Introduction To Cognitive Science

Project :

Audio Based Navigation System: Learnability & Effect Of Working Memory

Sanchit Aggarwal
Koustav Ghosal
Rishav Goyal

Mentor :
Dr. Priyanka Srivastava

Introduction

Objective : A system that helps visually challenged people to navigate using audio cues.

Why audio over videos for normal people ?

- Temporal Nature Of Data : No ‘Blur’ like in videos. [1]
- Independent of Orientation : We don’t have to fix our ears to a direction to listen to a sound. [1]

Other Applications : [1]

Fire Fighters

Mines

Geiger Counter

Pulse Oximeter

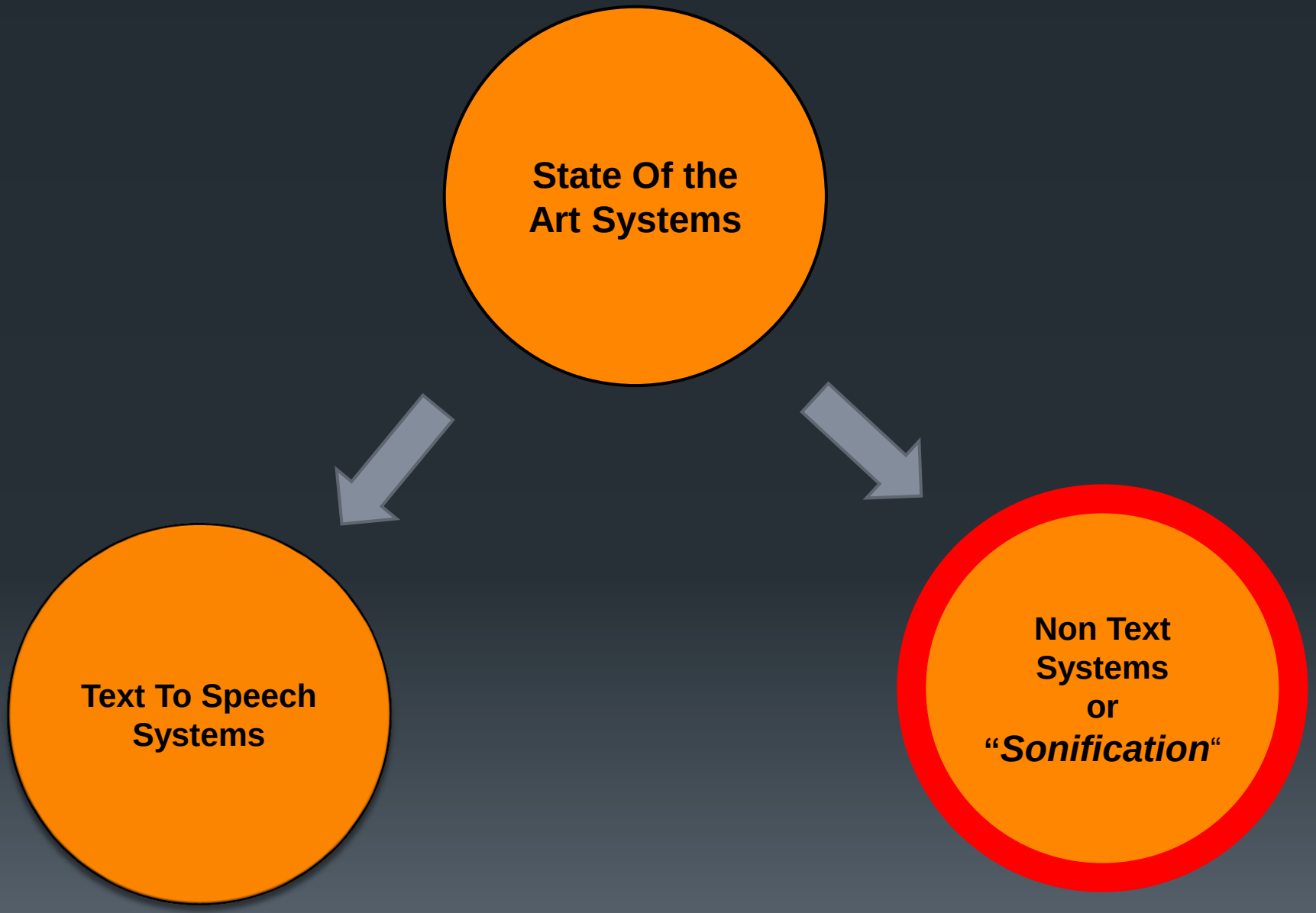
Voyager 2 Space Mission

There was a problem with the spacecraft as it began its traversal of the rings of Saturn.

The controllers were unable to pinpoint the problem using visual displays, which just showed a lot of noise.

