CyberPsychology & Behavior
Volume 12, Number 3, 2009

DOI: 10.1089/cpb.2009.0104

Rapid Communication

Auditory Cues Increase the Hippocampal Response to Unimodal Virtual Reality

Joseph Andreano,[1] Kevin Liang,[1] Lingjun Kong,[1] David Hubbard,[2]
Brenda K Wiederhold,[1] and Mark D. Wiederhold[1]

Abstract

Previous research suggests that the effectiveness of virtual reality exposure therapy should increase as the experience becomes more immersive. However, the neural mechanisms underlying the experience of immersion are not yet well understood. To address this question, neural activity during exposure to two virtual worlds was measured by functional magnetic resonance imaging (fMRI). Two levels of immersion were used: unimodal (video only) and multimodal (video plus audio). The results indicated increased activity in both auditory and visual sensory cortices during multimodal presentation. Additionally, multimodal presentation elicited increased activity in the hippocampus, a region well known to be involved in learning and memory. The implications of this finding for exposure therapy are discussed.

Introduction

Virtual reality (VR), in which a participant is exposed through one or more senses to a simulated environment, is an increasingly popular clinical tool. One application of VR techniques that has shown particular promise is exposure therapy for phobias and posttraumatic stress disorder (PTSD). While meta-analysis has shown that real-world exposure to the object of phobia (known as *in vivo* exposure therapy) is most effective,[1] this is often not feasible, particularly in cases of PTSD, where exposure therapy may require the re-creation of traumatic settings such as battlefields or automobile accidents. In these cases, virtual reality exposure therapy (VRET) provides a practical alternative that has been shown to consistently reduce symptoms of anxiety.[2] Furthermore, VRET has been shown to be significantly more effective than imaginal therapy, another common approach in which the phobic object is simply imagined.[3]

Because *in vivo* exposure therapy appears to be more effective than other, less realistic treatment, it stands to reason that VRET will be most effective when the virtual experience is as realistic and immersive as possible. While no study to date has directly related the experience of realism in VRET to treatment outcomes, a significant relationship between the feeling of presence, which is directly related to immersion, and subjective anxiety after VR exposure to phobic stimuli has been reported[4,5] (but see also Slater[6]). Others have found that a sense of presence significantly correlates to experienced anxiety but not to therapeutic outcomes, leading the authors to conclude that immersive experience is a necessary, if not sufficient, condition for therapeutic improvement.[7]

A major difficulty in identifying the factors related to success in exposure therapy is that most studies measure both the feeling of immersion and evoked anxiety by self-report, which can be unreliable. However, it may be possible to find other markers predicting therapeutic outcomes in the literature of a related field. The theoretical assumption underlying exposure therapy is that phobias are fundamentally a problem of learning and memory, in which the normal enhancement of memory by emotion becomes maladaptive. The goal of exposure therapy, therefore, is to reduce the association between the phobic cue and conditioned fear through either extinction training or reconsolidation.

While physiological factors predicting success in exposure therapy have not been consistently identified,[8] the neural correlates of successful memory encoding have been characterized by converging evidence from functional magnetic resonance imaging (fMRI).[9,10] Therefore if the success of exposure therapy depends on effective learning, and a more immersive VR experience contributes to the success of

[1]Virtual Reality Medical Center, San Diego, California.
[2]Applied fMRI Institute, San Diego, California.

Table 1. Immersion versus Fixation

Region		*x*	*y*	*z*	*Z*	*Voxels*	p
Middle temporal gyrus	R	48	−52	10	5.15	12	<0.001
Postcentral gyrus (area 2)	L	−48	−34	46	4.96	47	<0.001
Hippocampus (subiculum)	L	−20	−26	−16	4.67	44	<0.001
Superior temporal gyrus (TE1)	L	−48	−18	0	4.56	19	<0.001
Superior parietal lobule	R	16	−68	66	4.54	31	<0.001
Hippocampus (subiculum)	R	18	−22	−10	4.5	93	<0.001
Cerebellum (area 18)	R	16	−74	−18	4.47	17	<0.001
Supramarginal gyrus	R	52	−38	36	4.44	149	<0.001
Superior parietal lobule	L	−20	−74	52	4.4	28	<0.001
Precentral gyrus (area 6)	L	−56	2	26	4.35	137	<0.001
Superior frontal gyrus	R	24	−6	−66	4.2	25	<0.001
Middle occipital gyrus	L	−16	−88	−6	4.14	152	<0.001
Fusiform gyrus	R	36	−54	−12	4.13	70	<0.001
Middle occipital gyrus	L	−38	−70	2	4.06	15	<0.001
Cerebellum	R	8	−52	−46	4.05	39	<0.001
Inferior frontal gyrus (area 45)	L	−52	20	16	4.05	16	<0.001
OP3	R	48	−10	22	4.04	8	<0.001
Fusiform gyrus	L	−36	−62	−14	4.03	18	<0.001
TE 1.1	L	−30	−30	24	4	26	<0.001
Superior temporal gyrus	R	58	−22	2	3.95	16	<0.001
Postcentral gyrus (area 3b)	R	66	0	18	3.87	20	<0.001
Precentral gyrus	R	60	−8	−47	3.85	20	<0.001
Postcentral gyrus	R	−50	−8	40	3.84	5	<0.001
Supplementary motor area (area 6)	R	2	−18	72	3.78	6	<0.001
Middle occipital gyrus	R	48	−78	6	3.76	24	<0.001
Amygdala	L	−34	−2	−24	3.62	10	<0.001
Area 17	R	12	−92	−14	3.46	25	<0.001
Thalamus	R	20	−18	2	3.22	8	<0.001

