Effect of visual cues on human performance in navigating through a virtual maze

Nikhil Iyengar[1] Industrial Engineering
Deepak Vembar[2] Computer Science
Keith Pauls[3] Computer Science
Jason Hewitt[4] Computer Science
Kevin Clark[5] Computer Science

Abstract

Many applications of virtual environments (VE's) require people to be able to have some kind of spatial knowledge and use visual cues to perform well in the VE. Thus, the design and presentation of cues is important for the efficient use of a VE. In this paper, we report the results of an experiment in which groups were allowed to use three different visual cues (only 2-D map cue, 2-D map with non directional cue, 2-D with an arrow cue) and were asked to apply their spatial knowledge in navigating a 3-D virtual maze. The underlying purpose of the study was to evaluate the effectiveness of the cue types on the performance of subjects in the VE. It was observed that the cue type affected the performance of the participant in the 3D maze. The arrow cue was most effective in terms of time taken to reach the center of the maze as compared to the other cues.

1 Introduction

Virtual Reality (VR) is an artificial environment that is experienced through sensory stimuli as sights and sounds provided by a computer. In such environments, the user's actions partially determine what happens in the environment. With modern advances in technology, VR is being increasingly utilized in a wide range of applications such as avionics, visualization of architectural modeling, and games. Users navigate these spaces through immersion interfaces; however, not much research has been done on the usability and adaptability of these interfaces. An area of interest that was investigated by Vila et al (2003) was that of human navigation and wayfinding in VR. Having a better understanding of navigational patterns employed by humans while interacting with these environments will enable software developers to design more intuitive and effective VR interfaces.

An effective visualization system must accurately communicate shape and other spatial information to the user. To successfully improve on this communication, the visualization community needs to better understand the basic issues underlying spatial perception. While immersive systems can be a useful tool in portraying spatial information, it is important that they do so in a natural and intuitive manner. The perceptual cues driving this are purely visual. Research has been done to explore some of the visual cues likely to be effective in signaling imminent contact in immersive displays (Hu et al, 2000). Further, it must also be noted here that as the demand for quality products has increased and the incidence of product liability litigation has increased, human performance is becoming increasingly important in various industries (Micalizzi and Goldberg, 1989).

Visual cues have been considered as one of the most effective ways of improving human performance. Human reliability for a decision making task is much higher, though less than perfect, than for a search task (Hou et al, 1993). We define search as "a procedure adopted to carefully examine in an attempt to find something". The goal of a search may be to look for nonconformity, an anomaly or a particular route. Various strategies are available to improve the human performance during search. In particular, there have been relatively few efforts to understand the cognitive skill components, especially the perceptual skills, critical to the various aspects of search.

There is enormous variation between individuals in a variety of spatial behaviors, and it is a challenge to identify the sources of this variation (David et al, 1999). Goldin and Thorndyke (1982) examined the different types of knowledge developed from a direct navigation experience and from a simulated experience. The results of the study indicated that people could learn about environments from a simulated medium such as film. Thorndyke suggests that interaction, either by a person driving the tour or interacting with the environment in a simulation, might show different results. The results of this study also show that within a condition, such as the film condition, adding an additional navigational aid has consequences and that the consequences are task dependent. In one test we see that having a map or narration hinders performance, while in another task the map improves performance and narration lowers it. These results are important when deciding what aid to use to introduce someone to a new environment.

In 1985 Streeter, Vitello and Wonsiewicz performed a wayfinding study comparing navigational aids for people driving in a car. This study showed that while wayfinding it is more difficult to interpret a map than it is to receive directions. The finding that the combination of tools did not produce better performance than the narrative alone is what is important, since more tools may not mean better performance.

