Probing the structure of working memory representations in vision & audition

Abigail **Noyce**, Raina **Alam**, Will Yuhang **Li**, Eli **Bulger**, Stephanie **Bronfman**, Kammiee **Ardo**, Barbara **Shinn-Cunningham**

Carnegie Mellon University; University of Wisconsin; University of Pittsburgh

anoyce@cmu.edu

Patterns within and among scene elements profoundly affect perceptual organization, as in Gestalt processing [1], gist estimation and ensemble coding [2,3], or auditory scene analysis [4].

Perceptual organization of scene elements into objects affects memory representations, as in chunking [5], structured memory representations [6], or category labels.

Similarity among memoranda affects memory representations, as in false memory [7,8] or exemplar models [9].

Memory structures afford different kinds of access for retrieval, such as single-item, whole-display, or serial order [10].

Does **perceptual organization** shape **WM structures** similarly across **vision** and **hearing?**

Study Architecture

All studies compared two tasks in which participants were asked to remember a short sequence. Stimuli varied across studies, but were never mixed within a trial. All comparisons were within-subject.

Item recognition task

Sternberg-type old-new judgment; 50% probability "old". Measures WM access to **individual stimulus items**.

Serial order task

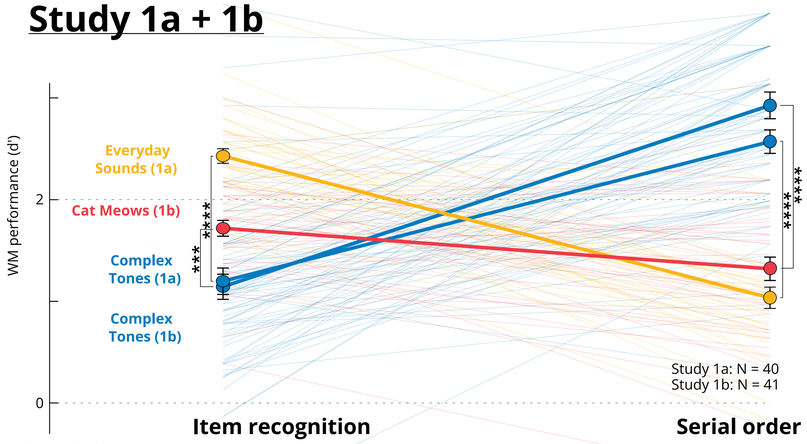
Serial order same-different judgment; 50% probability "same". Measures WM access to **sequence as a whole**.

[1] Baker, ed. (2012). *The Oxford Handbook of the History of Psychology: Global Perspectives.*
[2] Xu (2002). Limitations in object-based feature encoding in visual short-term memory. *JEP: HPP.*
[3] Oliva & Torralba (2006). Building the gist of a scene: The role of global image features in recognition. *Progress In Brain Research.*
[4] Bregman (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.*
[5] Miller (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psych. Review.*
[6] Brady & Tenenbaum (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psych. Review.*
[7] Deese (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *JEP.*
[8] Roediger & McDermott (1995). Creating false memories: Remembering words not presented in lists. *JEP:LMC.*
[9] Kahana & Sekuler (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research.*
[10] Noyce, Varghese, et al. (2024). Perceptual organization and task demands jointly shape auditory working memory capacity. *JASA Express Letters.*

Auditory Working Memory

Temporal grouping of discrete stimuli is natural in audition; however, some stimuli are more "groupable" than others.

