Effect of learning on Audio-Spatial Working Memory

Sanchit Aggarwal, Koustav Ghosal, Pulkit Singhal, Priyanka Srivastava

International Institute of Information Technology, Hyderabad, India

{sanchit.aggarwal@research.,koustav.ghosal@research.,pulkit_s@students.,priyanka.srivastava@}iiit.ac.in

CogSci Lab

IIIT-H

Abstract

We aimed to investigate the relationship between modes of training and audio-spatial working memory. Modes of training were manipulated by unsupervised with no error feedback and supervised with error feedback. Furthermore, supervised was divided into audio and visual error feedback. An experiment was conducted in three phases, which consisted of working memory task before and after various modes of training. Set size of working memory was varied from 2-8. The pilot result shows a trend that training enhances the performance of WM task.

Introduction

- Navigation in unfamiliar environments is a major problem encountered by the visually challenged and has been the focus of many assistive technology based studies [1, 2].
- Auditory cues used to describe the spatial map of walkable region. We studied the trade-off between the amount of information being passed and ability to store it i.e. ***audio-spatial WM***.
- ***Objective*** - To study the effect of learning on WM performance.
- We developed ***ViShruti***, a simulation of indoor environment where walkable path is unknown.

Experiment

Participants

- 16 Male and 8 Female, Mean age=23 with normal vision.

Apparatus and Stimuli

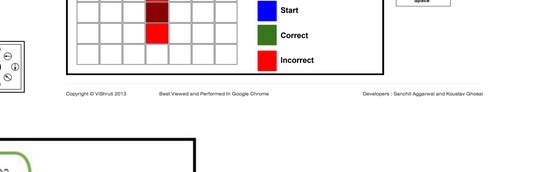
- Tones of frequency 440Hz, 880 Hz, and 1760 Hz (spanned in 3 octaves).
- Combined with pan to produce 8 different audio tones. E.g. ***North*** = Higher pitch (1760 Hz) on both the channels, ***SouthEast*** = Low pitch (440 Hz) on right channel.

Design and Procedure

- 3-Phase experiment: ***WM-Training-WM*** to observe the difference between the WM performances due to training.

a) Mixed design: 4 training modes as between group x 7 (WM set size varying from 2 - 8) as within group.

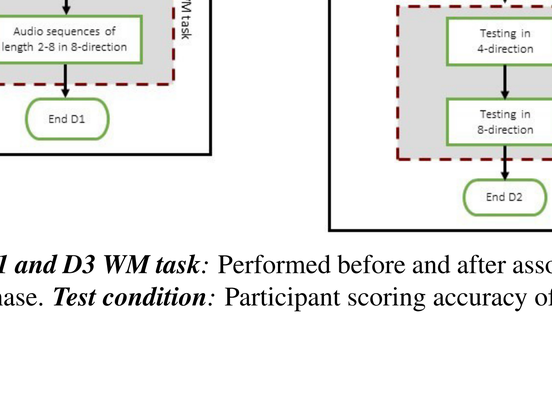
b) No-training introduced as a control group.



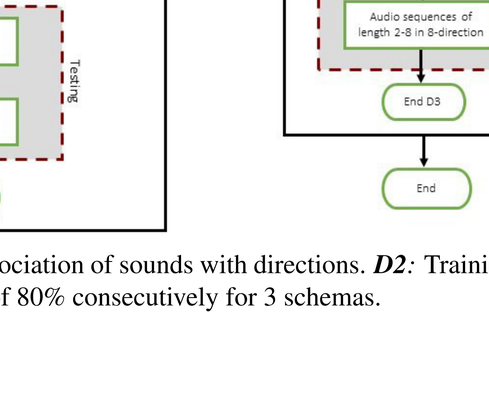
Working Memory Task

- Tones represents unit steps of the path. A trial consisted of list of tones varying from 2-8 set size. The task was to ***recall*** the list of tones in order.
- Keys for 4-directional cues: $\rightarrow$, $\leftarrow$, $\uparrow$, $\downarrow$ mapped with respective directions. Keys for 8 directional cues: NW: 7, N: 8, NE: 9, W: 4, E: 6, SW: 1, S: 2, SE: 3.
- Presented 5 schemas, 7 trials of length (2-8) per schema. ***Inter Stimuli interval:*** Determined by staircasing. ***Inter trial interval*** = 500 ms.
- ***Performance measure*:** Number of trials required for training and recall accuracy.

Illustration of Working Memory phase.



How ViShruti Works.



D1 and D3 WM task*:* Performed before and after association of sounds with directions. ***D2****:* Training phase. ***Test condition****:* Participant scoring accuracy of 80% consecutively for 3 schemas.