Regian and Shebilske (1990) conducted studies of the use of VR as a training medium for visual-spatial tasks. One of their experiments involved wayfinding. The environment used in the wayfinding study was a virtual maze. The authors claim that this shows that subjects can learn navigation spatial-navigational skill in a virtual environment. However, their design and methodology does not support this claim. Previous research has already shown that "people do not act like randomly moving automata that make unbiased decisions at each point where a decision has to be made" (Peponis, et. al., 1990). Two rules that Peponis observed in wayfinding behavior, is people avoid unnecessary backtracking and also that people tend to find the area that gives them the best visual access to other areas. Regian, et. al., should have run a

---

email : [1]niyenga@clemson.edu, [2] dvembar@clemson.edu,
[3]kpauls@clemson.edu, [4]hjason@clemson.edu,
[5]wckevin@clemson.edu
Clemson University, Clemson SC 29631

control group for the same wayfinding task. The control group would have had no experience in the building, and therefore would have had to search for the unique object, using strategies that would probably be much more efficient than a random search.

The wayfinding literature indicates that receiving directions verbally or through signs is advantageous when trying to find a goal. In contrast, the use of maps can be disadvantageous. Darken and Sibert, (1993) reported on an informal study looking at toolsets for wayfinding in virtual environments. The tools available to the participants were flying (the ability to rise above the virtual environment), spatial audio markers, visual markers (breadcrumbs), coordinate feedback, grid navigation, and two map-views of the world. Informal observations indicated that people used the different tools in a variety of ways. The conclusion of this report was that subjects showed different behaviors when they used different tools in wayfinding.

The literature on navigation and wayfinding shows us that at present we are unsure of how to introduce a person to new environments so that they will gain navigational awareness efficiently. The Thorndyke and Hayes-Roth (1982) along with the Goldin and Thorndyke (1982) studies show that exposure time and navigational tools can affect the process. Their studies revealed that map study before entering an environment can be beneficial but that when used alone it does not give someone complete navigational awareness. The research also points out that depending on the type of introduction to an environment, an additional navigational aid can have a positive or negative effect. The results also show that learning of an environment can occur without actually being there but the results do not reveal whether it is better to be an active participant or a passive observer when being exposed to a new environment.

The wayfinding literature shows that people are better at finding a target location when using signs or narrative directions than with using maps. Also, by adding navigational tools, it can have a positive effect or a negative effect depending on the task and the original tool or cue used. The research in wayfinding also showed that given a choice people prefer shorter routes to longer routes to a destination despite the complexity of a shorter route.

This brings out the need to examine in further detail how to introduce someone to a new environment, what navigational aids are beneficial during initial exposure and their after-effects when the aid is may no longer available.

This eye tracking study paid particular attention to human performance and his/her ability to find their way using visual cues (feedback). It also presented the importance of spatial knowledge. One important variable that was considered for this experiment to measure the effect of visual cues on transfer of spatial knowledge is speed. For this experiment, speed was measured as the time required to search and identify the route needed to reach the center of the maze. This study also conducted subjective evaluations of visual cues used for the experiment. One important issue surrounding the use of visual cues was the type of visual cue that the participants found most useful in navigation and determining the format for presenting this visual information in a way easy to interpret and utilize. The 3D maze will be developed in a VE where the human users are completely immersed in a 3D computer generated maze.

2 Hypothesis

Ho: There is no effect of visual cues on human performance in the 3-D maze

Ha: There is an effect of visual cues on human performance in the 3-D maze

3 Methodology

a. Subjects

The experiment tested the importance of visual cues in helping people navigate a 3D maze. There were three different groups, each group being tested for a different cue. The cues presented were either no cue, a positional cue or positional as well as directional cue. Each group contained 5 subjects, for a total of 15 subjects. In total 24 participants were tested, with 9 participants being excluded from the statistics due to incorrect eye-tracking data recording. The subjects were drawn largely from a population of graduate and undergraduate students at Clemson University and were randomly assigned to only one of the three different groups.

