Applied Human Factors and Ergonomics International

The Role of Three-dimensional Immersive Environments in Assessment and Training of Spatial Skills

Maria Kozhevnikov [a] *, Ido Amihai* [a]*, and 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

Recently, increasingly realistic three-dimensional (3D) visual displays have been designed to serve as new, more ecologically valid alternatives to conventional two-dimensional (2D) displays. However, research has thus far provided inconsistent evidence regarding the effectiveness of 3D displays in facilitating training and task performance. Here we investigated how individuals generate and transform mental images within 3D immersive environments, in which the viewers perceive themselves as being surrounded by a 3D world. In Experiment 1, we compared participants' performance on the mental rotation task across the following types of visual presentation environments: traditional 2D non-immersive (2DNI), 3D non-immersive (3DNI – anaglyphic glasses), and two 3D immersive (3DI) environments: 3DI/HMD - head mounted display with position and head orientation tracking, and 3DI/A –augmented virtual reality technology implementing "see-though" head mounted display. In Experiment 2, we compared electroencephalogram (EEG) data recorded while participants were mentally rotating spatial images presented in 3DI/HMD vs. 2DNI environments. The findings suggest that in a non-immersive environment, participants may utilize more "artificial" encoding, during which the 3D images are encoded with respect to a scene-based frame of reference (i.e. the computer screen). On the other hand, in an immersive environment, participants use egocentric encoding strategy, during which the 3D images are encoded in relation to the observer, the same strategy they would use in a real-world. Overall, the results of both experiments indicate that immersivity aspect of an environment might be one of the most important aspects to be considered for assessment and training in domains that rely on visual-spatial performance (e.g., robotics, navigation, medical surgery).

Keywords: Immersivity, Virtual and Augmented Reality, Egocentric and Allocentric Spatial Processing, Viewer-centered and Scene-based Encoding, Electroencephalogram (EEG).

INTRODUCTION

Our ability to generate and transform three-dimensional (3D) visual-spatial images is important not only for our every-day activities (locomotion, navigation) but also for a variety of professional activities, such as architecture, air-traffic control, and telerobotics. Difficulties of studying visual-spatial cognition within real world environments, where controlling the experimental stimuli and recording participants' behavior is often impossible, have led researchers to increasingly employ 3D immersive (3DI) virtual environments (Chance et al., 1998; Klatzky et al., 1998; Loomis et al., 1999; Macuga et al., 2007; Kozhevnikov and Garcia, 2011).

In a 3D immersive environment the observer perceives himself/herself as being surrounded by a 3D world. In contrast, in a non-immersive environment (i.e. traditional computer display, either 2D monocular or 3D-stereoscopic), the observer is placed outside of the scene looking in. Currently, most human visual processing research has been conducted using two-dimensional non-immersive computer displays. However, brain processing

https://openaccess.cms-conferences.org/#/publications/book/978-1-4951-2101-2

in 2DNI environments might vary from those during human interactions in the real-world. Indeed, recent evidence suggests that visual processing in non-immersive environments significantly differs from that in immersive environments, and that 3DI environments are more similar to natural real-world environments with respect to visual image encoding and transformation processes (Kozhevnikov & Dhond, 2012).

However, regarding assessment and training applications, although these 3D environments are both more appealing to the user and richer in spatial information, research thus far has not reached a strong conclusion regarding the effectiveness of 3D environments for promoting visual-spatial learning and task performance (e.g., Van Orden & Broyles 2000). For some tasks, such as collision avoidance, 3D displays have been found to facilitate learning and performance better than 2D displays (e.g., Van Orden & Broyles, 2000), but other researchers have reported that 3D displays as less efficient than 2D displays for a number of visual tasks involved in airtraffic control and aviation (e.g., Alexander & Wickens, 2005; Hollands et al., 1998). Currently, too little is known about the cognitive processes that underlie learning and training in 3D vs. 2D environments to fully justify using 3D immersive virtual reality displays. The focus of the current research is to understand how individuals process visual-spatial information in 3D immersive environments vs. 3D non-immersive (stereo-glasses) and conventional 2D non-immersive visual displays, and how complex 3D immersive technology can facilitate assessment and training of spatial skills.

