Effects of Interactivity and 3D-motion on Mental Rotation Brain Activity in an Immersive Virtual Environment

Daniel Sjölie[1], Kenneth Bodin[1], Eva Elgh[2], Johan Eriksson[3], Lars-Erik Janlert[1], Lars Nyberg[3]

Department of Computing Science [1]
Umeå University
S-901 87 Umeå, Sweden
{daniel,bodin,lej}@cs.umu.se

Department of Community Medicine and Rehabilitation, Geriatric Medicine [2]
Umeå University
S-901 87 Umeå, Sweden
eva.elgh@germed.umu.se

Department of Integrative Medical Biology [3]
Umeå University
S-901 87 Umeå, Sweden
{johan.eriksson, lars.nyberg}@physiol.umu.se

ABSTRACT

The combination of virtual reality (VR) and brain measurements is a promising development of HCI, but the maturation of this paradigm requires more knowledge about how brain activity is influenced by parameters of VR applications. To this end we investigate the influence of two prominent VR parameters, 3d-motion and interactivity, while brain activity is measured for a mental rotation task, using functional MRI (fMRI). A mental rotation network of brain areas is identified, matching previous results. The addition of interactivity increases the activation in core areas of this network, with more profound effects in frontal and preparatory motor areas. The increases from 3d-motion are restricted to primarily visual areas. We relate these effects to emerging theories of cognition and potential applications for brain-computer interfaces (BCIs). Our results demonstrate one way to provoke increased activity in task-relevant areas, making it easier to detect and use for adaptation and development of HCI.

Author Keywords

Virtual Reality, fMRI, Reality-Based Interaction, VRfMRI, brain imaging, BCI.

ACM Classification Keywords

H5.1 [Information Interfaces and Presentation]: Multimedia Information Systems — Virtual Reality, H5.2 [Information Interfaces and Presentation]: User Interfaces, H.1.2 [Models and Principles]: User/Machine Systems — human factors;

General Terms

Human factors, Measurement, Theory.

INTRODUCTION

The advantages of virtual reality (VR) have been apparent for a long time, but it is only recently that the technological development has allowed widespread implementation [34]. Continued development of efficient VR-interaction is hampered by the difficulty of evaluating interaction with the level of complexity and freedom often desired in virtual environments (VEs), a problem recently described within the framework of reality-based interaction (RBI) [25,7]. One promising way forward is offered by modern neuroscience and developments in technologies for brain measurement and brain-computer interfaces (BCIs) [11,37,30]. Measuring brain activity while interacting naturally with a system makes it possible to correlate activity in specific brain areas, or patterns of activation in distributed networks, to hidden cognitive states, such as mental workload, and in turn relate these hidden states to aspects of the interface and the interaction [24]. This also opens the door for using these brain measurements as a passive or adaptive BCIs and modify the interaction and the environment "on the fly" [11,20], e.g., to lower the speed of a monitored process if the mental workload of the operator becomes too high [32].

We believe that VR and adaptive BCI together constitute a particularly potent combination for the implementation of human-computer interaction (HCI) when sensitivity and adaptability is critical. If we have an understanding of how the brain reacts to different aspects of a complex, realistic and interactive environment we can relate measurements of brain activity to these aspects and adapt the interaction to optimize it for the current user and her state of mind. Using brain measurements to adapt interaction in such a way, without requiring the user to take conscious action, differs from the common use of BCI as a control channel. This passive adaptation approach adds extra value in applications of VR that already focus on the function of the brain, such as neuropsychological tests, rehabilitation, or research, but the potential gains are widely applicable. It is often desirable to keep the difficulty of a task and the resulting cognitive workload within a certain span to maximize training [14], efficiency and/or enjoyment [38]. Systems building on this combination of VR and BCI are becoming more common, but in order for the paradigm to become mature and reliable we need more research into the function of the brain in a VR environment and the influence of VR application parameters [29]. We need to know how brain measurements are affected by such parameters as 3d-motion, interactivity, presence, etc.

Functional MRI

The measurements of brain activity presented in this paper were made using functional magnetic resonance imaging (fMRI). Compared to other brain measurement techniques fMRI has a number of important advantages. It is possible to combine measurements with good spatial resolution from the entire brain (not possible with electroencephalography (EEG) or functional near-infrared spectroscopy (fNIRS)) with subject safety and decent temporal resolution (compared to positron emission tomography (PET)). A combination of VR and fMRI was first used by Aguirre [1] in 1996 and has slowly become more popular since then. Many of the published results have focused on the feasibility of the method and on describing problems and solutions; Beck et al [4] presents a review of this research. In a recent paper Jäncke et al [26] present a review of the neural underpinnings of presence, thus addressing brain functions in relation to one of the most central parameters in VR application. Presence is the subjective experience of being present in a real place. Still, the setup and management of a complete VRfMRI-system remains a complex procedure requiring a multidisciplinary research team and much remains to be investigated about the function of the brain while interacting in a VR environment.

Cognitive Grounding

When investigating the function of the brain in VR it is appropriate to consider theories of cognition and brain function in general. What support can be found for the idea that the way the brain works in a computer-generated reality is similar to how it works in our physical reality? Can theories of cognition be leveraged to improve the ability to interpret brain activity and adapt interaction? An interesting way to address these questions, prevalent in discussions about cognition in general, is to consider the nature of representations in the brain. Can we expect representations of real and virtual objects or phenomena to be similar? Research into the phenomenon of mental imagery has gathered a considerable body of work relating to this question over several decades, and today the relation between imagining something (mental imagery) and perceiving something real is well established [28].