b. Equipment

The primary rendering computer for rendering the VE consisted of a 1.5 GHz dual-CPU Linux PC with 1 GB RAM and a NVidia GeForce4 Ti 4600 graphics card. Multimodal devices include a V8 Virtual Research Head Mounted Display (HMD) and ISCAN video-based Corneal Reflection eye tracker. The V8 HMD offers a 640x480 pixel resolution for each eye, with separate video feeds for the left and right eye. Eye tracking is provided by the video based corneal reflection ISCAN eye-tracker unit mounted within the HMD. Each of the eye trackers is composed of a miniature camera and an infrared LED. The ISCAN RK-726PCI High Resolution Pupil/Corneal Reflection Processor uses corneal reflections (first Purkinje images) of infra-red LEDs mounted within the helmet to measure eye movements. The processor operates at a rate of 60Hz (30Hz when both eye movements are tracked) and the subject's eye position is determined with an accuracy of approximately 0.3 degrees over a 20 degree horizontal and vertical range using the pupil/corneal reflection difference. The maximum spatial resolution of the calculated POR provided by the tracker is 512 x 512 pixels per eye. The HMD and the eye tacking cameras are shown in Figure 1.

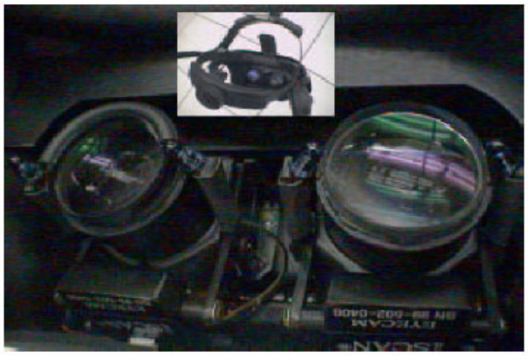


Figure 1: Head mounted Display with binocular eye tracker

HMD position and orientation was measured using Ascension 6 – Degree of Freedom (6-DOF) Flock of Birds electromagnetic tracker, which is mounted on the top of the HMD as shown in Figure 2. A 3-DOF mouse was used for user navigation. Left button presses allowed the subject to move forward along the line of view while the right button presses allowed the subject to move backward along the direction of view. To change the direction of motion, the subject was asked to turn their head in the appropriate direction. The eye tracking apparatus and the ISCAN computer are shown in Figure 3.

Figure 2: Virtual Research V8 Head Mounted Display and 3DOF mouse

Figure 3: Eye tracking Control Computer with left and right eye video screens

c. Procedure

The participants had to perform only one task in the VE using the maze. Through the HMD, they saw a 640 x 480 window split vertically down the middle. See Figure 4. On the right side of the screen was the 3D maze that the subjects were trying to get through. The left side of the screen displayed a 2D top-view of the same maze.

All three groups were given the opportunity to familiarize themselves in the environment and the screen elements. See Figure 4. Note that the same type of maze was shown to all the participants in order to eliminate treatment variability. Consequently, we adopted a between group testing strategy (3 X 3) to eliminate the effects of “learning.” Thus, no participant belonging to one type of cue group was tested for another cue. Also, none of the users were allowed to view the maze used in the actual test phase before it had begun.

The first test group was the control group. This group was asked to navigate through the maze with no help other than the 2D map on the right. They had no directional cues or point of reference once they began. They were timed to see how long it takes them to get through the maze.

