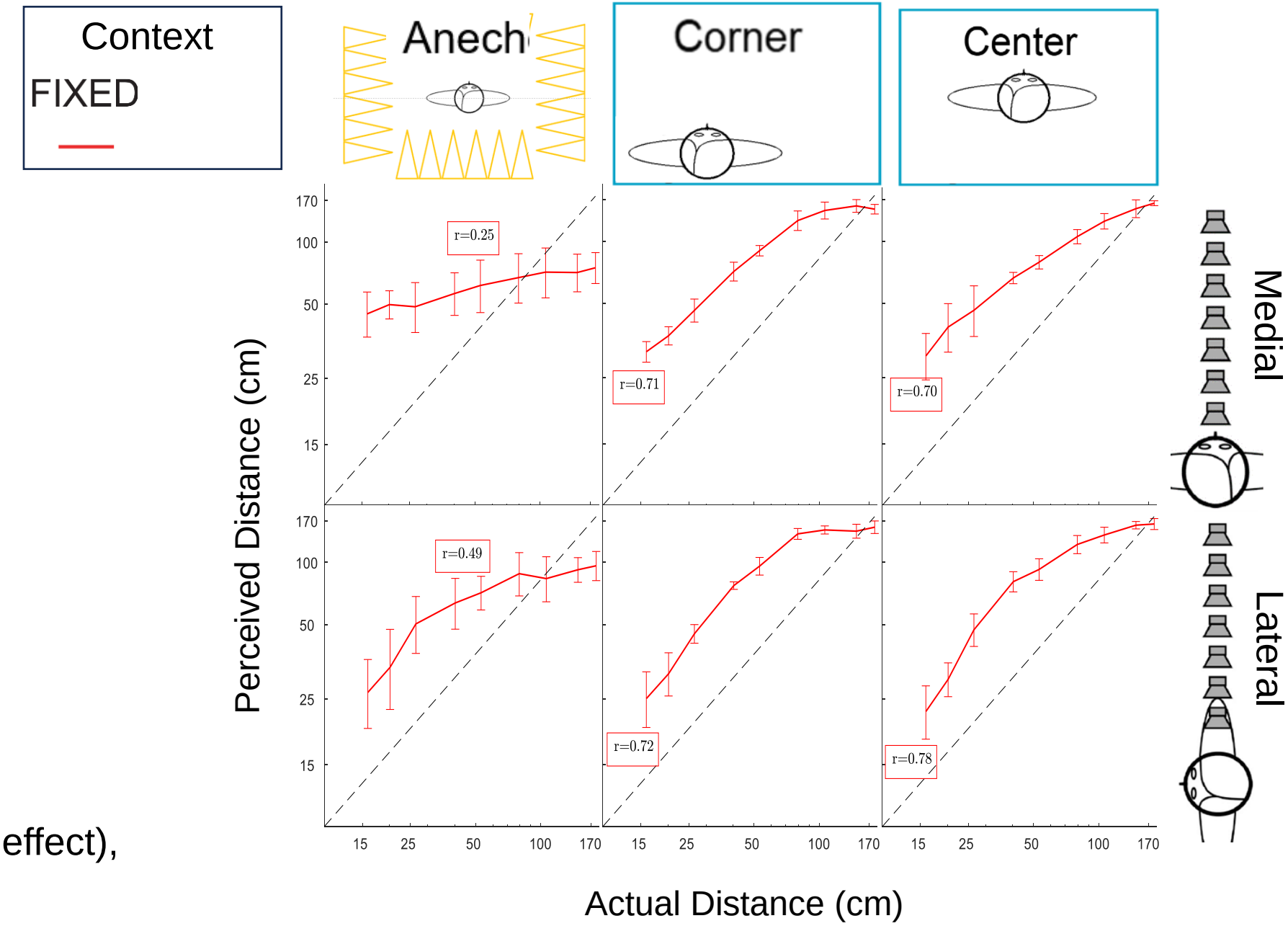
Session 1: room FIXED within a block.

