Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 3 (2015) 5059 – 5066

Procedia MANUFACTURING

6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

Effect of Environment Immersivity on Encoding Strategies of Spatial Tasks

Maria Kozhevnikov[a,]*, Lee Rong Cheng[a], Michael Kozhevnikov[b]

[a]*Department of Psychology, National University of Singapore, Singapore, 117570*
[b]*Department of Engineering, Norfolk State University, Norfolk, VA 23504, USA*

Abstract

The main objective of the current project was to examine the effect of different viewing environments on individual's use of spatial frames of references while performing mental rotation task. Twenty five participants were administered a computerized mental rotation task presented in three different viewing environments: two-dimensional computer desktop (2D), three-dimensional large projector screen (3DP), and three-dimensional large projector screen with tracking (3DPT).The participants' responses were compared to those obtained in fully-immersive virtual environment (VE) offered through the use of head-mounted displays. In 2D condition, the participants encoded the stimuli allocentrically, in relation to a computer screen. This is in contrast to a fully immersive VE, where participants encoded the images egocentrically, in relation to a viewer-centered frame of reference. In 3DP and 3DPT conditions, the participants tended to rely more on egocentric rather than allocentric encoding, although not to the same extent as in fully-immersive VEs. The change in participants' responses in 3DP and 3DPTconditions, as compared to 2D condition, implies that large screen displays provide higher degree of immersivity than conventional 2D displays. However, in contrast to fully immersive VEs, large screen displays are insufficient to ensure the dominant use of a viewer-centered frame of reference.

Peer-review under responsibility of AHFE Conference

*Keywords:*Immersivity, Egocentric and Allocentric Spatial Processing, Viewer-centered and Scene-based Encoding

* Corresponding author. Tel.: +65-6601-1190; fax: +65-6773-1843.
E-mail address: psymaria@nus.edu.sg

Peer-review under responsibility of AHFE Conference
doi:10.1016/j.promfg.2015.07.519

1. Introduction

Spatial ability is broadly defined as the cognitive functions that allow people to process mental rotation of object, orientation in space, spatial relations, and spatial visualisations[1-3]. Research has revealed that environment immersivity has an impact on individual's spatial ability performance [4-6]. In particular, mental rotation, the ability to manipulate mental representation of perceived physical object in space [7,8], was demonstrated to be affected by different environment immersivity[6]. Some studies have argued that a fully-immersive VE is not required to attain the desired immersivity in the test conditions; large screen display can be a good substitute for immersive VE (i.e., in terms of the degree of immersivity provided) for improving spatial task performance [9,10]. However, other researchers have suggested that large screen displays may not bring about better results for all spatial tasks [11]. Thus, in view of the mixed results on whether large screen displays can render comparable immersivity as immersive VE, one of the aims of this study was to investigate how the use of large screen displays may affect an individual's mental rotation task performance.

1.1. Mental Rotation

Shepard and Metzler's mental rotation (MR) task [7], where participants compare if two three-dimensional (3D) shapes presented on a 2D display (see Fig. 1) were identical or different by mentally rotating the said stimuli, is one of the main tests used in the examination of mental rotation ability. The angular difference (i.e., 20 to 180 degrees) applied to the shape also differed in terms of the orientation axes the shape was rotated around [7]. There are three rotational axes which the shapes could possibly rotate around in the Shepard and Metzler's task, namely rotation in the X-, Y-, and Z-axes (i.e., line-of-sight, vertical, and depth rotation respectively) or otherwise known as principal rotations.

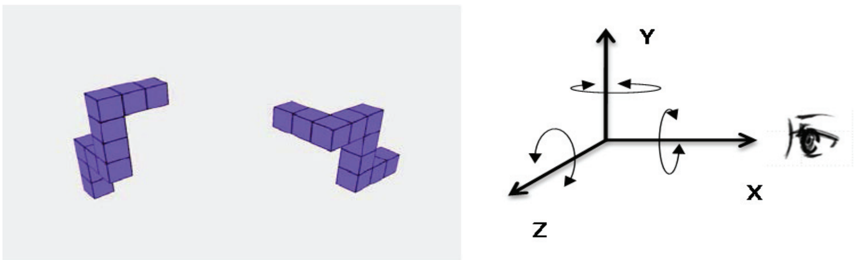

Fig. 1. Example of the mental rotation task showing the observer's view of the stimulus (left) and the three principle axes of rotation (right).