Results

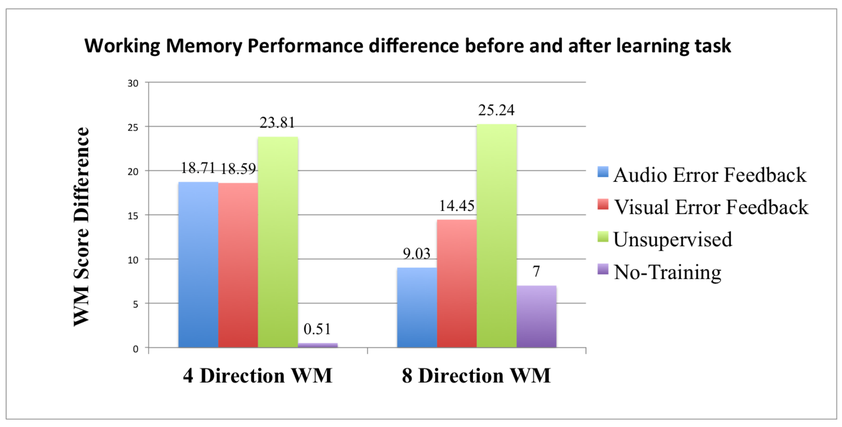
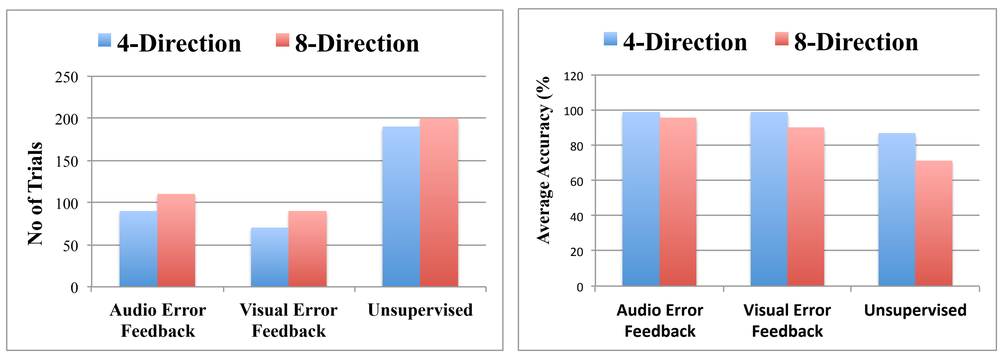
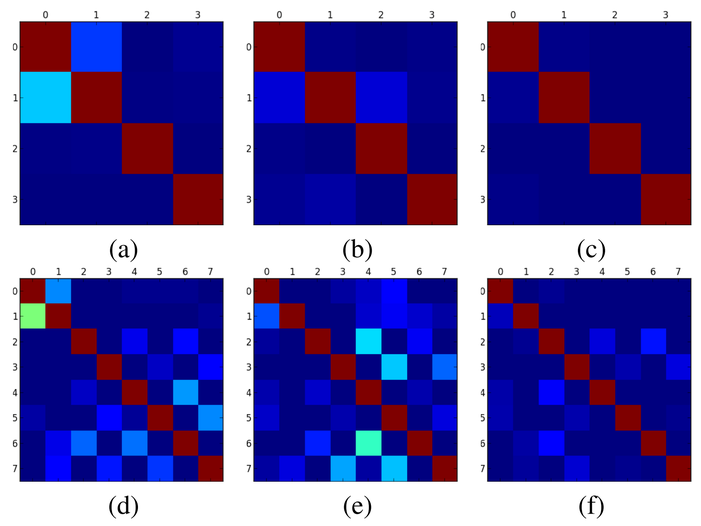


Fig. 1. Difference in performance in the first and third phase WM task. There is improvement in all the modes of training. The best performance is observed in unsupervised.



(i) (ii)

Fig. 2. Training Task Results: i) No. of Training trials vs. Training Modes. ii) Average Testing Accuracy vs. Training Modes. Blue: 4-directions, Red: 8-directions. (i) and (ii) depicts visual fedback as more efficient trend than audio feedback. Also in unsupervised, absence of feedback results in more number of training trials and lesser accuracy.



(a) (b) (c)

(d) (e) (f)

Fig. 3. Confusion Matrices. Row 1) 4-directions, Row 2) 8-directions. Column 1) Unsupervised, Column 2) Audio Error Feedback, Column 3) Visual Error Feedback. In each 8×8 matrix, rows 0,1,2,3,4,5,6 and 7 represent N, S, E, W, NE, NW, SE and SW respectively. For 4×4 matrices, rows 0-3 represent N, S, E, W. Confusions can be seen in unsupervised training which were lesser in audio error feedback and least in visual error feedback. Also, participants were confused between North and South directions in case of 4-direction in (a) and (b) respectively. Additionally, in (d) and (e) the confusing directions were North-East and South-East.

Conclusion

Based on what has been observed, we came up with three main trends.

- The methods of training help in constructing the spatial map, which enhances the performance of audio spatial working memory.
- Visual Error Feedback and Audio Error Feedback are better training methods.
- Combination of Pan and Pitch can be used effectively for representing the different directions.

Future Work

The limitation of the study is its small sample size due to which it is hard to generalize. We would like to continue the study with visually-challenged. We also plan to study the effect of sound cues in presence of noise. We plan to build a prototype of *ViShruti*, where the visual information about the schema of the indoor environment would come from a mounted camera and it will be processed to segment out walkable regions using machine learning and computer vision.

References

[1] Wilson, Jeff, *et al.* "Swan: System for wearable audio navigation." In Wearable Computers, 2007.

[2] Brock, Michael, and Per Ola Kristensson. "Supporting blind navigation using depth sensing and sonification." In Proc. Pervasive and ubiquitous computing, 2013.

[3] Dingler, *et al.* "Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech." In Proc. International Conference on Auditory Display, Paris, France, 2008.

[4] Tran *et al.* "Evaluation of acoustic beacon characteristics for navigation tasks." Ergonomics, 2000.