In a recent review, Postma and Barsalou [33] develop the connection between mental imagery and working memory, and relate these phenomena to the concepts of grounded cognition and mental simulation. Grounded cognition focus on the importance of grounding higher cognition in the modal systems of the brain and mental simulation is put forward as a central aspect of cognition. Mental simulation

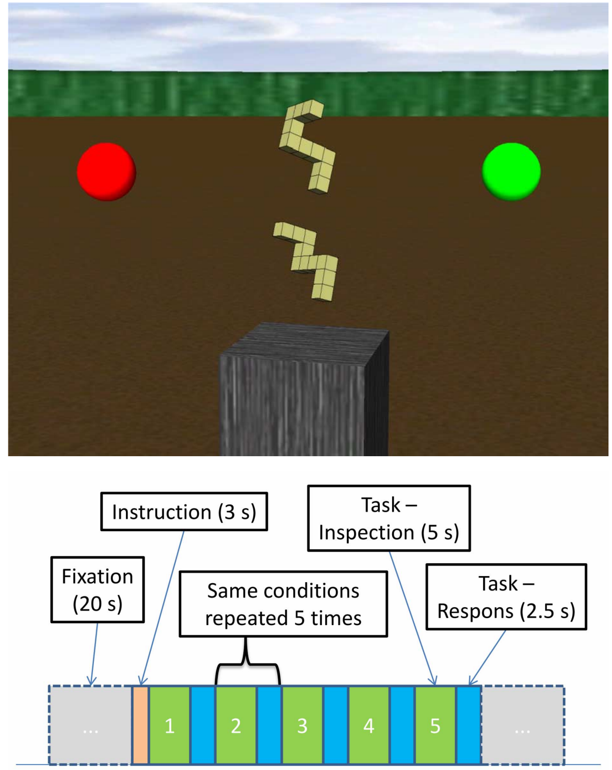

Figure 1. Top: View of the two 3d-objects to be compared, hovering over a pedestal placed in a 3d-environment with textured ground, sky and horizon. The spheres only serve as a reminder on how to respond. In this image, the 3d-objects are identical. Bottom: Time course of a block of tasks with the same condition. The experiment consists of 18 such blocks, divided into three repetitions of all 6 variations of the task, randomly ordered within each repetition.

extends mental imagery to encompass the unconscious and flexible reactivation of memories to recognize the current context and predict possible actions and expected results (see [3] for a review). This idea of predictions building on previous experience as a critical aspect of how the brain works has gathered increasing support in recent years. For example, following a discussion of the striking similarities in brain activation when remembering the past and when imagining the future, Schacter et al remark that they "find it helpful to think of the brain as a fundamentally prospective organ that is designed to use information from the past and the present to generate predictions about the future" (page 660) [35]. Considering that this view of cognition is currently being applied to areas such as artificial intelligence [21,22] and model-based analysis of brain imaging data [13], this view may indeed be helpful when combining measurements of brain activity with HCI.

In a study with particular relevance for VR and HCI, Decety and Jeannerod [12] examine mentally simulated actions in a VR environment and discuss the results in relation to Fitts's law, one of the most successful and well-studied models within HCI. Fitts's law models interaction

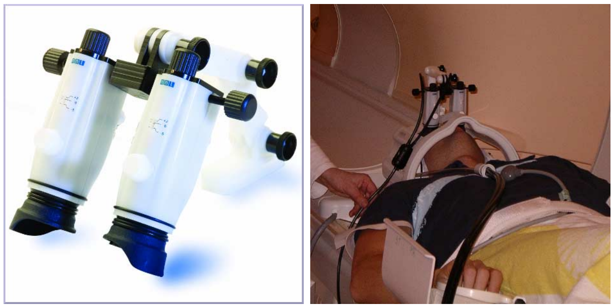

Figure 2. Left: MR-compatible video goggles from NordicNeuroLab. Right: The goggles setup in our study.

by relating the speed at which an action can be executed to the required accuracy [17]. This law was evaluated by instructing subjects to imagine walking through virtual gates with different widths, starting at different distances. The time required to imagine this action correlated with both the distance to and the width of the gate, further supporting the idea that mental simulation of action is governed by the same rules as motor action in general. In total, the previous work reviewed here provides a foundation for reasoning about brain functions in a VR environment based on what we know about cognition in general.

Mental Rotation

In order to improve our understanding of the impact of different aspects of VR on resulting brain measurements, we wanted to study a known cognitive process to ensure a well founded expectation on what the related brain activity should look like. We decided to use "mental rotation", the process of imagining an object being rotated to mentally see it from another angle. This task was first studied by Shepard & Metzler [36] in 1971 and gained some fame because the results showed a clear linear correlation between the angle of mental rotation and the time required to complete the task. This suggests that mental rotation is done at a constant speed, corresponding to normal rotation in physical reality, and served as early support for the idea that the brain functions underlying imagination relate to functions underlying perception of reality, thus encouraging further research into mental imagery.

Over the last decade mental rotation has been studied several times using fMRI [39,8,31] and a network of recurrent brain areas has emerged. The most prominent activations are found in the parietal lobe (intraparietal sulcus (IPS) and superior parietal lobe (SPL)) and in motor-preparation areas (premotor cortex (PM) and the supplementary motor area (SMA)), often together with activations in dorsolateral prefrontal cortex (DLPFC). The DLPFC is of particular interest since it has been suggested as a key node in a network of brain areas associated with the experience of presence [26], while the parietal cortex and motor preparation areas have been related more specifically to the mental rotation task [31].

In the present study we chose to focus on 3d-motion and interactivity as two important aspects of VR interaction. We created a scenario where the subjects perform a variation of the mental rotation task in a simple VE and recorded the brain activity for three different conditions with a varying degree of 3d-motion and interactivity. We also included conditions with and without stereo vision, another common parameter in VR applications. This setup allowed for a comparison of the resulting brain activity to determine the impact of these parameters on the cognitive processes underlying the mental rotation task.

METHOD

Population

Out of 11 initial subjects one was excluded because of poor performance and motion artifacts, leaving 10 healthy subjects, 5 male and 5 female, all students in ages 20-35, recruited on Umeå University campus.

The System

The MR-compatible hardware used was a combination of hardware delivered by NordicNeuroLab (Bergen, Norway), and hardware developed in-house at the department of Integrative Medical Biology, Umeå University. The visual system (Figure 2) consisted of a set of stereo-capable goggles, SVGA, 800 (x3) x 600 pixels, 16.7 million colors, a horizontal/vertical field of view (FOV) of 30°/23°, with accommodation distance at infinity and a possible diopter correction of -5 to +2 dpt. The OLED display used is less sensitive to the electromagnetic fields than alternatives, and all electronics are screened from electromagnetic fields using a Faraday cage. The metal net of the cage is out of focus in the display and introduces very little visual disturbance. The use of these goggles makes it possible for the subject to shut out the real-world surroundings in the MR-scanner and become immersed in our virtual environment. The VR-software-system was based on Colosseum3D [2], developed at VRlab, Umeå University. See Sjölie et al [37] for further details on the system.