When the data was played through a music synthesizer, a "machine gun" sound was heard during a critical period, leading to the discovery that the problem was caused by high-speed collisions with electromagnetically charged micrometeoroids

Existing systems are classified as



Sonified vs Text to Speech Systems

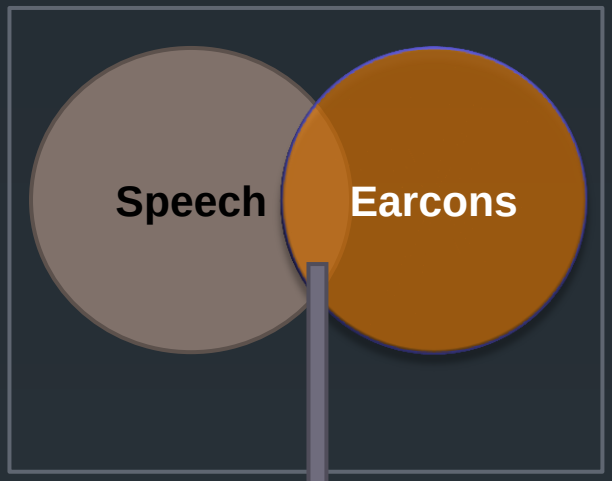
Language Independent
Speech beacons are harder to localize in a virtual environment than non-speech beacons. [2]
Users also give speech beacons low ratings for quality and acceptance [2].
Large amount of information . limited human speech processing capacity [3]
It is also difficult to use a speech-based interface for navigation and carry on a conversation at the same time [4]
Further, spoken messages in such a system are each generally more than a second long, so the system is often talking [5]
Reduced Cognitive Load [6]

Issues In Sonification

By its very nature, Sonification is interdisciplinary, integrating concepts from human perception, acoustics, design, the arts, and engineering. [1]

How can the mapping be done ? [7]

Earcons
Pan
Volume
Pitch
Timbre
etc…



How can these factors be effectively used in **navigation** ?

Spearcons

What have we done ?

Task A :

A detailed study of the Sonification principles, its advantages, disadvantages. We did a survey of the existing systems . The comparative study was on the basis of learnability, cost and efficiency.

We also tried to explore the scope of Sonification in Computer Vision and how these two can together be used to build an effective Navigation System

Task B :

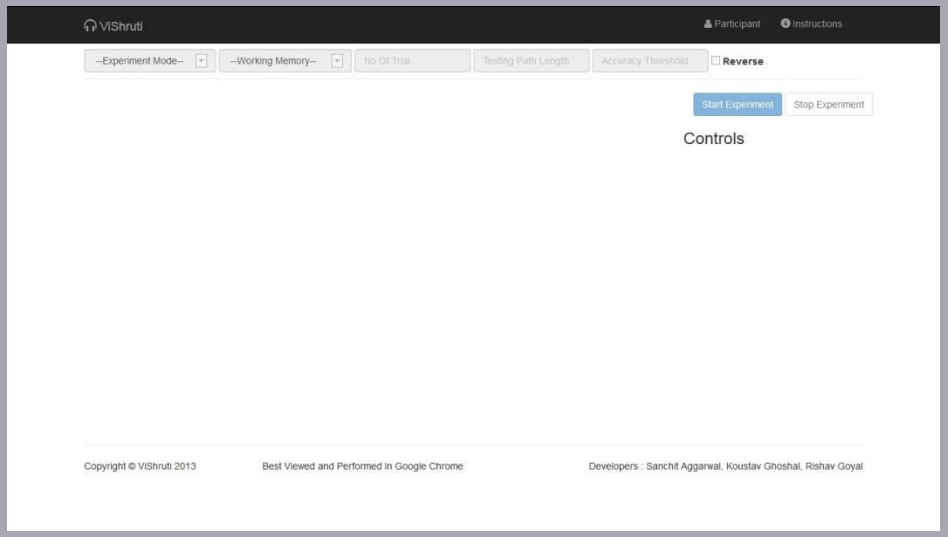
Effect of pan and pitch as sound parameters during navigation. The **efficiency and learnability** has been our area of focus.

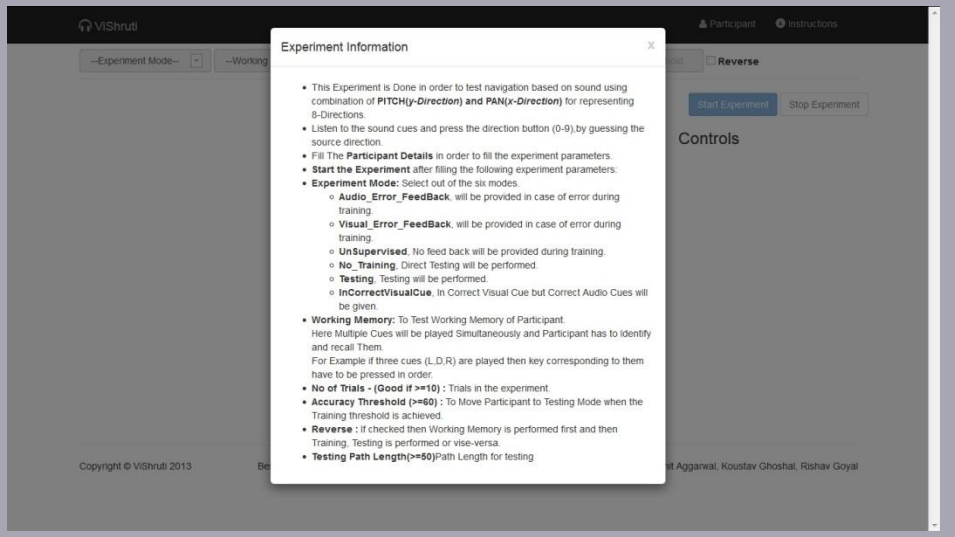
Task C :