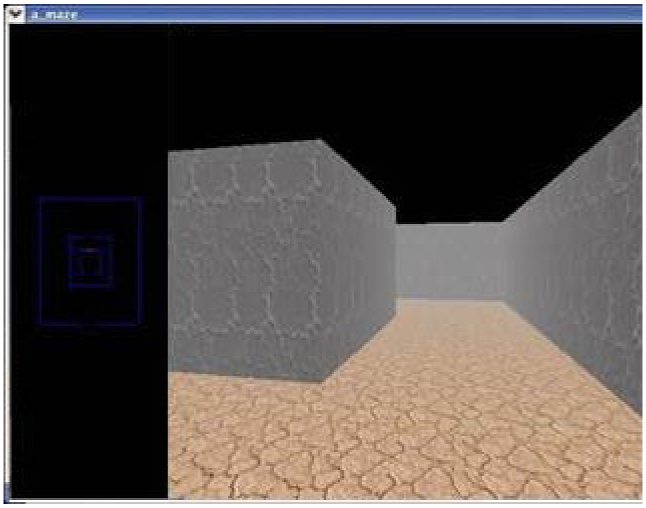
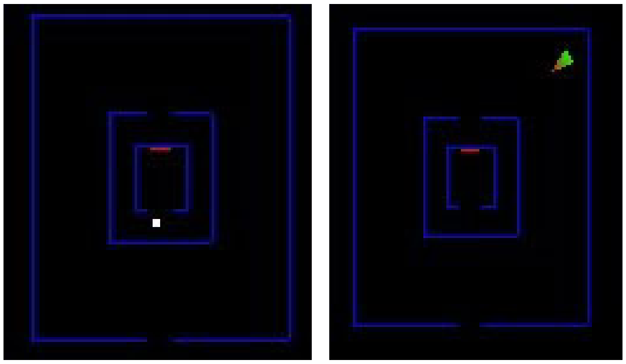


Figure 4: VE that the users familiarized themselves with.



(a) 2-D map with a dot **(b) 2-D map with an arrow**

Figure 5: Different visual cues on 2-D map for the users

The task assigned to the second group was the same as the control group, only they were given a point of reference. A small dot appeared on the 2D map, which represented the location of the participant in the 3D maze. See Figure 5 (a). This allowed them to know where they were inside the 3D maze at all times. The participant was able to use the point of reference while traveling through the maze as they chose.

The third group also performed the same task. However, the visual cue provided to them was represented by an arrow that pointed in the direction the subject was looking. See Figure 5 (b). This gave them a sense of orientation by knowing exactly where they are facing inside the 3D maze in addition to a point of reference.

Participants were timed to see how long it took them to get through the maze. The eye movements were studied to ascertain how often the subject used the 2D map and try to determine if there was a correlation between the different cues and how long it took to navigate through the maze.

d. Stimulus

OpenGL was used as the graphics API for rendering the virtual maze. The test consisted of three main phases:

1) The Familiarity phase: This was used to acclimatize the user to movement and navigation in the virtual maze. A simple version of the maze was presented to the user as shown in Figure 4. The visual cues we presented to the user were the same as in the test phase. This provided a learning opportunity for the user to work in tandem with the cues and the 3D world. The users were not timed during this phase.

2) The Calibration phase: To get proper overlaying of the eye tracker coordinates on the screen, we used this phase to calibrate the eye tracker. Five squares were displayed on various locations on the screen and the user was asked to fixate on them, while the operator performed the calibration routine. This is shown Figure 6 below.

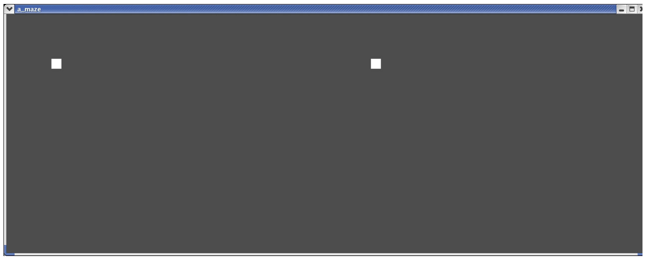


Figure 6: Calibration screen of the eye tracker

3) The Test phase: This was the main test scene for the program. The maze description was read into the program from a simple text file that listed out the (x, y) coordinates for the walls of the maze. The user was instructed to traverse from the outside of the maze to a predefined point, represented by the horizontal red line. The user was timed as soon as the test scene was started and his/her eye movements were recorded for offline analysis.

e. Role of Eye Tracking in the Experiment

The fixations and the saccades of the participants were studied to ascertain how often the subject used the 2D map and to determine if there is a correlation between the different cues and how long it took to navigate through the maze. We tracked only the left eye, with the assumption that both the eyes of the user have similar movements.