Results of Shepard and Metzler's study [7] showed a positive linear relationship between angular disparity (i.e., up to 180 degrees) and reaction time, a robust finding that was replicated in many subsequent studies (e.g., [7,12,13). As the angular divergence increases, it seemed that individuals needed more time to mentally reposition the shape, as though one was physically rotating them, to ascertain congruency with the paired shape. However, there was no significant difference found in terms of reaction times recorded for stimuli shown to be rotated around the X- or Z- axes, which seemed to contradict the idea that subjects in fact rotate mentally the shapes in 3D space. If the shapes were egocentrically encoded, it should imply that the shapes which had undergone Z-axis rotational transformation were harder to mentally rotate than those of X-axis rotational transformation, as the former would cause the individual to experience and perform occlusion and foreshortening of the three-dimensional shapes compared to the latter [7]. Thus, the higher cognitive load experienced through increased difficulty in visualizing and manipulating the shapes that had undergone Z-axis rotational transformation task should result in the individual using more time (i.e., longer reaction time) than those that had undergone X-axis rotational transformation [14]. Thus, similar rates of rotation around the X- and Z-axes suggested that the shapes may not be egocentrically encoded, in relation to the viewer, as previously thought, but rather encoded allocentrically, in relation to the computer screen.

1.2. Environment immersivity

Building on the results of Shepard and Metzler's landmark study [7], Kozhevnikov&Dhond[6] investigated the role of immersivity (i.e., immersive and non-immersive environments) on Shepard and Meztler's MR task. Different degrees of immersivity were operationalised through different experimental conditions, namely 2D desktop display, 3D non-immersive desktop display, and 3D immersive VE using a head-mounted display (HMD) [6]. From the results, Y-axis rotation was revealed to be significantly faster than X- and Z-axes rotations in all three conditions, which was consistent with Parsons's study[15]. Moreover, the identical reaction time for X- and Z-axes rotation rates in the 2D display and 3D non-immersive VE was similar to previous studies [7,15]. However, Z-axis rotations were significantly slower than X-axis rotations in 3D immersive VE condition. This suggested that perhaps egocentric encoding of shapes (i.e., such that Z-axis rotation is slower than X-axis rotations) was only elicited in immersive VE but not in 2D or 3D non-immersive environments [6]. Immersive VE arguably situates the participant within the 'problem' space (i.e., the same perceptual space as the stimuli) and thus, participants are biased to encode egocentrically. The paper further posited that in the 2D display and non-immersive VE objects were allocentirally encoded in that the display edges of the computer screen act as frames of reference for spatial relations encoding [6] and as such, there was no difference in the rotation speeds around the X- and Z- axes.

Previously, it was thought that immersivity is categorised dichotomously: either an environment is immersive or not immersive [16]. However, Bowman [16]also argued that immersivity should be considered along a continuum as there are various sub components that constitutes immersivity, one of which being the display size. Studies have shown that large screen displays seem to confer similar degree of immersivity as immersive VEs[9,10]. Large screen displays are typically rendered as wall-sized projections and the greater field-of-regard (FOR) as compared to conventional desktop monitor screen seems to add to the immersivity of the participants [11,17]. Patrick et al. asked participants to recreate the maps of the VE seen at comparable visual angles (i.e., same retinal image size) through the computer monitor screen, large projection display, or the HMD [9]. Results showed that performance of subjects in large projection screen and HMD conditions were not significantly different, suggesting that large screen may be as effective as HMD in conferring immersivity. Johnson and Steward II also found that measures of spatial knowledge (i.e., heliport map viewing) of military personnel assigned to either HMD or large screen display conditions were not significantly different in their scores[18]. However, as the aforementioned examples concentrated only on tasks concerning map formation and not on other types of spatial ability, it is unclear if the results obtained could be generalized over all spatial image encodingand transformation. Taking the nature of spatial tasks into consideration, Tan at al.'s study [11] tested a variety of spatial tasks and they argued that large screen displays, being more immersive compared to small screen, would bias participants into using more egocentric approach of solving spatial tasks, which in turn would improve spatial task performance. They also showed that performance of spatial task that are 'exocentric' in nature (e.g., MR task) would not be improved (i.e., higher proportion correct) by increased environment immersivity as "egocentric strategies bias from more immersive displays do not help exocentric tasks performance" [11]. However, no comparison of reaction time of different rotational axes between small and large screen displays was conducted, which would have elucidated the possible effects of immersivity on rotational rates of different axes.