Spatial Frames of Reference in Immersive vs. Non-immersive Environments

Two different spatial frames of reference, *environmental* and *viewer-centered*, can be used for encoding and transforming visual-spatial images. An environmental frame may involve the “permanent environment”, which is bound by standard orthogonal planes, i.e., the floor, walls, ceiling, and perceived direction of gravity or the local “scene-based” spatial environment, where the target object’s components are encoded allocentrically in relation to another object, i.e., table-top, blackboard, computer screen, etc. In contrast to environmental frames of reference, the viewer-centered frame is egocentric, that is, it defines object configurations and orientations relative to the viewer’s gaze and it includes an embedded retinal coordinate system. In the case of imagined spatial transformations such as mental rotation (MR), the prevailing hypothesis is that individuals rely more upon an environmental, scene-based, rather than a viewer-centered frame of reference (Corballis et al., 1976, 1978; Rock, 1986; Hinton and Parsons, 1988; Palmer, 1989; Pani and Dupree, 1994).

Evidence for primacy of scene-based reference frames for spatial image encoding and transformations comes from experiments (e.g., Parsons, 1987, 1995) comparing the speed of MR of classical Shepard and Metzler’s (1971) 3D forms around different axes (see Fig. 1).

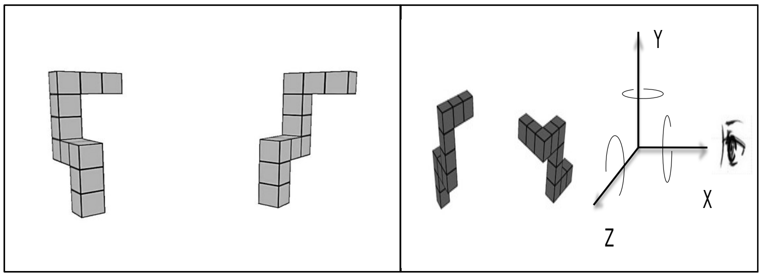

Figure 1.Example of the mental rotation task showing the observer's view of the stimulus (left) and the three principle axes of rotation (right)

MR around different axes places different demands on the transformation processes, and results in different brain activity (Gauthier et al., 2002). Rotation in the picture plane preserves the feasibility of all the features of a shape, but perturbs the top-bottom relations between features. Rotation in depth around the vertical axis alters side-to-side relationships between features and the visibility of features, some coming into view and others becoming occluded. Rotation in depth around a horizontal axis is the most demanding rotation; it alters top-bottom relations between features and feature visibility.

It has been consistently reported that participants mentally rotate shapes in the depth plane just as fast as or even faster than in the picture plane (Shepard and Metzler, 1971; Parsons, 1987, 1995). If participants were in fact rotating viewer-centered 2D retina-based visual representations, the depth rotation would take longer than rotation in the picture plane, since rotation in depth would have to carry out additional foreshortening and hidden line removal operations, not required during picture plane rotation. Interestingly, in a recent study on mental rotation in 3DI, Kozhevnikov & Dhond (2012) showed that depth rotation takes longer than rotation in the picture-plane only in a 3DI environment, suggesting that participants rotate mental images in 3D space only in this environment. Kozhevnikov & Dhond (2012) suggested that a limited and fixed field of view may encourage the use of structural scene-based encoding, during which the parts of the 3D image are encoded in relation to the sides of the computer screen or another salient object in the environment. However, because 3DI environments enclose an individual within the scene and allow images to be updated with respect to the observer's head orientation, egocentric, viewer-centered encoding may predominate. Thus, the controversial results from the previous studies on MR could be explained by the fact that they have been conducted using traditional non-immersive environments, where the stimuli were presented on a 2D computer screen or another flat surface (e.g., a table-top), which defines a fixed local frame of reference.

The primary goal of the current research was to examine how individuals process visual-spatial information (specifically encode and rotate 3D images) and what spatial frames of reference they rely upon in 3DI virtual environments vs. conventional non-immersive displays. In our first experiment, we replicated the data from Kozhevnikov & Dhond's (2012) study regarding the unique pattern of mental rotation exhibited by participants in a 3DI (HMD-based) environment, by comparing participants' performance on the Shepard and Metzler (1971) MR task across the following environments: traditional 2DNI, 3DNI (anaglyphic glasses), 3DI/HMD [head-mounted display (HMD) with position and head orientation tracking]. In addition we examine whether the patterns of participants responses exhibited in 3DI/HMD environments remains valid for a 3D immersive augmented reality environment (3DI/A). In the second experiment, electroencephalogram (EEG) responses were recorded while participants performed the MR task in 3DI and 2DNI environments to examine further how neurocognitive correlates of visual-spatial imagery are affected by immersivity of visual presentation environment.