General Procedure

Our experiment placed the subject in a 3-dimensional virtual environment with textured ground and sky, a distant horizon, and a central 1-meter high pedestal (Figure 1, top). Any movement in the virtual environment was restricted to the subject's viewpoint circling around the pedestal or moving up or down, giving the subject different perspectives but always centered at the pedestal. The rendered perspective was constructed with a viewpoint 3 meters away from and slightly above the pedestal, with a vertical FOV of 60°. A FOV greater than the display FOV was chosen to reduce the apparent tunnel effect and enable a focus on the behavioral realism of the environment. The interpupillary distance (IPD) used for the stereo rendering was set to 26 mm (less than normal human IPD) based on subjective judgments of a comfortable stereo effect during pilot testing. When the mental rotation task was performed

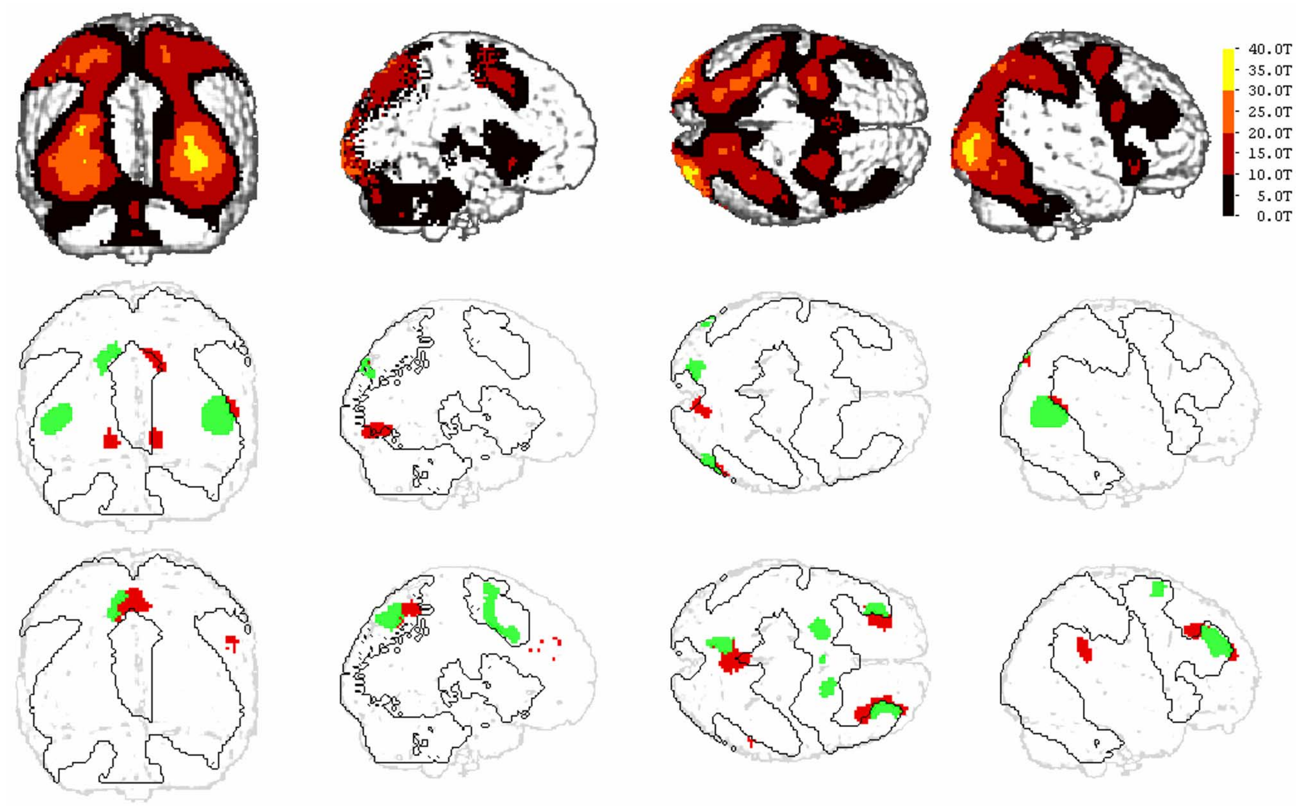

Figure 3. Brain areas significantly activated for the mental rotation task in general (top), and areas with increased activity for 3d-motion (middle) and interactivity (bottom). Increased activations are displayed as within (green/bright) or outside (red/dark) of the network (black outline). Images are surface renderings showing activations to a depth of 20 mm, with caudal, right medial, dorsal and right lateral views, from left to right. Figure is in color in the electronic version.

two geometric 3d-objects, either identical or mirrored, were placed above the pedestal (Figure 1, top). The objects start out with a random orientation (selected from a pre-randomized set, excluding occluding orientations) and the task was to determine whether or not the objects were identical.

To examine the effects of 3d-motion and interaction the task was performed under three different conditions; one condition without motion (*still*), one with automatic, non-interactive, motion (*auto*) and one where the motion was controlled interactively by the subject (*interactive*). Each of these conditions was presented with and without stereo-vision. The automatic motion circles the pedestal at a constant speed while moving smoothly up and down with the pedestal in the center. When interactive rotation was possible the subject could use the two buttons in each hand to control left/right circling and up/down motion respectively.

For each condition the subjects were presented with a blocked sequence of five pairs of objects to judge, with 5 seconds to inspect the objects, followed by 2.5 seconds to respond, for each pair (Figure 1, bottom). Before each block a sign was displayed for 3 seconds, instructing the subject to inspect and compare the objects and, in the interactive condition, to rotate interactively around the pedestal. We also had a 20 seconds rest-period between each block, to use as a baseline. This allows us to compare the brain activity in a resting state against the brain activity during the task. During this rest-period the virtual environment was still visible and the subject was free to look around (gaze only, the head was always fixed) but there were no complex objects and no motion. In summary, these three conditions, with and without stereo-vision, gave a total of six variations of task blocks. These six variations were presented in random order and this was repeated three times, giving a total of 18 blocks with 90 pairs of pieces to compare per experiment.