This paper aimed to examine the effects of varying environment immersivity on the performance of MR task. The test conditions were conventional 2D desktop display (2D), 3D stereoscopic large projection screen (3DP), and 3D stereoscopic large projection with tracking (3DPT) in ascending order of increasing environment immersivity. 2D was used as a baseline condition to show how participants perform in the conventional MR task and subsequently be used as a reference for comparison for any possible departure from the baseline response patterns in both 3DP and 3DPT. Both 3DP and 3DPT have equivalent size of projector display screen, which was larger than that of 2D. Head-tracking system was added in 3DPT condition to simulate spatial updating by adjusting the display screen scene perceived in accordance to the head movement of the individuals.

The expectation was that with different levels of immersivity in the three test conditions, individual's performance across conditions would be different. The focus was on the reaction time of the rotations around X- and Z-axes so as to determine the type of encoding strategy used in solving the MR task and the expectation is that speed of Z-axis rotations will have a larger latency than X-axis rotation rates. The difference in reaction time (RT) around

X- and Z-axes will be contrasted with that of the 2D condition to show a departure from the usual pattern of RT in 2D setups (i.e., identical RT of rotation around X- and Z-axes), which would imply that a different encoding strategy was used for solving the MR task in different test conditions.

Theother hypothesis was to investigate if the gender difference in efficiency with regards to MR task is preserved in the three test conditions. Mental rotation has been one of the most established gender differences in the study of cognitive psychology [19-21]. Men generally perform better on average than women and the difference is attributed to a myriad of reasons, like inherent difficulty of task [22], different patterns of brain activation with men showing more parietal activation and women displaying more right frontal activation in electroencephalography (EEG) studies [23], or different strategies used for MR task, with men using a more spatial one compared to women using a more verbal one [5]. However, Seurinck et al [24] found only modest differences between males' and females' performance when they were asked to use egocentric encoding in performing egocentric mental rotation of hand shapes. This departure from the normal gender difference seen in MR task begs the question if different environment immersivity, which as aforementioned with the more real-world equivalent sensory modalities would draw out the use of egocentric encoding for both men and women. Parsons et al. [5] found that gender differences were only observed on traditional pen-and-paper mental rotation tasks, but not in immersive VE. This suggested that perhaps the switch to a more immersive condition would force both genders to use egocentric encoding, and the use of the same strategy may eliminate the gender difference previously seen [5].

2. Method

2.1. Participants

Twenty-five undergraduates (12 males, M_{age} = 22.92, age range: 20-25, and 13 females, M_{age} = 20.87, age range: 19-23) from the National University of Singapore participated in the study. All participants had normal or corrected-to-normal vision. Course credit was awarded for their participation.

2.2. Materials and Procedure

A computerised adaptation of the Shepard and Metzler's (1971) test was used in this study. For each trial, participants viewed two spatial figures, one of which was rotated relative to the position of the other (Fig. 1). Subjects were to imagine rotating one figure to determine whether or not it matched the other figure, and to indicate whether they thought the figures were the same or different by pressing a left (same) or right (different) button on a remote control device. Twelve rotation angles were used: 20, 30, 40, 60, 80, 90, 100, 120, 140, 150, 160, and 180 degrees. The figures were rotated around 3 principal spatial axes (X-, Y-, and Z-axes), each axis of rotation would have a total of 24 trials, which adds up to be 72 (24 x 3) trials in total. The Vizard Virtual Reality Toolkit v. 3.0 [25] was used to create the scenes and to record the dependent variables (response time and accuracy).