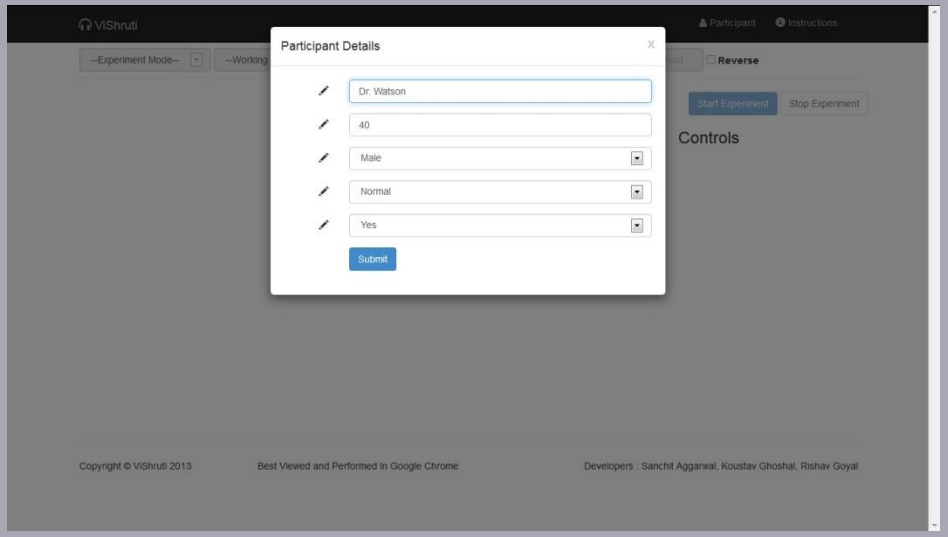
Effect of **working memory** on navigation task. We have tried to find a proper **Sonification-Pattern / Cue Length** based on the capacity of working memory.

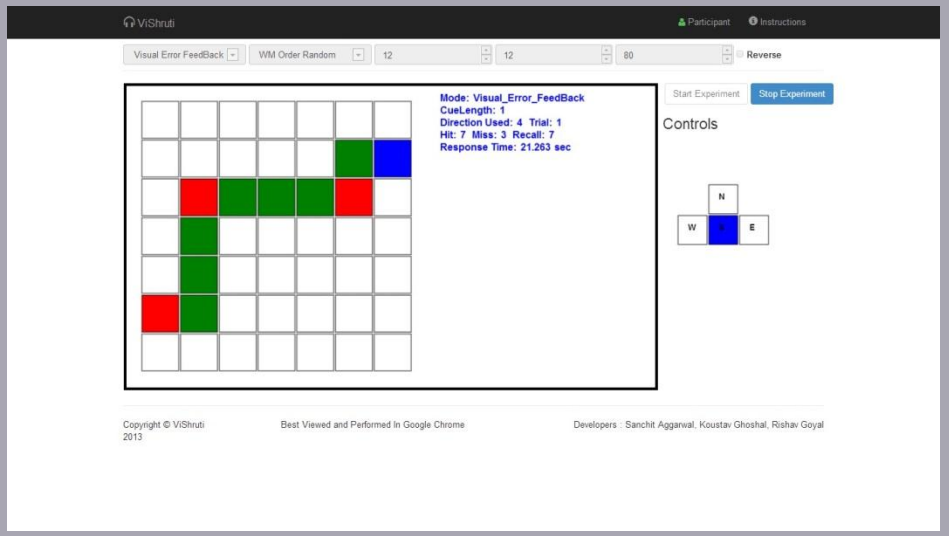
ViShruti

Technology Used : JavaScript, HTML5 , Riffle









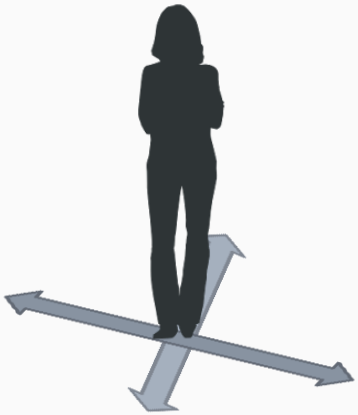
Task A : Learning ability of sound parameters pan & pitch for navigation of visually challenged people.