EXPERIMENT 1

Participants and materials

In the first part of the study, fourteen volunteers (7 females average age = 21.5) participated in the study for monetary compensation. Participants were asked about their ability to perceive stereoscopic images prior to the start of the experiment and those who reported difficulty with stereopsis were excluded from participation. Each of the above participants completed the Mental Rotation (MR) task--a computerized adaptation of Shepard & Metzler's (1971) task--in three different viewing environments: 3DI/HMD, 3D non-immersive (3DNI), and 2D non-immersive (2DNI). In the second part of the study, fifteen additional volunteers (8 females and 9 males) were recruited who performed the same MR task in a 3D immersive augmented reality (3DI/A).

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 spatial axes: line of sight (X), vertical (Y), and horizontal (Z) corresponding to rotations parallel with the frontal (YZ), horizontal (XZ), and midsagittal (XY) anatomical planes, respectively (Fig 1). The Vizard Virtual Reality Toolkit v. 3.0 was used to create the scenes and to record the dependent variables (response time and accuracy).

In the 3DI virtual environment, scenes are presented to the participant through an nVisor SX60 (by Nvis Inc) Head Mounted Display (HMD). The HMD has a 44" horizontal by 34" vertical field of view with a display resolution of 1280 x 1024 and under 15% geometric distortion. During the experiment, participants sat on a chair in the center of the room, wearing the HMD (Fig. 2, left panel) to view "virtual" Shepard & Metzler images in front them. Sensors on the HMD enabled real-time simulation in which any movement of the subject's head immediately caused a corresponding change to the image rendered in the HMD. The participant's head position was tracked by 4 cameras located in each corner of the experimental room that are sensitive to an infrared light mounted on the top of the HMD. The rotation of the participant's head was captured by a digital compass mounted on the back of the HMD.

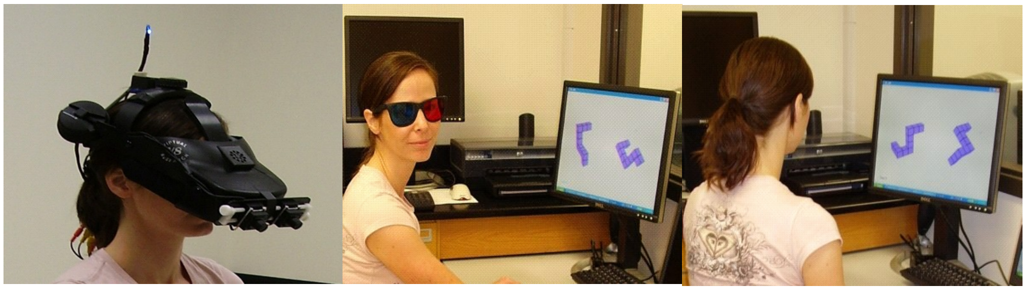

Figure 2. Three different viewing environments, 3DI (left), 3DNI (middle) and 2DNI (right)

In the 3DNI environment, images were presented to the participant on a computer screen. Stereoscopic depth was provided by means of stereo-glasses. In the 2DNI environment, scenes were presented on a standard computer screen. In the 3DI/A environment, virtual images of Shepard & Metzler forms were superimposed onto a real lab environment (Fig.3). The retinal image size of the Shepard & Metzler stimuli received by the participants' eyes was kept constant across all the environments (computed as the ratio between the image size and the participant's distance from the screen).

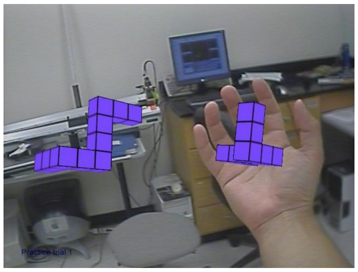

Figure 3. An example of a mental rotation trial in a 3DI/A environment

Research design

Each participant performed the MR task in all environments (the order of the environments – 3DI/HMD, 2DNI and3DNI was counterbalanced for the first group of the participants). Before beginning the MR trials, participants listened to verbal instructions while viewing example trials in each environment. In 3DI (both 3DI/HMD and 3DI/A), to familiarize the participants with immersive technology, there was also an exploratory phase prior to the practice trials in which the participants were given general instructions about virtual reality and the use of the remote control device. During the practice and test phases, the participants remained seated in the chair, but were allowed to move and rotate their head to view 3D Shepard & Metzler shape. The participants were also given similar time to familiarize themselves with the Shepard & Metzler shapes in the 3DNI and 2DNI environment.

Results

The accuracy level was relatively high for all the environments and all axes, ranging from 88% to 97% correct. Given the high accuracy rate, indicating that ceiling performance was reached for some rotations, we focused our remaining analyses on the response time during correct trials (RT).