Level-independent distance perception:

- better for lateral than medial sources
- better in room (CE/CO) than anechoic
- slightly better in CE vs CO

Nearby targets overestimated

Distant targets underestimated in AN, overestimated in CO (edge effect), accurately judged in CE



Results: Mixed (vs. Fixed) Room Context

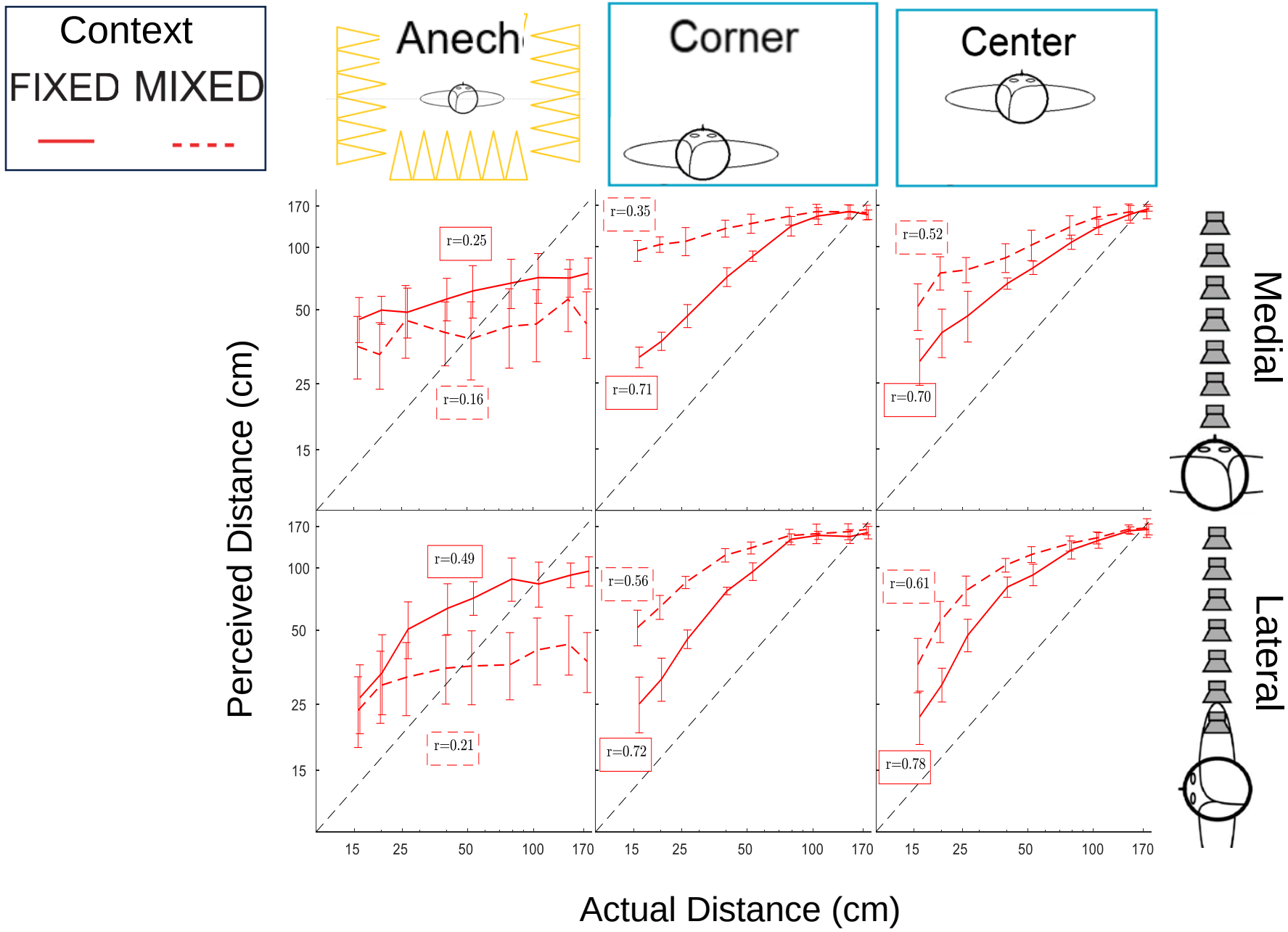
InitFixed Group

Session 2: room varying from trial-to-trial in block

Worse performance in all rooms & directions.