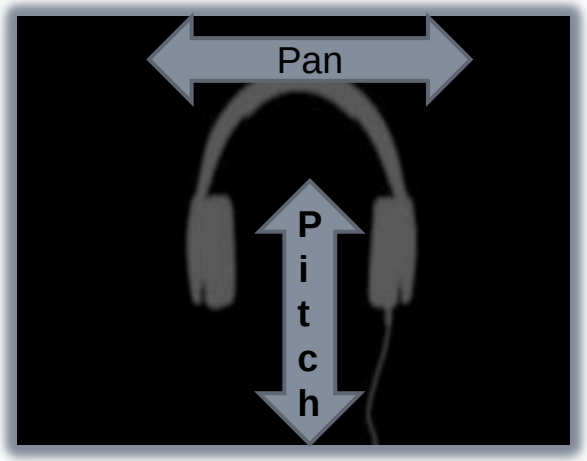
Simulation of real time navigation in two dimensions, in 4 and 8 directions.

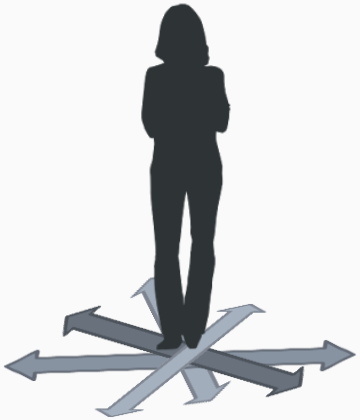
Questions:

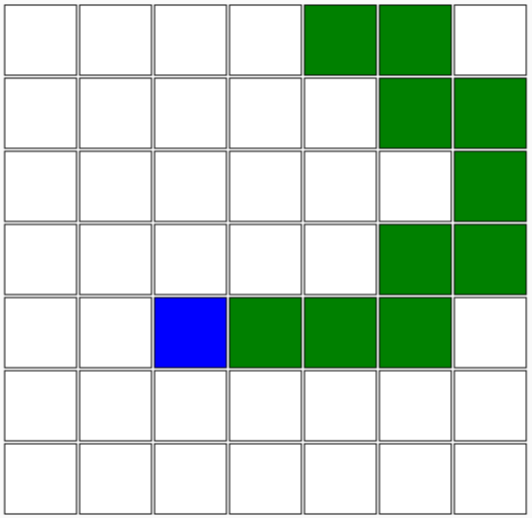
1. Do the different modes of training facilitate learning ?
2. What is the best/fastest way to train the user for this system?
3. What are the most common confusions made ?
4. Adaptability of experts & non-experts of Music to the system.
5. Adaptability of normal and visually challenged people to the system