First, to analyze RT as a as a function of axis of rotation and viewing environment, we performed 3X3 repeated-measures ANOVA for the first group of students (who were administered MR task in 3DI/HMD, 3DNI, and 2DNI), using axis (rotation around X, Y, or Z), and environment (3DI/HMD, 2DNI, or 3DNI) as independent factors. There was a significant effect of axis [$F(2, 22) = 16.20$, $p < .001$], where Y axis rotations were significantly faster than X and Z (p's $< .05$). There was no significant effect of environment ($F<1$), however, there was a significant interaction between axis and viewing environment ($p < .001$). Examination of simple main effects revealed that, consistently with previous studies (Kozhevnikov & Dhond, 2012; Parsons, 1987, 1995; Shepard & Metzler, 1971) in 3DNI and 2DNI, RT for rotation around the Y axis was significantly faster than either around X (all *ps < 0.05)* or Z (all *ps < 0.05).* However, rotations around X and Z axes were similar ($p = 0.95$ and 0.73 for 2DNI and 3DNI respectively). Interestingly, the pattern was opposite in 3DI/HMD: rotation around the Z axis was significantly longer than around the X ($p < .001$) or Y axis ($p < .01$). Thus, our central finding is that in 3DI/HMD, the RT of rotation differed between X and Z axes (Z was slower) and that rotation around the Y axis was faster than Z but not faster than X rotations. In contrast, response time patterns for 2DNI and 3DNI environments were similar to those found in previous MR studies that found Y rotations to be faster than X *and* Z rotations).

Furthermore, analysis of RT as a function of axis of rotation in 3DI/A environment for the second group of students (who took the mental rotation task in 3DI/A) showed that for the 3DI/A environment, RT for rotation around the Y axis was significantly faster than either around X (all $ps < 0.05$) or Z (all $ps < 0.01$). , However, rotation around Z was significantly longer than around X ($p < .001$) or Y ($p < .01$). Thus, in 3D/A, participants exhibited the same pattern of responses as in 3DI/HMD environment.

The slopes of the best-fit linear RT-Rotation Angle functions for each axis and each environment (representing *rates of rotation* around different axes in different environments) were computed and are presented in Table 1.

Table 1: Mean regression slopes of RT-Rotation angle function (sec/deg)

Environment	Rotation around X	Rotation around Y	Rotation around Z
3DI/A	0.028	0.013	0.052
3DI/HMD	0.029	0.014	0.043
3DNI	0.029	0.013	0.031
2DNI	0.032	0.016	0.036

As it can be seen from the table, for both 3DI/A and 3DI/HMD environments, the rate of rotation around Z was more than 1.5 times slower than around X. In both 3DNI and 2DNI, the rate of rotation around X and Z did not differ. There were no differences in the rate of rotation around the *Y* axis across environment, and the rate of rotation around *Y* seems to be the one of the fastest rotations independently of viewing environment, consistently with the findings of previous investigators (Parsons, 1987; Corballis, 1988) who argued that rotation around *Y*, a "gravitational vertical" axis, is the most common of all rotations in our ecology resulting from our extraexperimental familiarity with such rotation.

In summary, the results of Experiment 1 show that the rate of mental rotation about the horizontal axis (Z axis) in 3DI environments (and only 3DI) was significantly slower than the rate of rotation about the line of sight (X axis). This finding suggests that in the 3DI environments the participants were encoding and rotating 2D retina-based visual representations in relation to a *viewer-centered* frame of reference since only then would depth rotation take longer than rotation in the picture-plane, due to the involvement of additional foreshortening and hidden line removal transformations. In contrast, in 2DNI and 3DNI environments, the rates of mental rotation around the X and Z axes were not different, consistent with previous findings that investigated mental rotation using 2D traditional computer displays (Shepard & Metzler, 1971; Parsons, 1987). Thus, in non-immersive environments, participants seem to generate visual representations containing more allocentric information such as information about spatial relations among the elements of the object and their orientations with respect to the scene (i.e., the computer screen) in which the object is presented. The fact that there was equivalent performance in 2D and 3D non-immersive environments suggests that depth information per se, which is provided in a 3DNI environment is insufficient to encourage the use of viewer-centered frame of reference.