In the 3DP and 3DPT conditions, the stereoscopic stimuli display was beamed onto a wall by a projector and was measured at 237cm wide by 177cm tall (Fig. 2). NuVision 60GX wireless stereoscopic glasses were worn during both the 3D projector conditions over any viewing aid (e.g., spectacles) which the participants might have been wearing to enable perception of 3D shapes rendered by the projector. Specifically for the 3DPT condition, a head tracking bar apparatus with two infrared lights was attached to the top of the stereoscopic glasses worn (two infrared LED lights were required to compute the actual orientation of the participant's head as the stereoscopic glasses lack an internal digital compass). The participant's head position was tracked by 4 cameras located in each corner of the experimental room that are sensitive to the infrared light.

To enable a fair comparison of display screen size across conditions, the stimuli perceived on the participants' retinal image has to be of uniform size in all three conditions. Difference in retinal image size can possibly be a confounding variable as bigger (or smaller) retinal image size would enable participants to perform better (i.e., higher accuracy and lower RT) or worse (i.e., lower accuracy and higher RT) across test conditions) [11,14]. In order to maintain a constant visual angle across all three conditions, participants were seated at specific spot with the computer screen and projection screen displays set at varying calculated distances from the participant so that the retinal image size was kept constant in all three test conditions, and any differences in reaction time or response accuracy among the test condition are unlikely to be due to the discrepancy in the retinal image perceived. The distances from the participant to the computer monitor screen and projection screen were approximately 60cm and 365cm respectively.

Participants completed a series of computerised MR task across three different conditions (i.e., 2D, 3DP, and 3DPT). The conditions were counterbalanced. Each session comprised of two example trials (i.e., one match and one non-match) and eight practice trials to familiarise participants with the experimental conditions and mental rotation strategy. Participants were guided through the practice session to ensure that they understood how to solve the MR tasks by mentally rotating the stimuli shown and not through use of other strategies (e.g., verbal). Participants were informed to keep an upright position and minimise head movements while performing the mental rotation task. Fig. 2 below shows the experiment setup of the three test conditions aforementioned.

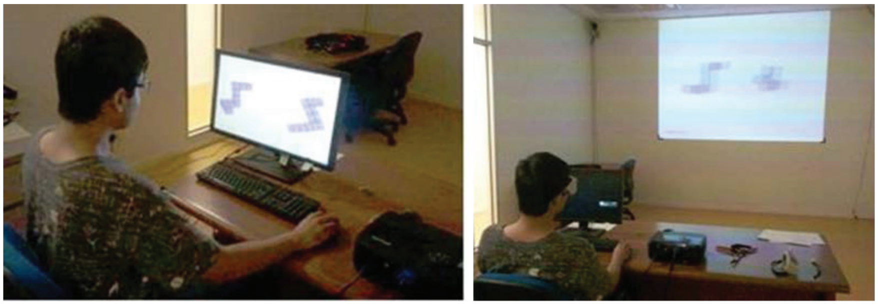

Fig. 2. The 2D condition where a desktop setup was used (left) and the 3DP and 3DPT conditions where a stereoscopic scene was projected onto a wall (right).

3. Results

3.1. Accuracy

An examination of participants' accuracy (Proportion correct) as a function of test conditions (2D, 3DP, and 3DPT) and rotational axes (X-, Y-, and Z-axis) was conducted by using a 3X3 repeated measures ANOVA. No significant result was found for interaction effect of test conditions and rotational axes on participants' accuracy, with $F(4, 96) = 1.56$, $p = .19$. The accuracy on the whole was very high across conditions and axes of rotations, with the proportion correct ranging from 0.77 - .092. As such, the remaining analyses conducted shall be focusing primarily on RT.

3.2. Reaction Time

An examination of participants' RT as a function of test conditions (2D, 3DP, and 3DPT) and rotational axes (X-, Y-, and Z-axis) was conducted by using a 3X3 repeated measures ANOVA. There was a significant interaction effect of condition and axes on RT, with $F(4, 96) = 15.43$, $p = .001$. There was also a significant main effect of axes on RT, with $F(2,48) = 35.11$, $p = .001$, where Z-axis rotations were significantly slower than Y-axis rotations in all three conditions (p's <.05).