Behavioral Data

The behavioral data collected was, for each piece pair: whether the answer given was correct, how fast the subjects responded and how much the subject interacted during examination. The subjects were also debriefed after the experiment. We used this data to ensure that any differences in activation between conditions were not caused by differences in difficulty, by comparing response time and error percentages between conditions. Questions about the experience of stereo effect were also included in the debriefing. One subject reporting abnormal stereo vision was excluded from the test for any significant differences in measured brain activity related to the use of stereo.

Neuroimaging Procedure

The current fMRI study was carried out on a Philips 3.0 tesla imaging device (MR-scanner), with a repetition time of 1512 ms and all other functional scanning parameters as in Eriksson et al [15]. All images were sent to a PC and converted to Analyze format. The VR-application was synced to the scan intervals of the MR-scanner before each block.

Statistical Analysis

The statistical analysis of the data was done following fMRI-analysis practice. The brain activity data was treated as a set of images with three dimensions, consisting of voxels representing areas of the brain. The data from each subject consisted of a series of such images (scans) for all time points from the experiment. See Beck et al [4] for additional background. In this study, the time between each image was approximately 1.5 seconds and a total of 758 images per subject were recorded during the designed experiment.

The recorded data was analyzed using SPM5 (Wellcome Department of Cognitive Neurology, London, UK) on Matlab 7.7 (Mathworks Inc, MA, US). The pre-processing applied to all images included slice-timing correction, realignment, normalization to standard anatomical space defined by the MNI atlas, and smoothing with an 8.0 mm FWHM Gaussian filter kernel. To estimate the effect of different conditions on brain activity we used the general linear model (GLM) to create statistical parametric maps with t-statistics. Since the mental rotation took place during the inspection phase we modeled six regressors to cover this period for all variations of the task, i.e., the three conditions, with and without stereo. To ensure that the subjects performed the mental rotation successfully we excluded any inspection phases that lead to an incorrect answer. The inspection phase for these incorrect pieces was modeled separately as a regressor of no interest as was the response phase. The rest-period was included as a baseline regressor. All these regressors were constructed as boxcars convolved with the canonical hemodynamic response function (HRF). We also added regressors of no interest for the motion correction acquired from the realignment pre-processing step. For the estimation of this model we used a high-pass filter with a cutoff of 720 s. Finally, contrasts against baseline were constructed for each condition and subject. After an initial examination of these contrasts we decided to pool the data from the variants with and without stereo-vision for each condition, since we failed to detect any stable effect of this manipulation. The small IPD used for the stereo rendering might have contributed to this. This left us with three contrasts of interest per subject, one for each variation of the task (stills, auto, interactive).

The single-subject contrasts representing the effect of the different conditions were entered into a group analysis

T-values for local maxima						
Stills	**Auto**	**Interactive**	**3d-motion**	**Interactivity**	**Coordinates**	
Mental rotation						**Visual cortex**
34.1	**37.7**	35.1	2.3	-0.5	-30 -90 16	left
29.0	**32.2**	**32.2**	2.0	1.5	-34 -88 -14	left
38.0	**42.9**	37.5	3.1	-2.2	36 -90 -2	right
						Parietal cortex
30.5	30.3	**33.3**	-0.6	2.5	-38 -40 46	left inferior
22.3	21.8	**22.7**	-0.7	0.6	36 -46 48	right inferior
25.0	23.9	**25.5**	-1.2	0.9	-20 -72 60	left superior
21.2	21.1	**22.3**	-0.3	1.0	20 -68 58	right superior
						Motor areas
21.5	21.8	**26.0**	0.0	3.7	-30 -6 56	left PM
20.2	18.5	**23.9**	-1.5	3.8	26 0 62	right PM
17.6	16.3	17.1	-1.2	0.2	-58 10 34	left inferior PM
13.3	12.0	13.1	-1.1	0.4	50 10 26	right inferior PM
13.6	13.6	**17.5**	-0.2	3.3	0 10 50	SMA
						DLPFC
10.1	10.5	**13.4**	0.2	2.6	-44 32 38	left
10.0	9.2	**13.6**	-0.7	3.3	42 48 26	right
						Subcortical areas
13.7	12.0	**14.0**	-1.4	1.0	30 24 -8	left basal ganglia
14.4	12.8	11.9	-1.4	-1.4	-16 14 -10	right insula
						Cerebellum
12.3	14.1	**17.4**	1.2	3.6	2 -78 -32	medial posterior
14.7	14.3	**17.2**	-0.5	2.3	36 -54 -28	right anterior
8.6	8.6	**10.2**	-0.2	1.4	-24 -40 -46	left inferior
6.4	8.1	**6.6**	1.1	-0.5	20 -36 -50	right inferior
Effect of 3d-motion						**MT/V5**
20.0	**36.3**	25.1	11.9	-2.3	-46 -72 2	left
22.5	**35.9**	26.6	9.7	-2.0	48 -78 6	right
						Cuneus
2.2	**10.8**	7.5	6.4	0.8	-16 -82 32	left
1.8	**11.6**	6.9	7.2	0.2	8 -86 44	right
						Cerebellum
-0.7	**8.2**	4.9	6.6	1.0	-14 -84 -10	left medial
-3.0	**5.3**	3.0	6.2	1.5	12 -80 -8	right medial
Effect of interactivity						**Parietal cortex**
1.0	3.2	**12.6**	1.6	8.6	2 -56 56	precuneus
-5.2	-5.0	**1.5**	0.2	5.4	54 -46 28	right angular gyrus
						DLPFC
-1.7	*-2.4*	**7.4**	-0.5	7.8	-30 38 40	left
5.7	3.6	**10.5**	-1.6	4.8	-28 46 12	left
3.4	4.8	**12.4**	1.0	6.8	34 38 40	right
						Motor areas
16.7	14.3	**23.3**	-2.0	6.4	-20 2 66	left PM
9.8	7.2	**15.3**	-2.1	5.6	18 6 68	right PM
8.8	9.2	**16.6**	0.2	6.3	-10 16 38	SMA/ACC