EXPERIMENT 2

In this experiment, electroencephalography (EEG) was recorded while participants performed the MR-task in 2DNI and 3DI/HMD environments. Since we didn't find significant differences between 3DI/A and 3DI/HMD in Experiment 1, we used only the 3DI/HMD environment in Experiment 2. Based on the results of Experiments 1, we hypothesized that brain responses should differ in these environments, indicating different mechanisms of spatial attention allocation during viewer-centered versus scene-based encoding in 3DI versus 2DNI environments. For instance, mental rotation in a 3DI environment might require participants to utilize a viewer-centered egocentric reference frame, preferentially engaging dorsal visual processing streams (e.g. occipito-parietal brain areas). In contrast, the allocentric reference frame which dominates within the 2DNI environment may preferentially engage ventral occipito-temporal brain areas (Gramann et al., 2006). Furthermore, the act of mental rotation engages not only spatial processing, but motor processing as well (Cohen et al., 1996; Wexler et al., 1998). Previous work has suggested that motor areas are most engaged when subjects imagine rotation as a consequence of manual manipulation (Kosslyn et al., 2001). Therefore an immersive, egocentric environment *by nature* may increasingly modulate sensorimotor (centroparietal) activity in comparison to non-immersive environments.

Furthermore, previous EEG research suggests that the MR task involves four sequential cognitive stages which may also differentially modulate frontal and posterior brain regions (Desrocher et al., 1995). The first stage occurs at ~200–300 ms post-stimulus, and is independent of the object's angular disparity and involves early sensory processing and simple stimulus evaluation. Subsequently, at ~300–400 ms a pre-rotation "set-up" stage involves evaluation of object orientation and rotation strategy selection. Third, is the act of MR that occurs at ~ 400–800 ms post-stimulus which is followed by response selection and execution from ~1000 ms onward. Object encoding with respect to a specific frame of reference occurs prior to the actual process of MR. Thus, the selection of a frame of reference should begin in the earliest cognitive stages between 200 and 400 ms post-stimulus. Therefore, we hypothesized that when performing a MR task in 2DNI vs. 3DI, brain response differences should be largest at early sensory and/or pre-rotation "set-up" stages occurring at ~200–400 ms post-stimulus.

Participants and materials

Twelve undergraduate psychology students (six males and six females, average age =22) from the National University of Singapore participated in the study. Electroencephalogram was recorded while subjects completed the MR task in 2DNI and 3DI viewing environments. The order of environments was counterbalanced and, in general, the procedures were similar to the first experiment except as follows: EEG was recorded using a 256-channel HydroCel Geodesic Sensor Net (Electrical Geodesics, Inc.). Signals were amplified using the EGI NetAmps 300 amplifier. The signal was sampled at 250 Hz and bandpass filtered online at 1.0−100 Hz. For the 3DI condition, the HMD was placed directly on top of the sensor net (Fig. 4).

Figure 4. EEG recording in a 3DI environment.

Each participant completed the Mental Rotation (MR) task used in Experiment 1 in two different viewing environments (similar to those used in Experiment 1): 2DNI and 3DI. The only difference was that the participants were administered 4 sets of 72 randomly ordered trials (overall 288 trials) for the MR test in each environment (in contrast to 1 set of 72 trials administered in Experiment 1). The order of environments was counterbalanced. The virtual reality set-up (for 3DI) as well as for 2DNI was similar to that used in Experiment 1.

Preprocessing and analysis were done in MATLAB (R2011b, The Mathworks Inc., Natick, MA), using a combination of EEGLAB (Delorme & Makeig, 2004), Neuromag software (Elekta, Stockholm) and MNE Software (http://www.martinos.org/mne/).

Results

All participants demonstrated clear centroparietal responses while performing the MR task in the 2DNI and also the

3DI environment (Fig. 5). As can be seen in Fig. 5, evoked potential (μV) peaks were largest in centroparietal electrodes regardless of visual environment. In the 2DNI environment (solid lines), parietal ERPs were highly similar for shapes rotated around the X and Z axes but more negative for Y axis rotations from ~250 ms onwards (Fig.5A). In the 3DI environment (dashed-lines) X, Z and Y rotation demonstrated ERP peaks at ~350 ms (Fig. 5B). When comparing rotations between 2DNI and 3DI environments, 3DI rotations demonstrated a trend for greater negativity prior to 350 ms but slightly larger positivity at longer latencies >350 ms (Fig.5C, D, E). The largest differences were between Z-rotations which were more negative for 3DI at ~270–300 ms (paired t-test: $p < 0.03$).