Bias induced by **mixed context** in all rooms, independent of direction:

- in AN, responses shifted closer,
- in CO and CE, responses shifted further away.



Group starting with MIXED context

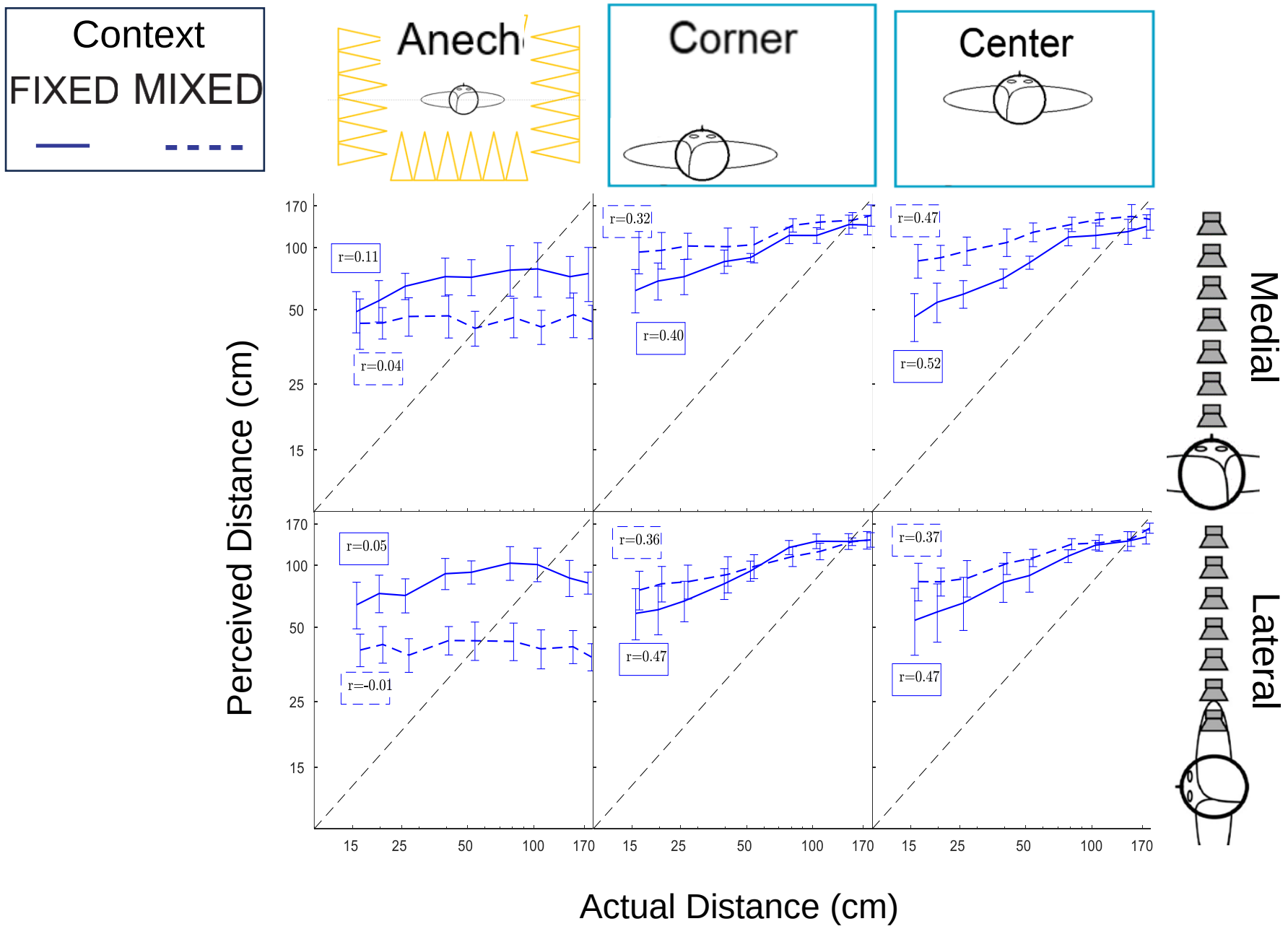
initMixed Group
- Session 1: MIXED
- Session 2: FIXED