How ViShruti Works





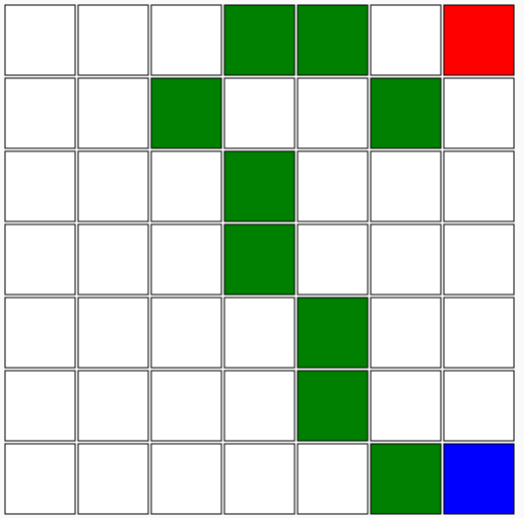


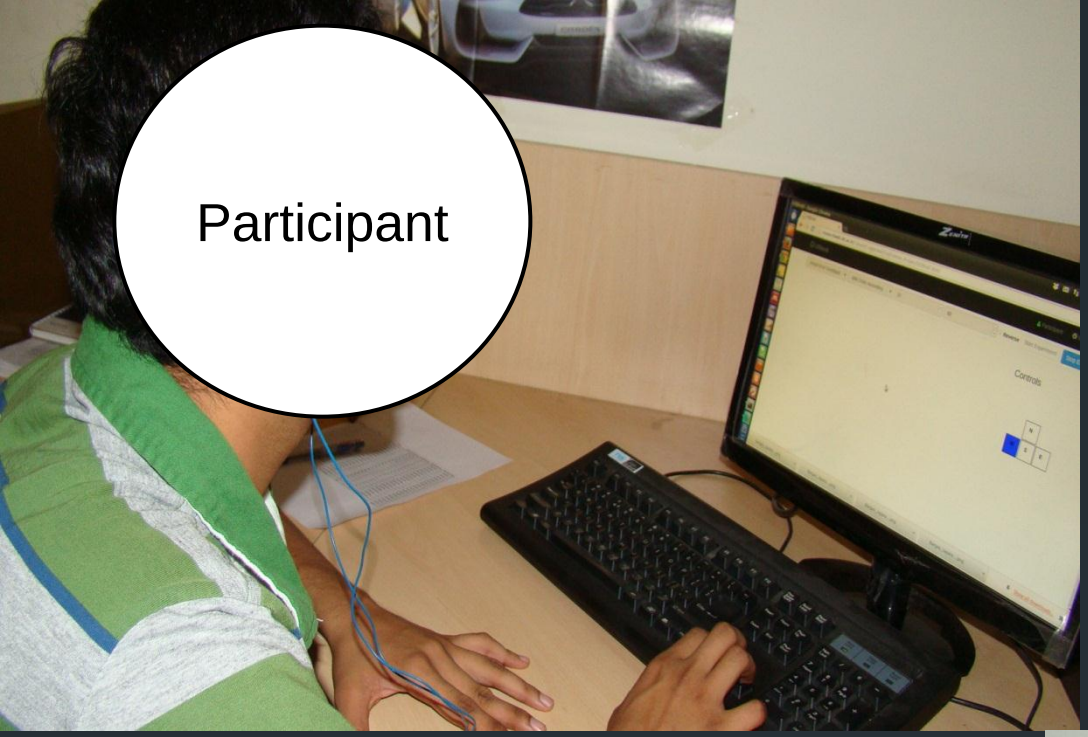


Start

Incorrect

Correct





Training

Trials : Maximum 10 or
Accuracy > 80 %
Path length : 10 units
Audio Error,
Visual Error,
Unsupervised &
No Training
Both 4 & 8 Directions

Testing

Trials : 4
Path Length : 10 +50*3
Both 4 & 8 Directions

<u>Subject Details :</u>

A & B : Unsupervised
C & D : No Training
E & F : Audio Error
G & H : Visual Error

Q 1 : Do these different modes of training facilitate learning ?

4-Directions

Mode Of Experiment	Average No. Of Trials Taken (80 % Accuracy)	Average Training Accuracy	Average Training Response Time	Average Test Accuracy	Average Testing Response Time
Audio – Error Feedback	1.5	82.08	5.17	81.6	2.14
Visual – Error Feedback	1	95	3.14	94.1	2.16
Un - Supervised Training	5.5	76.36	2.45	80.75	2.22
***No Training**	N/A	N/A	N/A	91	2.17

* No Training was done with musically trained people

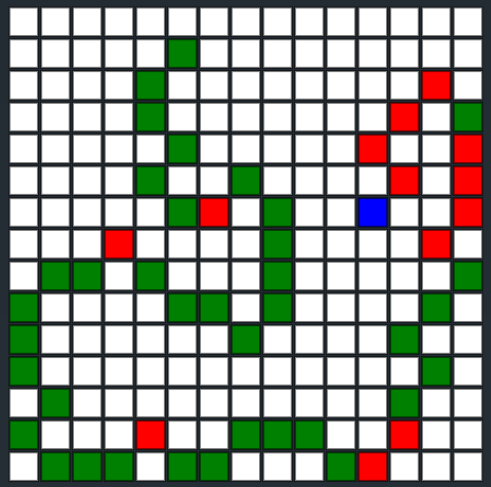
Q 1 : Do these different modes of training facilitate learning ?

8-Directions

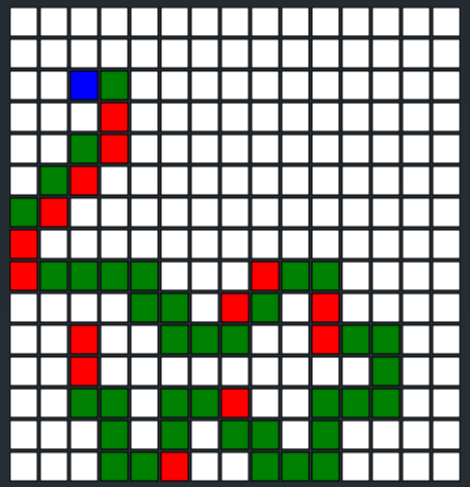
Mode Of Experiment	Average No. Of Trials Taken (80 % Accuracy)	Average Training Accuracy	Average Training Response Time	Average Test Accuracy	Average Testing Response Time
Audio – Error Feedback	10	59	2.817	62	2.27
Visual – Error Feedback	1	85	2.93	76.4	2.34
Un - Supervised Training	10	39.5	2.74	47.75	2.51
***No Training**	N/A	N/A	N/A	61.75	3.01

* No Training was done with musically trained people

Learning while navigating in unsupervised learning :