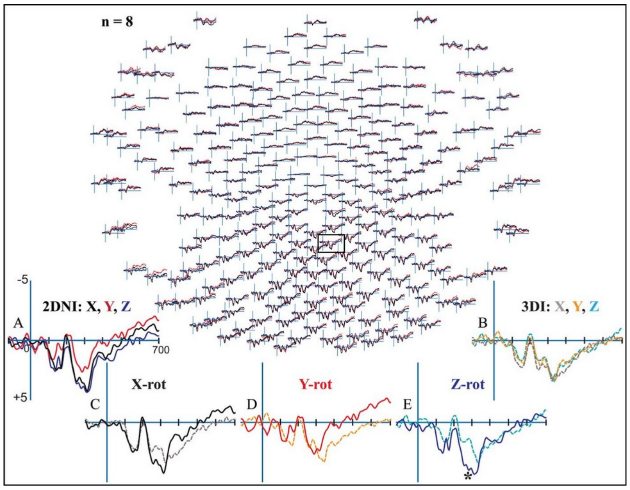

Figure 5. EEG for mental rotation task in 2DNI and 3DI viewing environments. 2DNI results are plotted as solid lines, and 3DI as dashed-lines

It is well known that mental rotation modulates parietal ERPs (Heil, 2002). Specifically, as the angle of rotation from midline increases, parietal ERPs become more negative at ~300 ms onwards, thereby *indexing the process of mental rotation* (Wijers et al., 1989; Peronnet and Farah, 1989). If mental rotation in a 3DI environment indeed forces participants to use an egocentric reference frame, then it should differentially involve dorsal visual processing streams (e.g. parietal brain areas). The results demonstrate that 2DNI and 3DI environments do evoke differential parietal ERP response. Specifically, ERPs for Z-rotations were more negative at ~270-300 ms post-stimulus for MR in the 3DI vs. 2DNI environment. One interpretation is that differential response represents egocentric orienting in preparation for subsequent mental rotation from 350ms onwards. This could be related to allocation of spatial attention for encoding an image from a viewer-centered vs scene-based frame of reference.

Furthermore, it has been shown that motor cortex mu and beta rhythms decrease in power when a participant observes, imagines, or performs movements (McFarland et al., 2000; Pfurtscheller et al., 2006). Therefore, we also evaluated how spectral power in mu and beta frequency bands differed for task performance in these environments (Fig. 6A). PSD maps of mental rotation with X, Y, and Z data combined, demonstrated a classical parietal-occipital mu (top row) and parietocentral beta (bottom row) power distribution in both 2DNI (first column) and 3DI (second column) environments. Stronger mu power was noted in the 2DNI environment over central electrodes (Fig. 6B) while increased beta band response was present for MR in 3DI over right occipitoparietal areas (Fig. 6C).

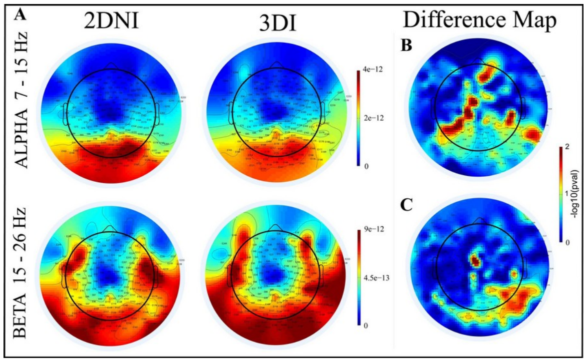

Figure 6. Spectral power in the mu (alpha) and beta frequency bands for MR processing in3Di and 2DNI

environments.

The analysis showed that the participants demonstrated decreased mu response in parietocentral electrodes for 3DI versus 2DI (Fig. 6B). One possibility is that the immersive, egocentric, environment more extensively engages motor imagery during MR. It should be noted that participants were discouraged from moving their heads while in the 3DI environment. Thus, the decreased mu observed in our experiment is unlikely due to small movements, but may rather indicate increased use of motor imagery within the egocentric environment. Finally, extended differences were found not in motor/sensory electrodes but instead, beta power was strongest for 3DI processing within right parietoccipital electrodes (Fig. 6C).

CONCLUSIONS

Overall, the findings of all three experiments reported in this paper support the idea that immersivity that implies viewing the scene from inside and being a part of it, is necessary to provide adequate information for building the spatial reference frame crucial for egocentric encoding and transformations. The results of Experiment 1 show that depth information per se, which is provided in a 3DNI environment, is insufficient to encourage the use of a viewer-centered frame of reference. Our findings suggest that in non-immersive 2D and 3D displays, where the viewer is observing the scene from the "outside", encoding and spatial transformations rely on a scene-based frame of reference that would not typically be used in a large-scale real environment.

In contrast, our results for 3D immersive environments (3DI/HMD and 3DI/A) suggested the use of an egocentric frame of reference while in these environments. In 3D immersive environments, the rate of rotation in horizontal depth (around Z) was significantly slower than the rate of rotation in the picture plane (around X). Furthermore, the rate of rotation in the picture plane (around X) was faster in immersive environments compared to non-immersive environments, which is expected for rotation in a plane where no object components are occluded. At the same time, the rate of rotation in horizontal depth (around Z) was slower in the 3DI environment compared to non-immersive environments, suggesting that subjects were in fact rotating 2D retina-based object depth representations, and were experiencing difficulties with foreshortening and occlusion.