Overall performance
- worse, especially for nearby & lateral sources

Effect of Mixed vs Fixed context:
- similar to initFixed gr.
- weaker mainly because the Fixed condition is worse

Bias effects not visible in corr. coef. *r* (e.g., AN)

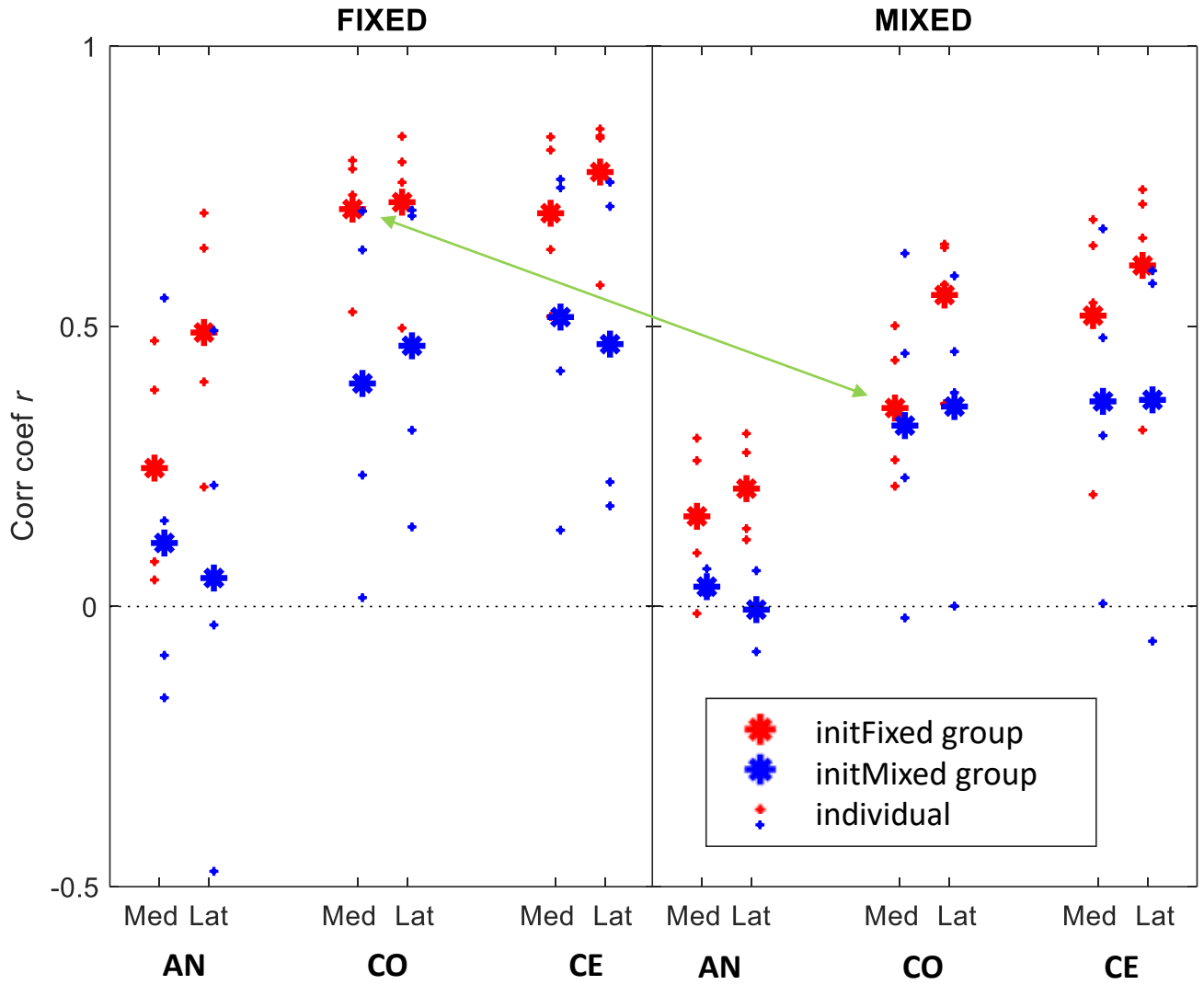
Again, AN < CO < CE



Summary of Results using Corr. Coef. *r*

Performance tends to be better for:
- initFixed group,
- Fixed context,
- lateral direction,
- room (CE>CO>AN)

Complex dependences between factors (4-way interaction: p = 0.034):
- initMixed group:
  no effect of direction
- initFixed group:
  effect of context (Mixed – Fixed):
  - varies with room and direction
  - is largest for **CO Med**



Not only the current context, but also initial/preceding context affects performance.