The basic assumption was that without any visual cues, the participant will make significantly longer fixations and saccades over the 2D maze and will rely more on his/her cognitive ability to navigate. In the second test group, the assumption was that there would be significantly larger number of switching between the 3D view and the 2D view, partly because of the lack of orientation information. In case of the third test group, it was expected that there would be lesser number of fixations and faster navigation through the maze. Predominantly the fixations in case of the third group would be at the beginning of the test program where the user would rely on his/her cognitive ability to chart out a predetermined course. This would be followed by rather fast "glances" towards the 2D map to make sure he/she was following the predetermined path.

f. Analysis of eye-tracker data

The data we were most interested in was the time it took for the subjects to traverse through the maze. This was recorded in the Test phase, which the subjects were presented after the Familiarization and the Calibration phase. Eye tracker coordinates were recorded every 40 ms for only the left eye. Once recorded, the eye tracker data was stored for later offline analysis.

For measuring the number of crossings that the subjects made when shifting their gaze from the 3D to the 2D map and vice versa, we wrote our own analysis program. See Figures 7a and 7b. With this, we can measure the number of crossings from the 3D to the 2D world, as well as the time spent in the 2D world. Eye tracker coordinates beyond a clipping region were disregarded due to the mapping limitations. The analysis program has the feature of moving around a "virtual line" which represents the boundary between the 2D and the 3D world. Since the screen was 648 x 480 pixels in dimensions, with the 2D map occupying 160 pixels by 480 pixels, a virtual line of 200 pixels from the left was chosen to represent the boundary line.

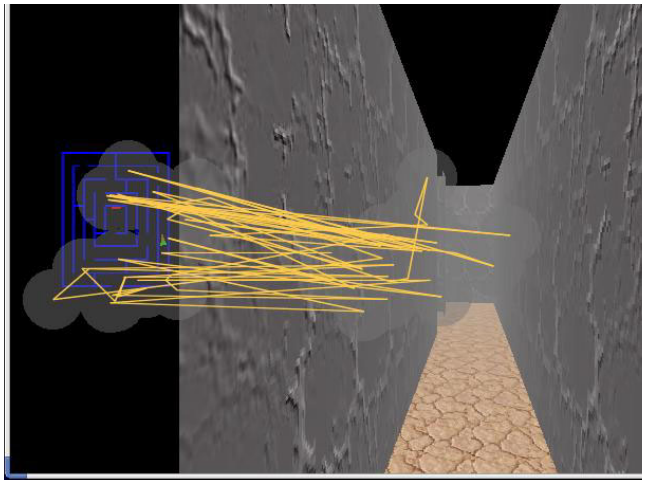
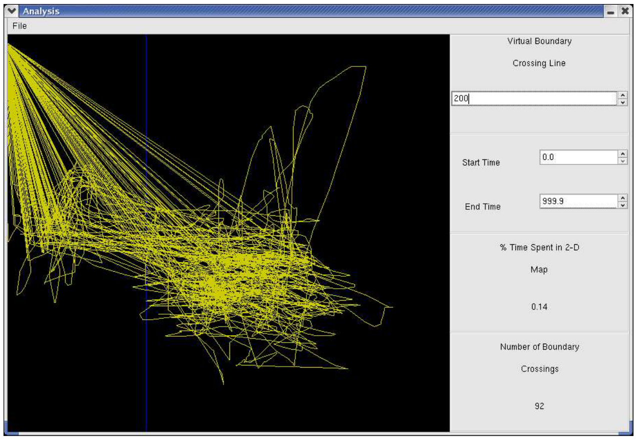


Figure 7a: Analysis Program with scanpaths superimposed on the 3D maze



Virtual Boundary Crossing Line: 200

Start Time: 0.0

End Time: 999.9

% Time Spent in 2-D Map: 0.14

Number of Boundary Crossings: 92

Figure 7b: Analysis Program for calculating the boundary crossing and the time spent on the 2D map.