Table 1. Activation correlated to the different conditions, compared to baseline, presented as t-values at locations of local maximum for the effect of mental rotation, 3d-motion and interactivity. The t-values correspond to the contrasts described in the Statistical Analysis section above. For the main effect contrasts the t-value for the maximum activation is in bold, tinted green, and the minimum activation (corresponding to the t-value for the network) is underlined, tinted red.

using a GLM with subject as one factor and condition as the second factor, thus modeling the between-subject variability and the group-effect of the conditions separately, using the "flexible factorials" option in SPM5. Contrasts for the main effect of the task for each of the conditions and contrasts for the difference between these effects at the group level were constructed. Thus the *main effect contrasts* correspond to the increased brain activity under each of the three conditions, compared to baseline. These are given as t-values describing the size of the effect in relation to the variation between subjects. I.e., large activations are effects that are consistently present across subjects. In this context,

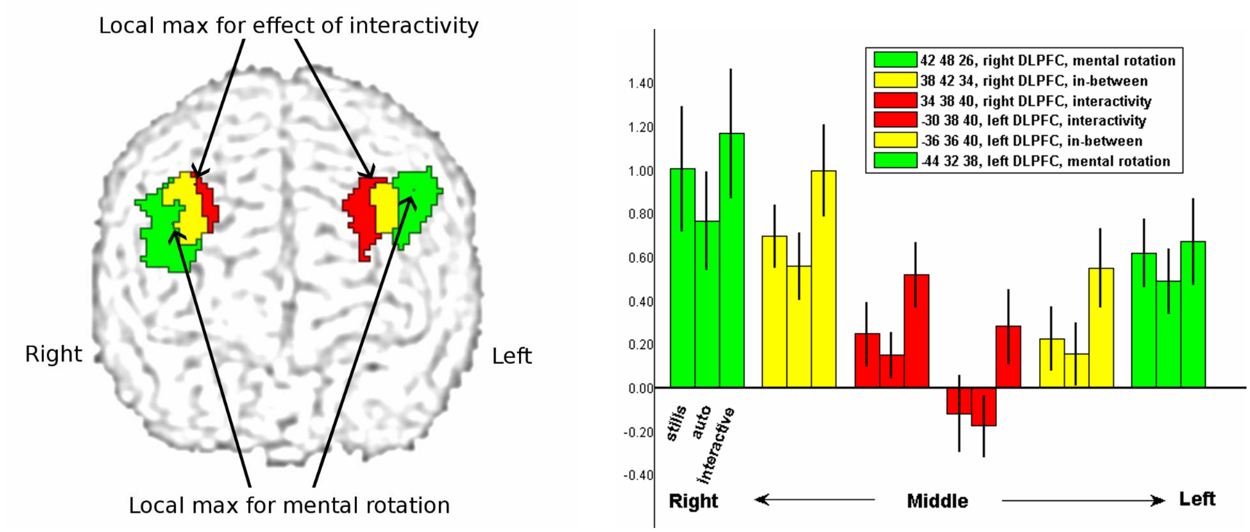

Figure 4. Left: A focused view (from the front, mirrored) of activation in the frontal cortex, masked to include only clusters intersecting the DLPFC. Lateral areas (to the sides, green) are only mental rotation and medial areas (to the middle, red) are only increases for interactivity. The areas in-between on each side (yellow) represent the overlap where the increase for interactivity adds onto the mental rotation activation. Right: Activation correlated to the different conditions, at locations of local maximum for the effect of mental rotation and interactivity in the DLPFC. In-between bars correspond to positions in-between the maxima indicated to the left. The ordering and colors of the bars correspond to the areas on the left.

the mental rotation network was defined as the areas of the brain that showed a significant increase in activation for all conditions of the task. We located this network by constructing a conjunction of the main effect contrasts described above, essentially taking the minimum of the condition effects at each voxel (red/underlined in Table 1).

To capture the effect of motion we contrasted *auto* against *still*, since these conditions were identical except for the automatic rotation. For interactivity we contrasted *interactive* against the average of *auto* and *still*, motivated by the fact that the difference in interactivity is the same, and the amount of motion in *interactive* is between *auto* and *still*. In order to examine how the effects of 3d-motion and interactivity were related to the mental rotation network we divided these effects into activations inside and outside of the rotation network. This was done by applying a mask containing the rotation network activations to these effect contrasts, inclusive and exclusive respectively to capture activations inside and outside of the network.

To ensure the statistical significance of results presented here we used a voxel-threshold of $p<0.01$, FDR corrected for multiple comparisons, and a cluster threshold of 0.01, uncorrected, for all contrasts. To further inspect areas of particular interest we plotted beta-values representing the activation for each condition at local maximum voxels in clusters of interest.

RESULTS

Behavioral Results

Although there is a significant correlation between response time and number of errors, an ANOVA on subject means for these variables does not show any significant effect of condition, $F(2,27) = 0.9$, $p = 0.42$, for errors and $F(2,27) = 2.17$, $p = 0.13$, for reaction time. This is also corroborated by data from the debriefing where the subjective reports on which condition was the most difficult varied greatly.

Mental Rotation Network

The mental rotation network includes large areas of bilateral activation (Figure 3, top). Most activated areas match previous results for mental rotation well. Apart from activations in occipital cortex the strongest activations are indeed found in SPL, PM and SMA, extending down towards anterior cingulate cortex (ACC). There are also weaker activations in DLPFC, cerebellum, insula and the basal ganglia. We present t-values for the most significant local maxima throughout this network (Table 1, top).

Effects of 3d-motion

Areas showing significant increases for the effect of motion are restricted to the posterior part of the brain (Figure 3, middle). The strongest activations can be seen in the MT/V5 areas, bilaterally. In addition to these MT/V5 activations we also find weaker activations on the medial borders of the occipital lobe, dorsally near the parietal cortex (cuneus) and ventrally near the cerebellum. We can see from the t-values for these locations (Table 1, middle) that the ventral positions have a deactivation in the *still* condition that we are comparing against, and are clearly outside of the mental rotation network (red/dark in Figure 3, middle). For the dorsal activations we see a different pattern where the left activation is primarily within the mental rotation network, adding onto the existing activity. The right cluster, however, is outside of the network with *auto* being the only condition with significant activation.