The results of Experiment 2 demonstrated that 2DNI and 3DI environments do evoke differential parietal ERP responses, and that ERPs for Z-rotations were more negative at ~270-300 ms post-stimulus for MR in the 3DI vs. 2DNI environment. One interpretation is that this early decreased positivity reflects a viewer-centered vs. scene-based orienting within the 3DI during image encoding stages in preparation for subsequent mental rotation from 300ms onwards. In addition, participants demonstrated decreased mu response in pericentral electrodes for 3DI versus 2DNI. It is possible that the immersive environment more extensively engages motor imagery during MR. However, more work should be done in the future to separate movement versus motor imagery related activity in the 3DI environments in order to support this conclusion. Overall much work remains to be done to further investigate electrophysiological differences and their underlying significance for cognitive processing in these different environments.

Furthermore, our study is the first attempt to understand the correlates of immersivity from a cognitive neuroscience perspective. Currently, there is no clear understanding of what "immersivity" means in cognitive terms. Most of the definitions are rather descriptive, such as "perceiving oneself to be enveloped by, included in, and interacting with an environment" (Witmer & Singer, 1998, p. 227), and often confounded with such terms as "immersion" and "presence" describing the "extent to which the human operator loses his or her awareness of being present at the site and instead feels present in the artificial environment" (Durlach & Mavor, 1995, p. 22). We believe that an *immersive* 3D environment, similar to the one used in the current experiments, will provide all the necessary information to encourage the use of egocentric (viewer-centered) frames of reference as in a real environment. These findings have implications for future studies on spatial transformations of mental images and the design of testing environments. They show that the results of the previous experiments on mental rotation, performed in laboratory conditions using a traditional 2D computer screen, might be limited and do not reflect the mental rotation patterns that occur in a natural, three-dimensional environment. In addition to its theoretical implications, this research could be of considerable interest from an applied perspective; specifically for the design of learning environments.

Although 3D environments might be more attractive to the user, the results of the current research show that there will probably be no significant differences between encoding and spatial transformation of images in 2DNI and

3DNI. On the other hand, a 3DI environment can provide a unique and more realistic learning environment, and is beneficial for those tasks that benefit from encoding from an egocentric frame of reference (e.g., navigation, wayfinding, laproscopic surgery, and telerobotics). Using desktop graphics to train users for real world egocentric spatial tasks might not be effective, and may actually be counterproductive due to the differences in encoding and transformation processes in immersive versus non-immersive environments. In fact, the findings of this research explain the results of previous studies that show no transfer from training in 2D environments to immersive VR. For instance, Pausch et al. (1997) reported that immersive prior practice with conventional 2D displays in visual search tasks impaired performance in immersive VR. The researchers suggested that using desktop graphics to train users for real world search tasks may not be efficient. The current study explains this finding by pointing out that the encoding of spatial relations and cognitive strategies applied to perform visual/spatial transformations in these two types of environments are different. We suggest that 3DI environments with a variety of simulated 3D stimuli will provide the most efficient environment for training egocentric visual-spatial skills that will generalize and transfer to real-world tasks.

REFERENCES

Alexander, A.L. and Wickens, C.D. (2005). "*3D Navigation and Integrated Hazard Display in Advanced Avionics: Performance, Situation Awareness, and Workload*". Institute of Aviation, Aviation Human Factors Division, Savoy, IL, USA.

Chance, S., Gaunet, F., Beall, A., and Loomis, J. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence* 7, 168-178.

Cohen, M.S., Kosslyn, S.M., Breiter, H.C., DiGirolamo, G.J., Thompson, W.L., Anderson, A.K., Brookheimer, S.Y., Rosen, B.R., Belliveau, J.W. (1996). Changes in cortical activity during mental rotation. A mapping study using functional MRI. *Brain* 119, 89-100.

Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychological Review*. 95, 115–123.

Corballis, M. C., Anuza, T., and Blake, L. (1978). Tachistoscopic perception under head tilt. *Perception & Psychophysics* 24, 274-284.

Corballis, M. C., Zbrodoff, J., and Roldan, C. E. (1976). What's up in mental rotation? *Perception & Psychophysics* 19, 525-530.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods* 134, 9-21.

Desrocher, M. E., Smith, M. L., and Taylor, M. J. (1995). Stimulus and sex differences in performance of mental rotation: evidence from event-related potentials. *Brain and Cognition*, 28, 14–38.