4 Results

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	214724.82	2	107362.412	5.285	.023
Within Groups	243793.26	12	20316.105		
Total	458518.08	14			

Figure 8: Results of ANOVA for Time taken in the 3D Maze with or without the cues

Based on the above results in Figure 8, there is sufficient evidence, at the 0.05 level of significance, to conclude that there is a difference in the three treatment means. The graph in Figure 9 clearly shows the difference in the time taken to complete the experiment between the control group (No Cues) and the other two groups that were provided cues.

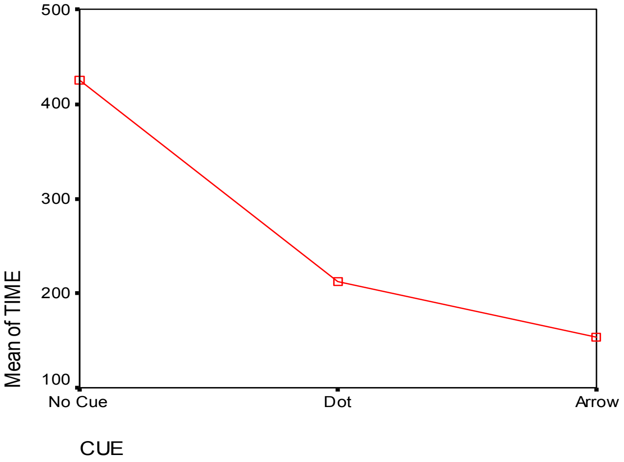


Figure 9: Results of ANOVA for Time taken in the 3D Maze with or without the cues

The descriptive statistics for other information collected, which was worth noting, have been presented in Figure 10 to 16. Figure 10 presents the percentile for the answers on the pretest and post test questionnaires provided by the 15 participants. The pretest questionnaire included only design related questions such as location of 2D map, visibility of the display etc. These questions are similar to the post test design questions. The reason for asking these questions in the pretest was to evaluate the participants' design rating of the maze early on. Thus, comparison of the pretest and the post test design questions allow us to understand if there was any change in their rating. Only a small change in the percentile ratings was noticed. A high percentile of participants strongly agreed on the design (45%) and usefulness (66%) of the VE. Some of the questions asked in post test's usefulness section were 'The 2D map was useful in finding Mona Lisa'; 'The visual cue presented in the 2D map was helpful in navigating through the maze' etc. There were some general complaints with using the eye tracking gear. The participants' comments on this issue are presented in the Discussions section in this paper. However, this is also seen in the percentile data in figure 9.

Question Categories	SDA	DA	N	A	SA
Pretest Design	0	13	17	35	**52**
Post Test Design	0	19	19	44	**45**
Post Test Usefulness	6	0	16	23	**66**
Post Test Eye Tracking	13	27	**33**	27	13

Figure 10: Responses for various Question Categories in percentages

SDA – Strongly Disagree; DA – Disagree; N – Neutral; A – Agree; SA – Strongly Agree

Figure 11 reveals that the subjective rating of majority of the participants (88%) for frequency of usage of the 2D map was 'Very Often.' It is interesting when this subjective rating is compared with the actual percentiles collected from the eye tracking analysis program. The eye tracking data shows a positive correlation between the group- type and the percentage time spent on the 2D map. This would mean that the time spent on the 2D map was higher for those with cues than for those without cues. But, the relation however is not significant (0.061). This observation was contrary to our expectation. We expected the control group (No Cue) to spend more time looking at the 2D map.