Analysis of simple main effects showed that in 2D, RT of X- and Z-axes rotation were not significantly different ($p = .424$) and Y-axis rotation RT was significantly smaller than that of X- ($p = .001$) and Z-axis rotations ($p = .001$). However, a different pattern of RT was observed in 3DP and 3DPT. In 3DP and 3DPT conditions, no significant difference was revealed between the RTs of X- and Y-axes rotation ($p = .642$ and $p = .079$ respectively), while RT of Z-axis rotations is significantly greater than that of X-axis rotation, (p's = .001), and also significantly larger than that of Y-axis rotations ($p = .017$ and $p = .001$ respectively). This reaction pattern was similar to Kozhevnikov&Dhond's 3D immersive VE condition [6] in which the X- and Y-axes rotation RT were not significantly different and the RT of Z-axis rotation was significantly larger than that of X- and Y axes rotations.

The reaction pattern of 2D condition effectively replicates the finding of previous studies, whereby RT of X- and Z-axes rotation are not significantly different and RT of Y-axis rotation is significantly slower than X-and Z-axes [6,7]. Fig. 3 shows the pattern of accuracy (proportion correct) and RT (seconds) of the three axes of rotation across the three test conditions.

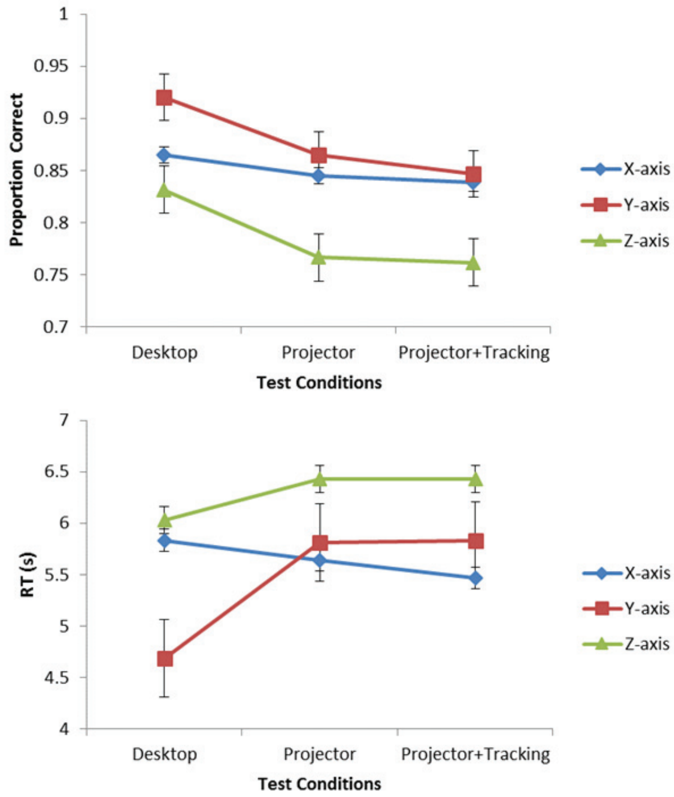

Fig. 3. Accuracy (top) and RT (bottom) as a function of environment immersivity. Note that the vertical axis does not start from 0.

Since the interaction of rotational axes and test conditions has a significant effect on RT, further analyses were conducted to see if the test conditions (i.e., environment immersivity) had a significant effect on the rate of interaction of each axis of rotation and compare the interaction pattern to that of Kozhevnikov &Dhondstudy [6]. Mean regression slopes were computed based on the average participant's regressions in a 3X3 repeated measures ANOVA. The mean regression slopes with standard error (SE) for each condition and axis or rotation are presented in Table 1.

Table 1**.**Mean regression slopes (sec/deg) and SE rotations around the X-, Y-, and Z-axes in 2D, 3DP, and 3DPT conditions. Results on immersive VE from Ref. [6] were included in the table to act as a reference for comparison

	Axis of Rotation		
Condition	X	Y	Z
2D Desktop	0.030 (0.002)	0.011 (0.002)	0.029 (0.004)
3D Projector	0.022 (0.003)	0.015 (0.003)	0.024 (0.004)
3D Projector with Tracking	0.024 (0.004)	0.009 (0.002)	0.027 (0.004)
3D Immersive VE [6]	0.028 (0.003)	0.014 (0.003)	0.043 (0.005)

Repeated measures ANOVA was also conducted to determine the effect of angle of rotation (i.e., 12 angles ranging from 20 to 180 degrees) around the X-, Y-, and Z-axes in the 2D, 3DP, and 3DPTtest conditions on participants' RT in completing the MR task. There was a significant interaction effects of axes and conditions, with F(22, 682) = 3.90, *p* = .001. Both main effects of test condition and angle of rotation were found to be significant, with F(2, 62) = 24.02, *p* = 0.001 and F(11, 341) = 38.30, *p* = 0.001 respectively. The results suggest that different environment immersivity does have a significant effect on rates of MR around different axes of rotation. In addition, it seems that the angular disparity between the stimuli pair presented and the RT needed forms a positive linear relationship as reported in previous studies [6,7].