Early reflections modulate effect of initial context for med sources.

Learning within a run: $r_{2nd\ half} - r_{1st\ half}$

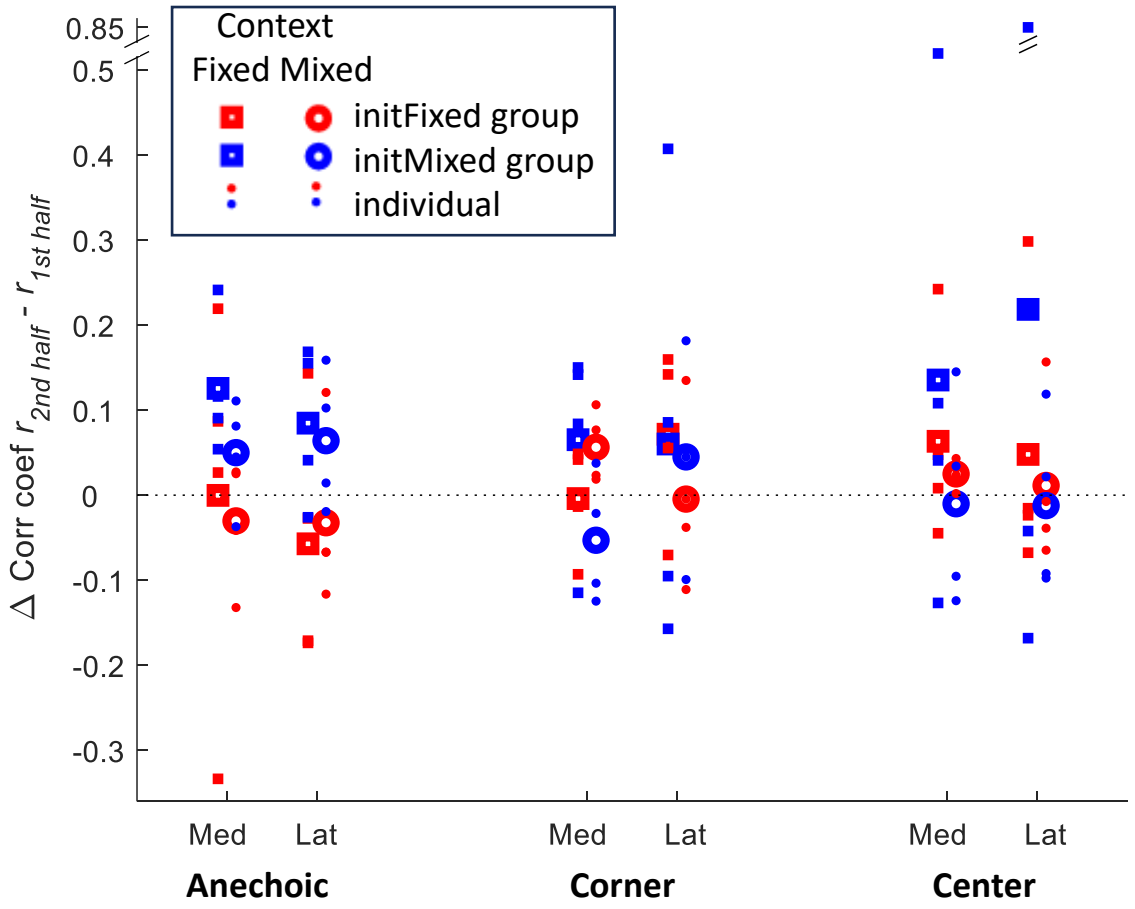
In room (CE/CO):

- **improvement** in **Fixed** context (□, □) but **not** in **Mixed** context (○,○)
- except in **CO Med** initFixed, where improvement is in mixed (○) but not in fixed (□) context (4-way interaction, p = 0.011).