Question Categories	Very Often	Often	Sometimes	Rarely	Very Rarely
Post Test Frequency	**88**	19	0	6	0

Figure 11: Subjective responses for Frequency of usage of 2D map during the experiment

Correlations

		CUE	PERC_2D
CUE	Pearson Correlation	1	.494
	Sig. (2-tailed)	.	.061
	N	15	15
PERC_2D	Pearson Correlation	.494	1
	Sig. (2-tailed)	.061	.
	N	15	15

Figure 12: Correlation of percentage time spent on the 2D map and cue type

The table in Figure 13 shows the mean time spent in the 3D maze by gender. As expected by us, the mean time spent in the maze was high for the control group and lowest for the group which received the arrow as the cue. A quick review of the means also shows that overall for all the groups, male participants took lesser time than female participants. In order to understand this relationship better a correlation analysis was conducted.

Figure 14 shows that the gender to cue relationship is significant (0.029).The analysis of the correlation matrix indicates that the observed relationship was strong (r= -0.563) which indicates that if the individual was a male, he is more likely to clock less for the experiment. This is because of the negative sign on the correlation coefficient and the coding of the gender question (0=Female, 1=Male). The negative correlation means that as X increases, Y decreases.

	Mean Time Spent in Seconds		
Gender	No Cue	Dot	Arrow
Male	276.415	154.5445	128.65725
Female	523.97	250.897	255.119
Total	400.1925	202.72075	191.88813

Figure 13: Mean time spent by Male and Female participants

Correlation

		SEX	TIME
SEX	Pearson Correlation	1	-.563(*)
	Sig. (2-tailed)	.	.029
	N	15	15
TIME	Pearson Correlation	-.563(*)	1
	Sig. (2-tailed)	.029	.
	N	15	15

* Correlation is significant at the 0.05 level (2-tailed).

Figure 14: Correlation of Gender and time taken to complete the experiment

The correlation reported in the Figure 15 is negative (-0.109) and it is significantly different (0.698) from 0, which suggests that experience using VR equipment did not have an appreciable effect on performance. The correlation coefficient was also low.

Similar observation was made in Figure 16. The correlation coefficient was low (-0.418) and the relation between boundary crossings and groups was not significant (0.121). The negative coefficient indicated an inverse relationship between boundary crossings and the group types. As the cue changes from none to a dot to finally an arrow, the crossings decreased. This result has been explained in the Discussion section of this paper.

Correlations

		TIME	EXP
TIME	Pearson Correlation	1	-.109
	Sig. (2-tailed)	.	.698
	N	15	15
EXP	Pearson Correlation	-.109	1
	Sig. (2-tailed)	.698	.
	N	15	15

Figure 15: Correlation of experience and time taken to complete the experiment

Correlations

		B_CROSS	CUE
B_CROSS	Pearson Correlation	1	-.418
	Sig. (2-tailed)	.	.121
	N	15	15
CUE	Pearson Correlation	-.418	1
	Sig. (2-tailed)	.121	.
	N	15	15

Figure 16: Correlation of boundary crossing and cue type

5 Discussion

Overall, the participants made good use of the 2D map. However, it was often noted that those users with the visual cues like the dot or the arrow, would use the 2D map to navigate through the 3D maze without looking consciously at the 3D maze. Consequently, it was observed that the participants would walk the 3D maze looking at the floor. This blind navigation could be attributed to the over-reliance of the participants on the 2D map. Through questionnaires the participants were asked to express their comments on the experiments and the cues provided. The general response to the color and design elements was positive. They complained on the inconvenience caused by the wires. A few users mentioned that often this distraction would result in them loosing their orientation and forgetting where they are in the maze thereby loosing precious time. This would be particularly bad for a participant who was working with no cues. The heat from the HMD was a common complaint made by all the users. There was disagreement between users on the use of the mouse in the study. Some felt it was very useful although they did not like the fact that the left mouse button would lock in a pressed position. Other group of participants felts that the mouse buttons did not necessarily match up with the direction of movement. A forward movement with a joystick which would be similar to the virtual movement in the VE was asked by the participants.