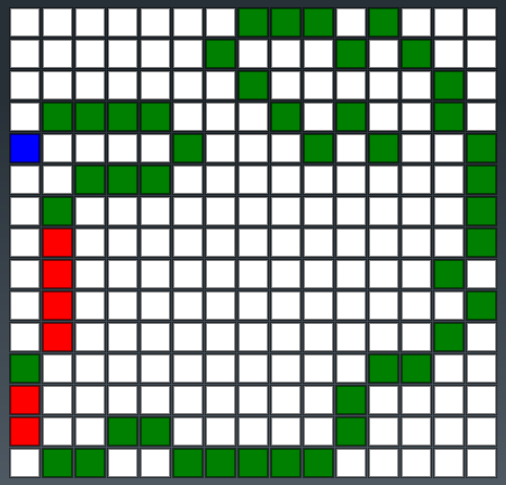


Participant A

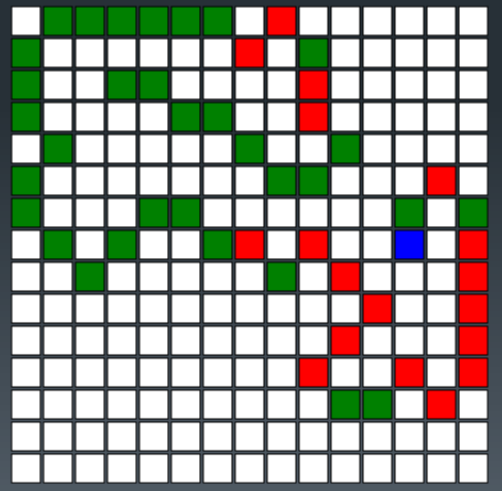


Participant B

Navigation without training for musically trained people

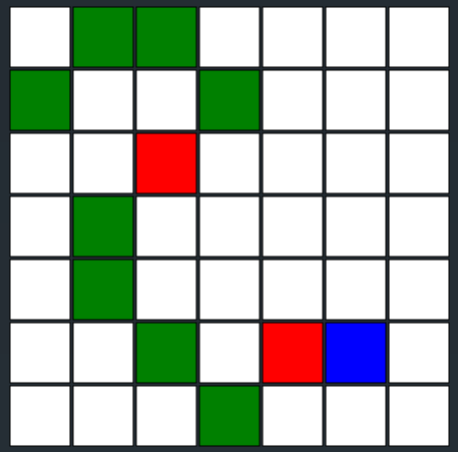


Participant C

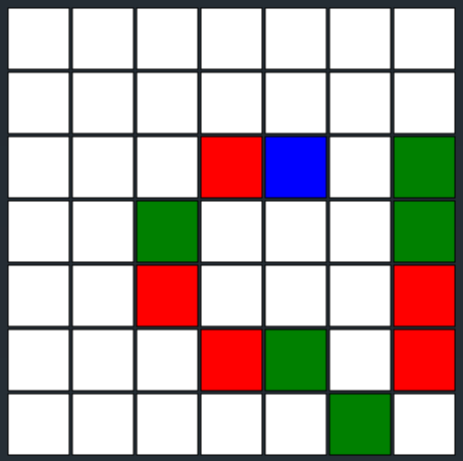


Participant D

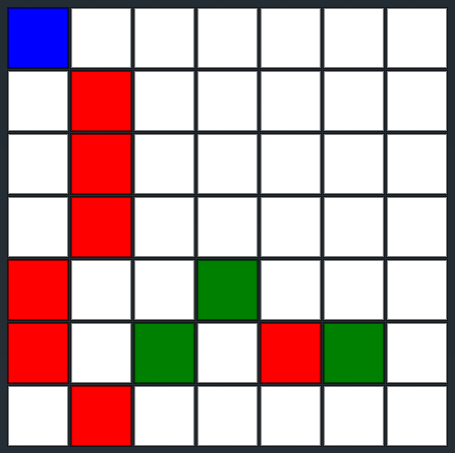
Different Training Methods : A Comparative Study