We can also see that the activation in *interactive* is in-between the activation for *still* and *auto*, for all the areas with an effect of 3d-motion. This reflects the fact that the amounts of motion in these conditions follow the same pattern, with a maximal amount of motion in *auto*, a minimal amount of motion (none) in *still*, and a varying amount of motion, in-between these, in *interactive*.

Effects of Interactivity

The general pattern of activation resulting from the addition of interactivity is that the affected areas are more frontal and more medial, with the most posterior activation, in posterior parietal cortex, having a clearly medial local maximum in the precuneus (Figure 3, bottom). All activations but one represent increases that lie within the mental rotation network, at least in part (green/bright in Figure 3, bottom). The exception is a cluster at the right angular gyrus where the *interactive* condition stands out only in relation to marked decreases in activation for the other two conditions (Table 1, bottom). For both parietal and DLPFC activations the clusters is partly inside of the network, extending into a more medial direction while increases in PM and SMA are completely within the network.

To understand how the brain activity at the local maxima listed in Table 1 relates to the overlapping clusters of activation we can consider the clusters intersecting DLPFC as an illustrative example (Figure 4). The local maxima for interactivity are found a small distance from the corresponding maxima for the general mental rotation network (Figure 4, left) and the changes from large general effect to more specific effect of interactivity is gradual in-between (Figure 4, right). The general effect of mental rotation is reduced while the increased activity from interactivity remains. This is true for all the interactivity effect activations within the mental rotation network. In the precuneus a large portion of the strong parietal activation for mental rotation is maintained with interactivity, at the medial maximum. It should be noted that the local maxima for the total effect of the *interactive* condition in the DLPFC clusters are within the network, and close to the local maxima for the *auto* and *still* conditions. It is the maxima of the *increase* of activity that, compared to the other conditions, lie outside of the network. In this context the medial parietal maxima stand out as the exception, with a local maximum outside of the mental rotation network for the *interactive* condition in this area.

DISCUSSION

The focus of this study was to investigate how different parameters of VR applications affect measurements of brain activity, and discuss related opportunities and potential problems. We specifically decided to look at variations in the presentation and interaction possibilities in an immersive VR environment, and at how these would affect the brain activity when performing a mental rotation task. Thus the brain activity results of greatest interest are the significant differences in activation when comparing the task with and without 3d-motion, and with and without interactivity. The general mental rotation network, constituted by areas significantly activated for all variations of the task, matches earlier results very well and we focus our initial discussion of the brain activity on the areas that are well established as part of this network in the literature.

The most significant activations within this mental rotation network were found in SPL and in the occipital lobe. In SPL the level of activation was high for all conditions with relatively little variation between them, further supporting the existing notion that this is the most critical area for mental rotation. However, the strongest activation was actually found in the occipital cortex, an area that has been inconsistently reported in previous mental rotation studies. This is probably because of differences in the design of experiments and the choice of baseline for task activation contrasts. In our study we have a baseline with no reason to attend to any particular part of the visual environment and no counterpart to the complex 3d-objects to be inspected in the task. We speculate that, in line with the results of Mourao-Miranda [31], the strong activation in visual areas is in large part related to the general visual inspection component of the task, and not specifically to the mental rotation component. In particular, this activation is probably enhanced by the requirement to attend to and encode the task-related visual stimuli. The subjects are free to look around and visually inspect the environment even during the baseline-period but there is nothing motivating them to do so and the environment is simple enough to become familiar and uninteresting in short order. Thus, excluding the occipital lobe, we focus on the *core mental rotation network* consisting of SPL, PM, SMA and DLPFC.

Since the effect of 3d-motion did not show any significant increase in activation within this *core* mental rotation network we suggest that the addition of the rotating 3d-objects did not make any significant difference at higher levels in the brain since the subjects were already imagining and mentally working with rotating 3d-objects even in the *still* condition. It seems reasonable that the areas that do show a significant increase in activity are more related to the perception of motion and moving visual stimuli in general, with the strongest activation in MT/V5, an area well documented as sensitive to motion [6]. It is interesting to note that these increased activations are almost completely within the larger mental rotation network (green in Figure 3), suggesting that there was motion sensitive activation even in the *still* condition, in spite of the fact that this condition had no visible motion.

Another possible factor is related to the issue of perspective, or agency. That is, whether you feel that you are controlling events yourself or if you are merely an observer. In the contrast used to evaluate the effect of motion we are comparing *auto* to *still*, meaning that we do have an addition of motion *outside of your control*, compared to the complete control when you are imagining the motion. This might contribute to the increase in activity in the cuneus, since this area is adjacent to the inferior parietal lobe which has been implicated as an area of increased importance when observing passively [16]. Another brain area reported in relation to the issue of agency is the anterior insula. In this case the correlation is in the other direction, with increased activation for egocentric perspectives [16]. The t-values for the local maxima in the right anterior insula provide some initial support for this interpretation, showing less activity for *auto* than for the *still* condition. However, if this factor was the cause of any significant effect in general we would expect a comparison of the *auto* condition and the *interactive* condition, where the subject controls the motion herself, to show a difference in activation. But there are no significant activations at all in such a contrast at the selected threshold and only MT/V5 activations at a slightly more liberal threshold. This speaks against any feeling of being unable to control events as being a significant contributor to the measured brain activity in general.