The interesting result from this experiment was the time spent by the participants on the 2D map. The eye tracking data shows us that the group which had the arrow cue (perceived as the easiest by our team), exploited the 2D map to the fullest by spending most of their time looking at the 2D map. This strategy allowed them to get to the center of the maze. The other groups spent relatively less time on the 2D map as they had to compare what they saw in the 3D maze with what they understood from the 2D map. Needless to say that the strategy adopted by the arrow group allowed them to get to Mona Lisa much quickly without making much use of the 3D maze as compared to the other groups.

The low number of crossings for the arrow group indicates that the users were less confused about their orientation or location and hence jumped less often between the 3D maze and the 2D map. This was as expected by us. The reason is fairly simple. The participants who received the arrow as a cue had a much better idea of where they were and which direction they were looking and hence had less doubt in their mind during the wayfinding process.

6 Future Work

There is a need for more research to determine other factors that affect navigational behavior in VR. One of the limitations of this paper is the size and the distribution of the data. A larger group size would provide a stronger foundation for the results and inferences made. Further, the distribution of data across the three different groups is not uniform. More research could be done in the area of gender. This will help us understand the effect of gender on wayfinding. Use of cues in environments other than maze would allow for further understanding of how humans perform in real world environments. It would also be interesting to explore how they employ their real world strategies in these conditions and if they are successful in doing so. This research is needed to gain a good understanding of wayfinding strategies employed by different types of users with the intent to formulate VR design guidelines that could aid VR designers in creating more effective VR applications for specific groups.

7 References

David W., Earl H., and David K., The Transfer of Spatial Knowledge in Virtual Environment Training. 1999

Darken, R.P., and Sibert, J.L., (1993). A Toolset for Navigation in Virtual Environments. *Proceedings of the ACM Symposium on User Interface Software and Technology*. November, 1993, Atlanta, GA 157-165.

Duchowski, A.T., Shivashankariah, V., Rawls, T., Gramopadhye, A. K., Melloy, B. J., and Kanki, B., Binocular eye tracking in virtual reality for inspection training. Eye Tracking Research and Applications Symposium, 2000

Goldin, S.E., and Thorndyke, P.W. (1982). Simulating Navigation for Spatial Knowledge Acquisition. *Human Factors, 24(4),* 457-471.

Hou, Tung-Hsu, Lin, Li, and Drury, C.G. and Prabhu, P.V. (1997), Training strategies for visual Inspection, Human Factors and Ergonomics in Manufacturing, pp. 351-367.

Hu, H., Gooch, A. A., Thompson, W. B., Smits, B. E., Rieser, J. J., and Shirley, P., Visual Cues for Imminent Object Contact in Realistic Virtual Environments Proceedings of the conference on Visualization 2000. pp. 179-185.

Micalizzi, J. and Goldberg, G., 1989. Knowledge of results in visual inspection decisions: Sensitivity or criterion effect. International Journal of Industrial Ergonomics, 4: 225-235.

Peponis, J., Zimring, C., and Choi, Y.K., (1990). Finding the Building in Wayfinding. *Environment and Behavior, 22 (5)* 555-590.

Regian, J.W., and Shebiske, W.L., (1990). Virtual Reality: A Instructional Medium for Visual-Spatial Tasks. *Journal of Communications 42(4)* 136-149.

Streeter, L.A., Vitello, D., and Wonsiewicz, S.A., (1985). How To Tell People Where to Go: Comparing Navigational Aids. *International Journal of Man-Machine Studied, 22,* 549-562.

Vila, J., Beccue, B., and Anandikar, S., (2002). The Gender Factor in Virtual Reality Navigation and Wayfinding. Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)