Visual Cue

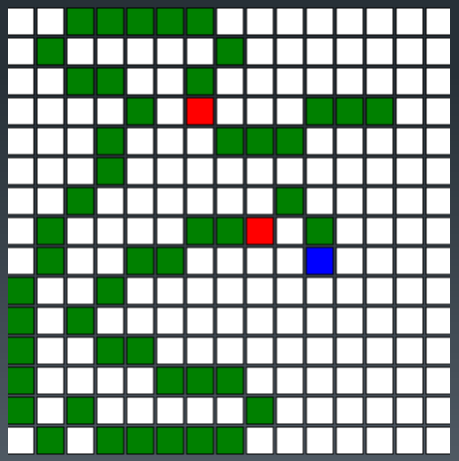


Audio Cue

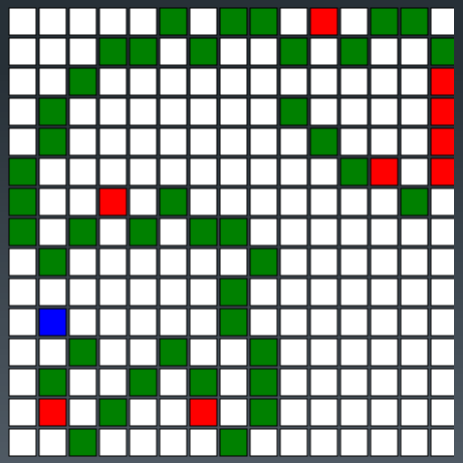


Unsupervised

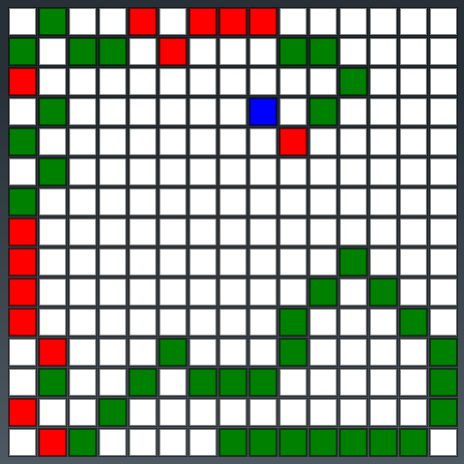
Testing with different training



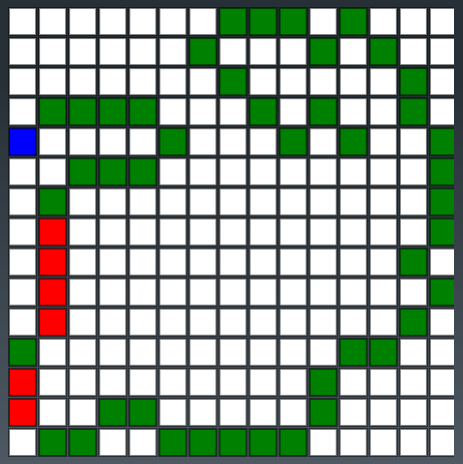
Visual Cue



Audio Cue

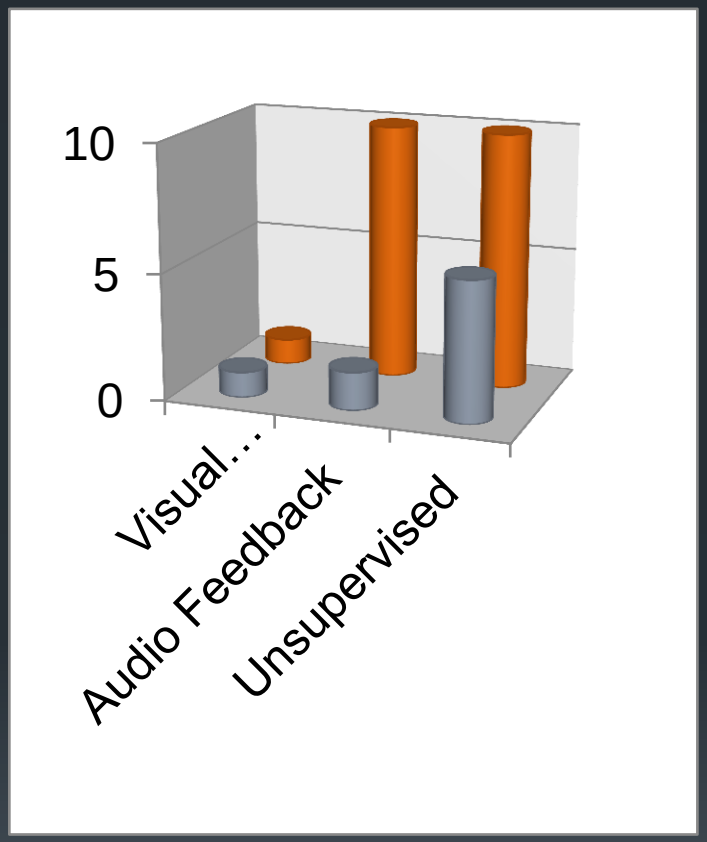


Unsupervised

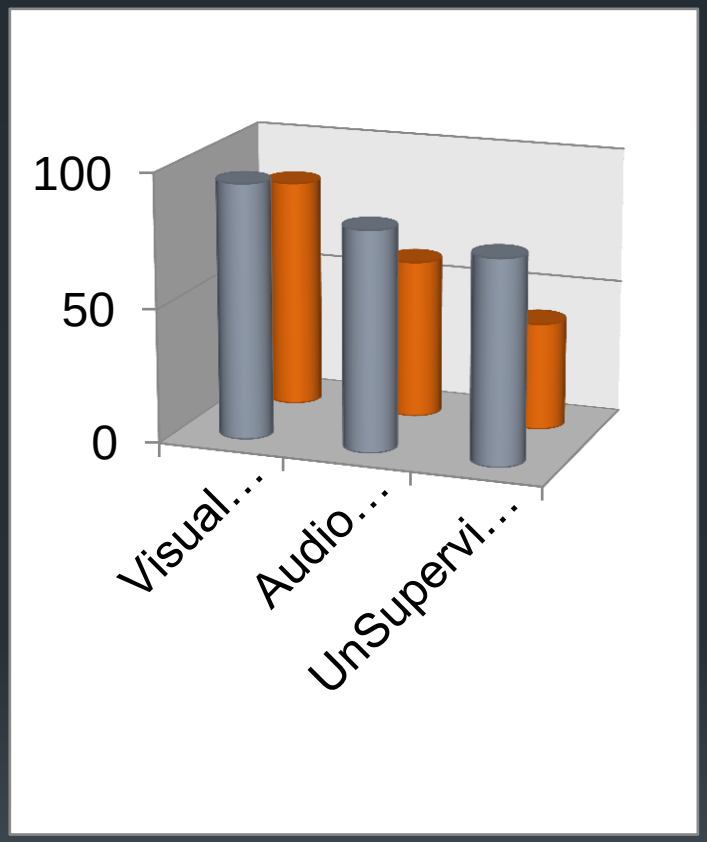


No - Training

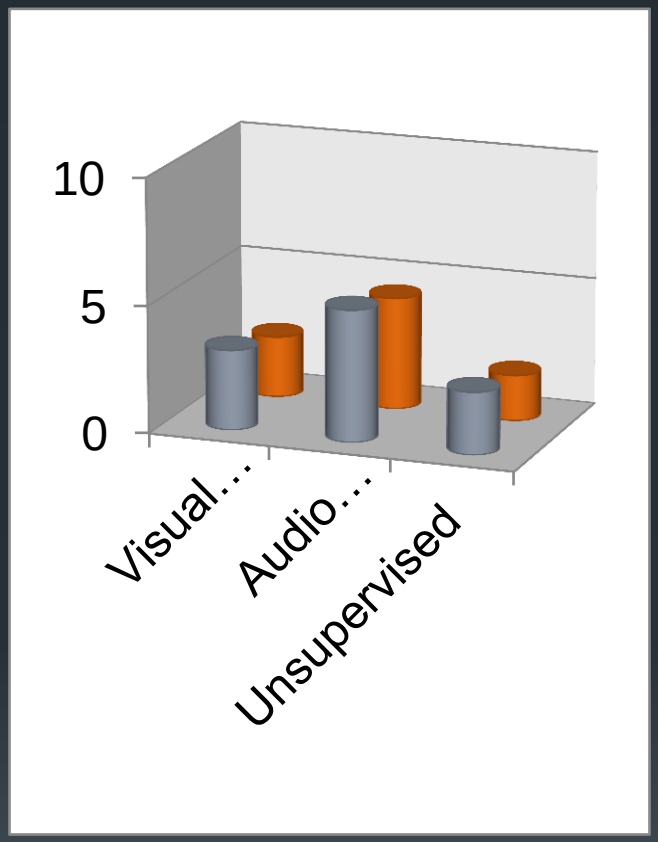
Q 2 : What is the best/fastest way to train the user for this system?



No. Of Trials



Accuracy



Response Time

Multiple Cue Navigation : Test Of Working Memory

Working memory is the system that actively holds multiple pieces of transitory information in the mind, where they can be manipulated. It is also known as ***Short Term Memory***

Multiple Cues were given to the participant at a time.

4-Dir	2	4	6	7	8
8-Dir	2	4	6	7	8

- Ascending Order
- Descending Order
- Random Order

The participants were divided into 2 groups :

- Working Memory with Training
- Working Memory without Training

- Working Memory with Training

4 - Directions

Ascending

Cue Length	Accuracy	Recall	Response Time
2	70	6	2.990625
4	73.75	6	2.24185
6	49.16667	0.5	2.36355
7	67.14286	3.5	2.656929
8	66.875	3.5	2.496969

Descending

Cue Length	Accuracy	Recall	Response Time
8	63.75	2	3.031713
7	74.28571	3	2.966471
6	100	10	2.683133
4	100	10	2.264125
2	90	8	2.77225

8 - Directions

Ascending

Cue Length	Accuracy	Recall	Response Time
2	25	1.5	3.927525
4	58.75	2.5	2.30535
6	41.66667	0.5	2.2565
7	37.14286	0.5	2.337407
8	33.75	0	2.160231

Descending

Cue Length	Accuracy	Recall	Response Time
8	32.5	1	3.116575
7	62.85714	2	3.280857
6	53.33333	0	2.967733
4	47.5	2	3.11655
2	75	7	2.99275

Future Work

- *Dichotic listening Task :* Using our system to study the effect of sound cues in presence of noise.