In AN, initFixed do **not improve**, while initMixed **improve** in both **Fixed and Mixed** contexts (p = 0.03)

Spontaneous learning within run influenced by presence of reverberation (AN vs CE/CO).

Early reflections influence the learning of reverberation cues for medial sources.



Discussion and Conclusions

1. Fixed rooms:

   Performance better for room than AN and, sometimes, for lat. than med. sources. → **Both DRR and ILD cues are used by listeners when available.**

   Performance slightly worse in CO than CE.
   → **Early reflections in CO are detrimental for distance judgments.**

2. Context effect:

   Mixing rooms from trial to trial induces biases: underestimation in AN and overestimation in CE/CO. → **Listeners cannot separately process distance information from different rooms on trial-by-trial basis.**
   Biases consistent with listeners creating a single DRR-to-distance mapping in Mixed context, since in such 1 combined room model:
   - AN ~ very large DRR → percepts biased closer,
   - CE/CO ~ smaller DRR → percepts biased away from listener.

Discussion and Conclusions (cont.)

3. Initial/Preceding context:

Starting in **Mixed** context tends to cause more deterioration re. **starting Fixed**.
However, the effect is complex:

- initMixed group performed equally for lateral and medial sources.
  → **If starting in a Mixed context, listeners did not benefit from ILD cue for lateral sources in the Mixed or in the Fixed context, even though in the Fixed session (performed as 2nd) all the cues were consistent.**

- initFixed group can benefit from ILD, but the effect of context (Mixed vs Fixed) varied with room and direction (largest for CO Med).
  → **How the cues are combined and weighted depends on the current context, the initial context, which cues (ILD/DRR) are available, as well as on early reflections.**

Factors that determine these complex interactions need further examination.

Can this result be used to enhance externalization?

Discussion and Conclusions (cont.)

4. Learning within a run (1st vs. 2nd half):
   Room (CE/CO):
   Learning observed in most conditions in fixed context but not in mixed context (except in CO Med initFixed, in which improvement was in mixed but not in fixed context). → **Stable reverberant environment is required for spontaneous learning / tuning of DRR-to-distance mapping. Early reflections can interfere with the learning.**

   AN:
   Unexpectedly, initMixed (but not initFixed) group improved in both Fixed and Mixed sessions. → **Possibly an effect of confusion: initial exposure to mixed context causing very poor initial performance from which subjects gradually recover during the rest of the study.**

Next steps: analyze learning over blocks.

Thank you!

Acknowledgement:

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

The 5th **workshop** on

Cognitive neuroscience of auditory and cross-modal perception

15-17 April 2024, Košice (pronounced KOH-shih-tse), Slovakia

https://pcl.upjs.sk/workshop-2024/

Main focus: ***spatial audio virtualization and gamification for hearing assessment and enhancement***.

Guest speakers and SAV consortium speakers (tentative):

G. Christopher Stecker, Boys Town National Research Hospital, United States
Antje Ihlefeld, Meta, “Spatial Audio” (tentative)
Eleni Vlahou, University of Thessaly, Greece
Mathieu Lavandier, ENTPE – University of Lyon “Modelling speech intelligibility in noise: from differences in SRTs to full psychometric functions?”
Robert Baumgartner, Austrian Academy of Sciences, “Short-term adaptation of spatial hearing.”
Bernhard Laback, Austrian Academy of Sciences, “Dilation of Auditory Space by Short-Term Context”
Piotr Majdak, Austrian Academy of Sciences, Austria
Anja Pahor, Univerza v Mariboru, “Development and validation of mobile measures of executive function.”
Frederick Gallun, Oregon Health and Science University, “Does Psychoacoustics Have to be Boring? Exploring Gamification of Auditory Testing”
Jorg Buchholz, Macquarie University, Australia
Jyrki Ahveninen, Mass General Research Institute / Harvard Medical School, United States
Virginia Best; Boston University, United States
Aaron Seitz, Northeastern University, “New games to train speech in competition; from psychoacoustics to music.”
Norbert Kopco, P. J. Šafárik University in Košice, “Other topics in SAV: Attention and distance in real and virtual environments”