The most interesting and strongest effect, however, is the effect of interactivity. The increased activation in the right angular gyrus is of some interest since this location is close to the temporoparietal junction (TPJ), an area previously related to the multisensory integration of body-related information, out-of-body experiences [5] and the related impact on the sense of presence [26]. However, this effect is hard to evaluate in the present study, since the difference in activation depends more on reduced activity for the *still* and *auto* conditions than any increase in activity for the *interactive* condition (Table 1, bottom). This area should be further investigated in future work, with a focus on multisensory integration and presence. Excluding this activation, we can see that all other activated areas overlap with the mental rotation network (Figure 3, bottom). All of these areas have been previously reported in conjunction with spatial working memory and attention. In a review of the neural circuit basis for spatial working memory [9], Constantinidis and Wang remark that imaging studies on working memory in humans almost invariably show concurrent prefrontal and parietal activation, and that spatial working memory in particular shows increased activation in PM and SMA. In particular, the DLPFC and posterior part of the parietal lobe are emphasized, thus matching our results well. This suggests that this increase in activation can be largely attributed to increased demands on working memory and increased attention to existing representations. However, given our present questions we suggest that it is useful to consider the implications of these activations in relation to the cognitive theories mentioned in the introduction. Importantly, the critical role of *prediction error* in these theories [18] allows us to reason about how different aspects of our environment give rise to increased brain activity. In short, it is hypothesized that the brain works by predicting what comes next, and if the prediction is correct little further reaction is needed. Objects and phenomena that are easy to predict should be represented with little effort, resulting in less brain activity compared to phenomena that are more unpredictable. From this perspective, the increased activity with interactivity can be related to an increased variability and unpredictability in the environment. In addition to the need to update representations while interacting with the environment the interaction itself was also somewhat unfamiliar to the subjects, giving rise to additional difficulty in prediction. The automatic rotation in the *auto* condition follows a smooth, regular, path that subjects can grasp very quickly, and in the *still* condition the subjects know that nothing will happen during the inspection phase. Thus perceived events are easily predictable in both of these conditions, leading to less brain activity compared to the *interactive* condition, and no significant difference in activation between the two in more frontal areas.

The more frontal nature of the effect of interactivity also fits well with an important aspect of how the prediction errors mentioned above are processed in the brain. According to this theory, predictions are made in relation to a context and the brain consists of a complex hierarchy where contexts are defined at higher levels and predictions are checked at lower levels [27]. When prediction errors occur they are passed upwards and assumptions about the context are adjusted as needed. Thus environments that are harder to predict lead to more prediction errors being fed upwards and in turn lead to increased activity in higher, more frontal, regions. One implication of this is that we should be able to adapt our environment and interaction in order to provoke increased activity in frontal areas, such as DLPFC. This is promising, both because frontal regions make other techniques for brain measurement, such as fNIRS [19], feasible, and because the DLPFC has been implicated in several relevant functions, such as spatial working memory and presence. However, while the importance of the DLPFC is promising for potential BCI applications, it should be noted that the ability of fMRI to measure brain activity across the entire brain is still very valuable. As exemplified by the mental rotation network, results from brain imaging studies on cognitive tasks and processes often present distributed networks of cooperating areas across the brain. If there is something to the ideas behind grounded cognition we can expect this pattern to continue. Cognition and representations are not stored in any one location of the brain but distributed across modalities and the corresponding brain areas. For example, in a review focusing on the DLPFC and its role in a working memory network, Curtis and D'Esposito [10] discuss the nature of representations for spatial working memory. They suggest that the representations in question are stored in more posterior regions and that the DLPFC primarily aids in their maintenance by continually directing attention to the relevant representations. These representations are critically related to preparing motor responses, for instance by reactivation of motor programs for eye movements related to certain spatial locations without actually executing these movements, i.e., mental simulations. We can recognize similar conclusions in a recent review of the role of the DLPFC for presence in VR. In this review presence, commonly described as the sense of "being there", is related to the ability to "do there", and the critical importance of being able to relate the virtual space to your "real motor space" and build on familiar simulations of real motor responses [26]. This focus on representations of familiar interactions and motor responses also serves as an additional grounding for the description of RBI [25], explicitly relating the themes of familiarity and skill with body, environment, etc, to a cognitive theory and brain function.

Finally, one primary challenge when we want to adapt interaction based on the mental workload of the user is that it can be hard to measure, especially when we consider that mental workload can be decomposed into multiple components. The increased activity in our results, correlated with interactivity, can be related to an increased mental workload, triggered by an environment that is more dynamic and harder to predict, thereby forcing the user's brain to work to make new predictions. This effect was seen even though we found no significant differences in behavior or reported subjective difficulty for the interactive condition. Thus we have a measure of mental workload, in a dynamic reality-based interaction environment, which we could not easily measure by considering behavior. The ability to measure mental workload that cannot be otherwise measured has great potential for the development of adaptive brain-computer interfaces [11,20] and evaluation of HCI models that are otherwise hard to evaluate, such as RBI [37,23,24]. In addition, our results and this discussion illustrate a promising way to provoke increased workload in specific modalities, providing a way to probe the limits of detection and calibrate measurements of brain activity to specific users and interaction solutions.

CONCLUSIONS

The present study illustrates the value of interactive and dynamic environments for obtaining brain activity that is easier to detect and thus the potential for reality-based brain-computer interaction. We show how one can adapt the interaction in a VR setting to provoke increased brain activity in areas identified as part of a core network for the central task (mental rotation) by increasing the level of interactivity. This is particularly interesting for training and diagnostics, where it can be of great value to be able to produce the desired level of mental workload to optimize efficiency or sensitivity. The continued grounding of these results in theories of cognition and brain function holds great promise, but more research is needed to evaluate what is truly possible. Additionally, it may be important to understand these functions in order to compensate for unintended effects.

ACKNOWLEDGMENTS

Thanks to Gösta Bucht, Anne Larsson and Anders Backman for involvement in the project.

Thanks to Swedish Brainpower, Umeå University, the Kempe foundations (JCK-2613) and Göran Gustafsson Foundation (award in medicine, 2007, to Professor Lars Nyberg) for funding.