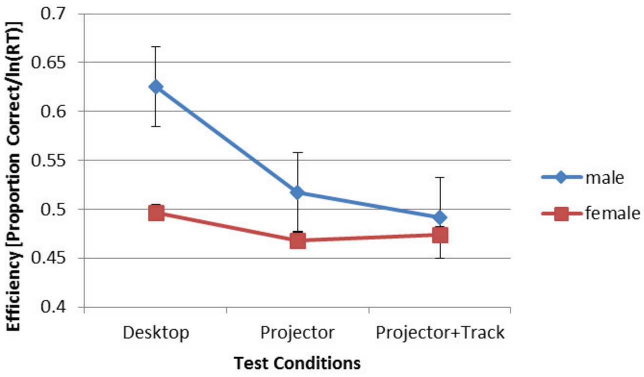

Fig. 4.Difference in efficiency between males and female in 2D, 3DP, and 3DPT conditions. Note that the vertical axis does not start from 0.

3.3. Gender Difference

One-way between subjects ANOVA was also calculated to see if gender difference had an effect on mental rotation efficiency rate (i.e. mean proportion correct/RT) in 2D, 3DP, and 3DPT conditions. The accuracy of participants was computed by summing up the correct responses across the three axes of rotation in each condition separately (i.e., proportion correct of trials around X-, Y-, and Z- axes in 2D, 3DP, and 3DPT to obtain total proportion correct in each condition). Total average RT was derived by summation of the average RT correct in each The RT was normalised using a logarithm (i.e., ln) function to transform the RT recorded. Efficiency was then calculated through dividing the total proportion correct in each condition by the normalised RT. The results (see Fig. 4) showed that a significant gender difference was found in the 2D condition, with F(1, 23) = 3.20, *p*< .05. However, efficiency of men and women in 3DP and 3DPT were both found to be non-significantly different, with F(1,23) = 0.98, *p* = .33 for 3DP and F(1, 23) = 0.16, *p* = .70. These results suggest that gender difference in efficiency rate is only shown in 2D condition, but not in 3DP and 3DPT. The results suggest that the more immersive the environment is, the more non-significant statistically the gender difference seems to be (i.e. *p* value increases from .33 in 3DP to .70 in 3DPT).

4. Discussion

From the resultsit can be seen that different environment immersivity does have a significant effect on participants' MR task performance. In 2D condition, the results showed that the RTs of rotation around the X- and Z-axes were essentially the same, indicating that, consistent with previous studies [6,7], the participants engaged in allocentric encoding (in relation to a computer screen) to solve the MR task. In 3DP and 3DPT conditions, the RT of MR around Z-axis was slower than around X-axis. This change in RT pattern implied that the participants were persuaded to rely more on egocentric encoding when situated in a more immersive VE, as compared to the use of more allocentric encoding in the conventional 2D condition. However, in contrast to fully immersive VEs, large screen displays are insufficient to ensure the dominant use of a viewer-centered frame of reference.

The results of this study do not fully agree neither with several previous studies [9,10,17] that found that large screen display is indeed a good substitute for fully-immersive VE in terms of the immersivity level afforded, nor with the results of Tan at al.'s paper [11] in that 'exocentric' task like MR task is not affected by large display screen size. In fact, the results of this study suggest that while large screen display is better than the standard 2D computer display in terms of the immersivity level afforded, only filly immersive environments seem to ensure for individuals to encode images in relation to egocentric (viewer-based) spatial frame of reference.This paper used a stereoscopic large screen display instead of a conventional 2D large screen display as used in Tan et al.'s study [11], so that some disparity in terms of stereoscopic depth cues provided by the display may have been the reason for the increased immersivity shown in 3DP and 3DPT conditions in the current study but not in Tan et al.'s [11]. At the same time, when presentation is on a large projection screen, but the edges of the screen are visible (forming a "frame"), target objects are embedded within a "fixed frame", suggesting that whenever a person is observing a scene fromthe "outside" (e.g., a scene is defined by the frame or computer screen), it relies more on the use of allocentric (scene-based) encoding.Further studies need to be carried out so as to delineate and explicate the actual effect of large screen display.