Durlach, N.I., and Mavor, A.S. (1995). "*Virtual Reality: Scientific and Technological Challenges*". Washington, DC: National Academy Press.

Gauthier, I., Hayward, W.G., Tarr, M.J., Anderson, A.W., Skudlarski, P., and Gore, J.C. (2002). BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron* 34, 161-171.

Gramann, K., Müller, H.J., Schönebeck, B. and Debus, G. (2006). The neural basis of egocentric and allocentric reference frames in spatial navigation: Evidence from spatio-temporal coupled current density reconstruction. *Brain Research*, 1118, 116-129.

Heil, M. (2002). The functional significance of ERP effects during mental rotation. *Psychophysiology* 39,535-545.

Hinton, G. E., and Parsons, L. M. (1988). Scene-based and viewer-centered representations for comparing shapes. *Cognition* 30, 1-35.

Hollands, J. G., Pierce, B. J., and Magee, L. E. (1998). Displaying information in two and three dimensions. *International Journal of Cognitive Ergonomics, 2*(4), 307-320.

Klatzky, R., Loomis, J., Beall, A., Chance, S., and Golledge, R. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science* 9, 293-298.

Kosslyn, S. M., Thompson, W. L., Wraga, M., and Alpert, N.M. (2001) Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. *Neuroreport* 12, 2519-2525.

Kozhevnikov, M. & Dhond, R. P. (2012). Understanding immersivity: image generation and transformation processes in 3D immersive environments. *Frontiers in Psychology*, doi: 10.3389/fpsyg.2012.00284

Kozhevnikov, M., and Garcia, A. (2011). "*Visual-Spatial Learning and Training in Collaborative Design in Virtual Environments,*" in Collaborative design in virtual environments, eds. X. Wang and J. Jen-Hung (Intelligent Systems, Control and Automation: Science and Engineering, Vol. 48. Springer Netherlands), 17-26.

Loomis, J., Blascovich, J., and Beall, A. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments & Computers* 31, 557-564.

Macuga, K.L., Beall, A.C., Kelly, J.W., Smith, R.S., and Loomis, J. M. (2007). Changing lanes: inertial cues and explicit path information facilitate steering performance when visual feedback is removed. *Experimental Brain Research* 178, 141-150.

McFarland, D.J., Miner, L.A., Vaughan, T.M., and Wolpaw, J.R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography* 12, 177-186.

Palmer, S. E. (1989). "*Reference frames in the perception of shape and orientation,*" in Object Perception: Structure and Process, eds. B. Shepp and S. Ballesteros (Hillsdale, NJ: Lawrence Erlbaum Associates), 121-163.

Pani, J. R., and Dupree, D. (1994). Spatial reference systems in the comprehension of rotational motion. *Perception* 23, 929-946.

Parsons, L. M. (1987). Visual discrimination of abstract mirror-reflected three-dimensional objects at many orientations.

Perception & Psychophysics 42, 49-59.

Parsons, L. M. (1995). Inability to reason about an object's orientation using an axis and angle of rotation. *Journal of Experimental Psychology: Human Perception and Performance* 21, 1259-1277.

Pausch, R., Proffitt, D., and Williams, G. (1997). *Quantifying Immersion in Virtual Reality*. ACM SIGGRAPH '97 Conference Proceedings, Computer Graphics.

Peronnet F., and Farah, M.J. (1989). Mental rotation: an event-related potential study with a validated mental rotation task. *Brain and Cognition* 9, 279-288.

Pfurtscheller, G., Brunner, C., Schlogl, A., and Lopes da Silva, F.H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage* 31,153-159.

Rock, I. (1986). *"The description and analysis of object and event perception,"* in Handbook of Perception and Human Performance, eds. K. Boff, L. Kaufman and J. Thomas (New York, NY: John Wiley).

Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701-703.

Van Orden, K, and Broyles, J. (2000). Visuospatial task performance as a function of two- and three-dimensional display presentation techniques, *Displays*, 21, 17-24.

Wexler, M., Kosslyn, S.M., and Berthoz, A. (1998). Motor processes in mental rotation. *Cognition* 68,77-94.

Wijers, A.A., Otten, L.J., Feenstra, S., Mulder, G., and Mulder, L.J. (1989). Brain potentials during selective attention, memory search, and mental rotation. *Psychophysiology* 26,452-467.

Witmer, B.G., and Singer, M.J. (1998) Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 225-240.

WorldViz. (2007). Vizard VR Toolkit R3: http://www.worldviz.com/products/vizard4/index.html