REFERENCES

1. Aguirre, G.K., Detre, J.A., Alsop, D.C., and D'Esposito, M. The parahippocampus subserves topographical learning in man. *Cereb. Cortex 6*, 6 (1996), 823-829.
2. Backman, A. Colosseum3D–Authoring Framework for Virtual Environments. In *Proc. EUROGRAPHICS Workshop IPT & EGVE Workshop*, (2005), 225-226.
3. Barsalou, L.W. Grounded Cognition. *Annual Review of Psychology 59*, 1 (2008), 617-645.
4. Beck, L., Wolter, M., Mungard, N., Kuhlen, T., and Sturm, W. Combining Virtual Reality and Functional Magnetic Resonance Imaging (fMRI): Problems and Solutions. In *HCI and Usability for Medicine and Health Care*. 2007, 335-348.
5. Blanke, O. and Arzy, S. The Out-of-Body Experience: Disturbed Self-Processing at the Temporo-Parietal Junction. *Neuroscientist 11*, 1 (2005), 16-24.
6. Born, R.T. and Bradley, D.C. Structure and Function of Visual Area MT. *Annual Review of Neuroscience 28*, 1 (2005), 157-189.
7. Christou, G., Law, E.L., Green, W., and Hornbaek, K. Challenges in evaluating usability and user experience of reality-based interaction. In *Proc. CHI 2009 Extended Abstracts*, ACM (2009), 4811-4814.
8. Cohen, M.S., Kosslyn, S.M., Breiter, H.C., et al. Changes in cortical activity during mental rotation A mapping study using functional MRI. *Brain 119*, 1 (1996), 89-100.
9. Constantinidis, C. and Wang, X. A neural circuit basis for spatial working memory. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry 10*, 6 (2004), 553-65.
10. Curtis, C.E. and D'Esposito, M. Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences 7*, 9 (2003), 415-423.
11. Cutrell, E. and Tan, D.S. BCI for passive input in HCI. In *Proc. CHI 2008 Workshop on Brain-Computer Interfaces for HCI and Games*, (2008).
12. Decety, J. and Jeannerod, M. Mentally simulated movements in virtual reality: does Fitt's law hold in motor imagery? *Behavioural Brain Research 72*, 1-2 (1995), 127-134.

13. Dolan, R. Neuroimaging of Cognition: Past, Present, and Future. *Neuron 60*, 3 (2008), 496-502.
14. Ericsson, K.A., Prietula, M.J., and Cokely, E.T. The making of an expert. *Harvard business review 85*, 7/8 (2007), 114.
15. Eriksson, J., Larsson, A., and Nyberg, L. Item-specific Training Reduces Prefrontal Cortical Involvement in Perceptual Awareness. *Journal of Cognitive Neuroscience 20*, 10 (2008), 1777-1787.
16. Farrer, C. and Frith, C.D. Experiencing Oneself vs Another Person as Being the Cause of an Action: The Neural Correlates of the Experience of Agency. *NeuroImage 15*, 3 (2002), 596-603.
17. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology 47*, 6 (1954), 381-391.
18. Friston, K. and Stephan, K. Free-energy and the brain. *Synthese 159*, 3 (2007), 417-458.
19. Girouard, A., Hirshfield, L.M., Solovey, E., and Jacob, R.J.K. Using functional Near-Infrared Spectroscopy in HCI: Toward evaluation methods and adaptive interfaces. In *Proc. CHI 2008 Workshop on Brain-Computer Interfaces for HCI and Games*, (2008).
20. Girouard, A. Adaptive brain-computer interface. In *Proc. CHI 2009 Extended Abstracts,* ACM Press (2009), 3097-3100.
21. Hawkins, J. *On Intelligence*. Owl Books, 2005.
22. Hawkins, J. and George, D. Hierarchical Temporal Memory. Numenta Inc., 2006.
23. Hirshfield, L.M., Solovey, E.T., Girouard, A., et al. Using Brain Measurement to Evaluate Reality Based Interactions. In *Proc. CHI 2009 Workshop on Challenges in the Evaluation of Usability and User Experience in Reality Based Interaction*, (2009), 19.
24. Hirshfield, L.M., Solovey, E.T., Girouard, A., et al. Brain measurement for usability testing and adaptive interfaces: an example of uncovering syntactic workload with functional near infrared spectroscopy. In *Proc. CHI 2009*, ACM Press (2009), 2185-2194.
25. Jacob, R.J.K., Girouard, A., Hirshfield, L.M., et al. Reality-based interaction: a framework for post-WIMP interfaces. In *Proc. CHI 2008,* ACM Press (2008), 201-210.
26. Jäncke, L. Virtual reality and the role of the prefrontal cortex in adults and children. *Frontiers in Neuroscience 3*, 1 (2009).
27. Kilner, J.M., Friston, K.J., and Frith, C.D. Predictive coding: an account of the mirror neuron system. *Cognitive Processing 8*, 3 (2007), 159-66.
28. Kosslyn, S.M., Ganis, G., and Thompson, W.L. Neural foundations of imagery. *Nat Rev Neurosci 2*, 9 (2001), 635-642.
29. Lécuyer, A., Lotte, F., Reilly, R.B., Leeb, R., Hirose, M., and Slater, M. Brain-Computer Interfaces, Virtual Reality, and Videogames. *Computer 41*, 10 (2008), 66-72.
30. Minnery, B.S. and Fine, M.S. Neuroscience and the future of human-computer interaction. *Interactions 16*, 2 (2009), 70-75.
31. Mourao-Miranda, J., Ecker, C., Sato, J.R., and Brammer, M. Dynamic Changes in the Mental Rotation Network Revealed by Pattern Recognition Analysis of fMRI Data. *Journal of Cognitive Neuroscience 21*, 5 (2009), 890-904.
32. Parasuraman, R. and Wilson, G.F. Putting the brain to work: neuroergonomics past, present, and future. *Human Factors 50*, 3 (2008), 468-474.
33. Postma, A. and Barsalou, L.W. Spatial working memory and imagery: From eye movements to grounded cognition. *Acta Psychologica 118,* 2-3 (2009).
34. Rizzo, A.A., Schultheis, M., Kerns, K.A., and Mateer, C. Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation 14*, 1 (2004), 207-239.
35. Schacter, D.L., Addis, D.R., and Buckner, R.L. Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci 8*, 9 (2007), 657-661.
36. Shepard, R.N. and Metzler, J. Mental Rotation of Three-Dimensional Objects. *Science 171*, 3972 (1971), 701-703.
37. Sjölie, D., Bodin, K., Eriksson, J., and Janlert, L. Using brain imaging to assess interaction in immersive VR. In *Proc. CHI 2009 Workshop on Challenges in the Evaluation of Usability and User Experience in Reality Based Interaction*, (2009), 23.
38. Sweetser, P. and Wyeth, P. GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain. 3*, 3 (2005), 3-3.
39. Tagaris, G.A., Kim, S., Strupp, J.P., Andersen, P., Uğurbil, K., and Georgopoulos, A.P. Mental Rotation Studied by Functional Magnetic Resonance Imaging at High Field (4 Tesla): Performance and Cortical Activation. *Journal of Cognitive Neuroscience 9*, 4 (1997), 419-432.