The second hypothesis sought to investigate if the robust gender difference in MR task performance would be preserved in the large screen display conditions. The results suggest that the gender difference is non-significant in more 3DP and 3DPT compared to 2D condition. As such, as the environment condition becomes more immersive, gender difference in MR task performance is seemingly absent.

References

[1]M. C. Linn and A. C. Petersen, Child Development56 (1985) 1479-1498.
[2]D. Voyer, S. Voyer and M. P. Bryden, Psychological Bulletin117 (1995) 250-270.
[3]M. Sjölinder, SICS Technical Report T98:01, ISRN: SICS-T-98/01-SE (1998) 46-58.
[4]J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt and J. N. Bailenson, Psychological Inquiry13 (2002) 103-124.
[5]T. D. Parsons, P. Larson, K. Kratz, M. Thiebaux, B. Bluestein, J. G. Buckwalter and A. A. Rizzo, Neuropsychologia42 (2004) 555-562.
[6]M. Kozhevnikov and R.P. Dhond, Frontiers in Psychology, doi: 10.3389/fpsyg.2012.00284 (2012).
[7]R. N. Shepard and J. Metzler, Science171 (1971) 701-703.
[8]J. E. Roberts, Developmental Neuropsychology17 (2000) 199-223.
[9]E. Patrick, D. Cosgrove, A. Slavkovie, J. A. Rode, T. Verratti and G. Chiselko, CHI Letters2(1) (2000) 478-485.
[10]D. A. Bowman and R. P. McMahan, Computer 40 (2007) 36-43.
[11]D. S. Tan, D. Gergle, P. Scupelli and R. Paush, Transactions on Computer-Human Interaction13 (2006) 71–99.
[12]M. C. Corballis and R. McLaren, Journal of Experimental Psychology8 (1982) 215-224.
[13]J. M. Zacks, J. Mires, B. Tversky and E. Hazeltine, Spatial Cognition and Computation2 (2000) 315-332.
[14]R. N. Shepard and L. A. Cooper, Mental images and their transformations, MA: MIT Press, 1982.
[15]L. M. Parsons, Attention, Perception, & Psychophysics42 (1987) 49-59.
[16]D. A. Bowman, "Tutorial Virtual Reality Resources," 2008. [Online]. Available: http://www.cs.panam.edu/~rfowler/VR-papers/Bowman_2008_lecture_Immersion-and-presence.pdf. [Accessed 15 February 2012].
[17]J. Z. Bakdash, J. S. Augustyn and D. R. Proffitt, "Large displays enhance spatial knowledge of a virtual environment," in APGV '06 Proceedings of the 3rd symposium on Applied perception in graphics and visualization, 2006.
[18]D. M. Johnson and J. E. Steward II, Military Psychology11 (1999) 129-148.
[19]D. Kimura, Sex and Cognition, MA, USA: MIT Press, 1999.
[20]K. Kucian, T. Loenneker, T. Dietrich, E. Martin and M. Von Aster, Psychology Science 47 (2005) 112 - 131.
[21]M. Brosnan, I. Walker and J. Collomosse, European Journal of Cognitive Psychology22 (2010) 172-189.
[22]D. W. Collins and D. Kimura, Brief Communications 111(4) (1997) 845-849.
[23]E. Weiss, C. Siedentopf, A. Hofer, E. Deisenhammer, M. Hoptman, C. Kremser, S. Golaszewski, S. Felber, W. Fleischhacker and M. Delazer, Neuroscience Letters344 (2003) 169–172.
[24]R. Seurinck, G. Vingerhoets, F. de Lange and E. Achten, NeuroImage23 (2004) 1440–1449.
[25]WorldViz. (2007). Vizard VR Toolkit R3: http://www.worldviz.com/products/vizard4/index.html.