- *Effect of Incorrect Visual Cue :* on learning and testing response time.
- *Auditory Attention & Sonification :* multi-tasking like navigating while talking.
- *Musical Properties & Usability* : effect of different musical properties like rhythm, scale and timbre on Sonification.
- *Sounds for Color & Shape :* Using sound properties to describe color and shape of objects.
- *Bag of Sounds :* creating descriptors with sounds both at local (object mapping) and global (scene mapping).

Final Goal …To find an answer …

Will the visually challenged be able to Drive a Car, Play Table-Tennis, or Listen to Nature ?

Acknowledgements

- Tejaswinee Kelkar
- Falak Chhaya
- Dr. Priyanka Srivastava
- Dr. Anoop Namboodiri
- Google (for its images).

References

1. Sonification report : Kramer, Walker et al
2. T. V. Tran, T. Letowski, and K. S. Abouchacra, "Evaluation of acoustic beacon characteristics for navigation tasks," Ergonomics, vol. 43, pp. 807-827, 2000.
3. G. H. Mowbray, "Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty," Journal of Experimental Psychology, vol. 46, pp. 365-372, 1953.
4. C. D. Wickens, Engineering psychology and human performance, 2nd ed. New York, NY: HarperCollins Publishers, 1992.
5. J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, and F. Dellaert, "SWAN: System for Wearable Audio Navigation," presented at International Symposium on Wearable Computers, Boston, MA, 2007.
6. B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in International Conference on Auditory Display, London, England, 2006, pp. 63-68.
7. T. Dingler, J. Lindsay, and B. N. Walker. Learnabiltiy of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech. In Proceedings of the 14th International Conference on Auditory Display.