Study 1a + 1b



Will Li & Eli Bulger

Complex Tones

262 - 623 Hz in semitone steps

High inter-item similarity

Same attributed source

Musical pitch **present**

Cat Meows

9 vocalizations (from Noyce 2017)

Moderate inter-item similarity

Different attributed source

Musical pitch **absent**

Everyday Sounds

15 samples (from ESC-50, Piczak 2015)

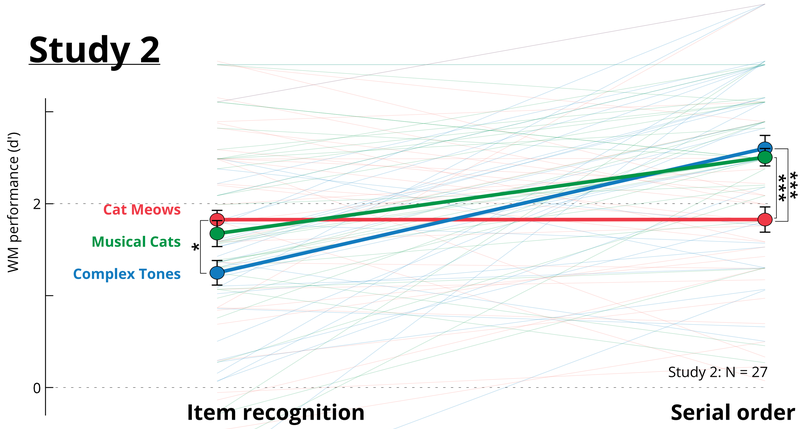
Low inter-item similarity

Different attributed source

Musical pitch **absent**

All stimuli were 300 ms duration and loudness normalized (RMS). Stimulus types were never mixed within a trial.

Study 2



Steph Bronfman & Kammiee Ardo

Complex Tones + **Cat Meows** =

"Musical Cats"

Each cat exemplar monotonized and pitch-adjusted (Praat).

Moderate inter-item similarity

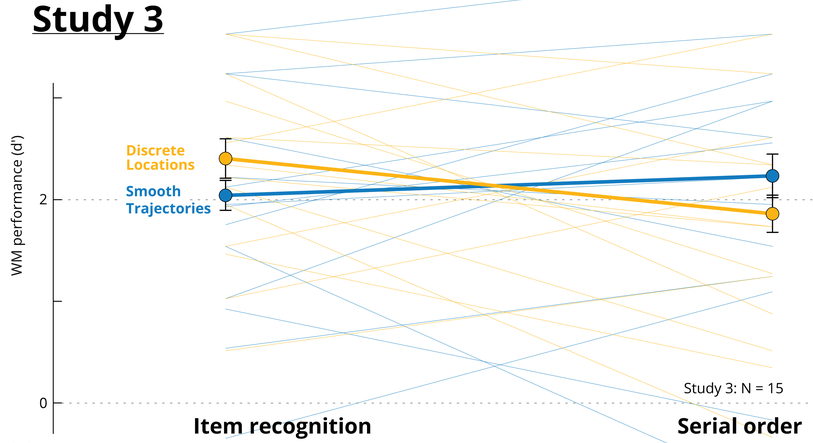
Different attributed source

Musical pitch **present**

Visual Working Memory

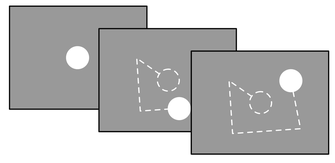
In vision, sequence representations are more accessible for smooth trajectories than discrete locations, but the effect is smaller.

Study 3

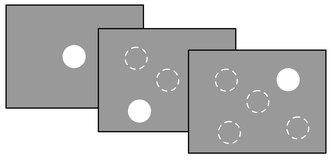


Raina Alam

Smooth Trajectories



Discrete Locations



Conclusions

Item recognition depends on "ungrouped" memory representations, with elements stored as single items.

Serial order depends on structured memory representations, which are facilitated by perceptual grouping.

Similar effects across vision and hearing. Stimulus by task interaction was observed in both visual and auditory WM.

Temporal structure is weaker in visual than auditory WM. Hearing has an affinity for temporal and sequential cognition.

Thank you to Sahil Luthra for advice on measuring similarity, and to Cherie Hua for early work on a precursor to this project.

Funded in part by the Office of Naval Research (MURI N00014-19-12332), by an NSERC postgraduate fellowship to EB, by a CMU HURAY grant to KA, and by NIDCD R25 DC-020922 funding to SB.

Poster Link