exposure therapy, as VR becomes more immersive, it should produce a pattern of neural activation increasingly similar to that seen with successful encoding. To test this hypothesis, neural activity during two levels of VR immersion was compared, and these results were related to those seen in imaging studies of learning and memory.

Methods

Participants

Five healthy individuals, all male and between the ages of 20 and 30, were recruited for the study. All participants were right handed and had at least a high school education.

Stimuli

Participants viewed clips showing movement through two virtual worlds designed by the Virtual Reality Medical Center (San Diego, CA). The first depicts movement through an icy environment in which ice blocks are opened, revealing penguins. The second world depicts movement along a beach in which the participant collects seashells. In both of these worlds, an auditory signal can be associated with locating a penguin or shell. Conditions with and without the addition of auditory cues were investigated in this study.

Experimental design

Each scan session began with a high resolution T1-weighted anatomical scan upon which functional activations would be overlaid. This was followed by a scanning session lasting 17.6 minutes. During functional scanning, each participant viewed a total of 48 film clips, each 16 seconds long. A 6-second rest period during which a fixation cross was displayed on the screen was presented after each film clip. The film clips were divided into four categories: clips of the icy world with and without sound and clips of the beach world with and without sound. Twelve 16-second clips were displayed in each of these categories. While viewing the clips, participants were instructed to count the penguins or seashells discovered in each trip. This task insured attention and cognitive engagement.

Imaging

Scanning was performed using a 3T Siemens scanner. During anatomical scanning, 160 T1-weighted slices covering the whole brain were acquired in descending order, with a repetition time of 8 milliseconds. Field of vision was $240\times240\times160$, with a voxel size of $1\,mm^3$.

During functional scanning, 33 T2*-weighted slices covering the whole brain were acquired every 2 seconds. Slice thickness was 3 mm, and functional resolution was $3\,mm^3$. Over the 1056-second scanning session, 528 time points were acquired for each participant.

Data analysis

fMRI data was preprocessed in statistical parametric mapping software, SPM5,[11] running in MATLAB 2007b (MathWorks, Natick, MA). Preprocessing began with spatial realignment of functional and anatomical images to correct

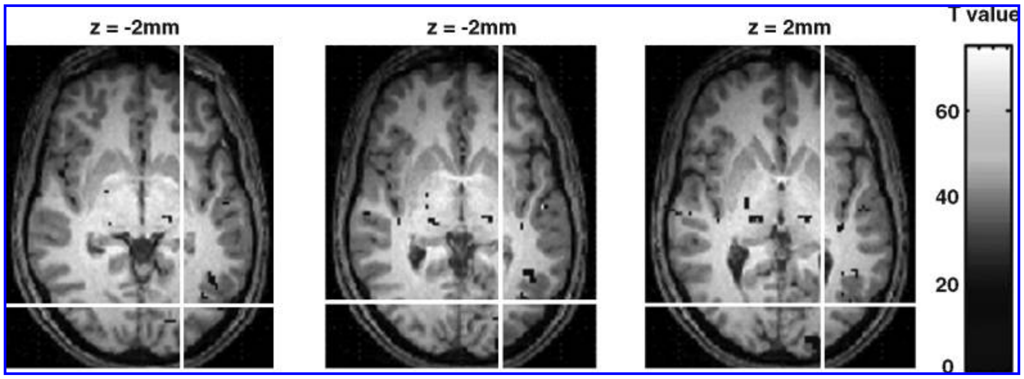

FIG. 1. Regions activated during virtual world viewing compared to fixation. Clusters can be seen in visual processing areas, including the occipital cortex, fusiform gyri, and thalamus.

for head motion. All participants included in this study showed no head motion greater than 2 mm. Anatomical and functional images were reoriented to be aligned with the AC-PC line and were then coregistered. After reorientation, functional images were normalized to the minimum number of individuals (MNI) functional template, to place all images in a common anatomical space. Normalized voxel size was 3 mm^3. Lastly, functional images were smoothed using an 8-mm full width at half maximum (FWHM) Gaussian smoothing kernel.

Functional data was analyzed as a block design, and each epoch of trials was modeled using a boxcar function. Comparisons were performed using *t* contrasts in SPM.

Results from the individual participant level were submitted to a second-level analysis in which participants were treated as a random effect. Second-level analysis was performed by *t* tests using the exploratory threshold of $p < 0.001$, uncorrected, with a minimum cluster size of 5 voxels.

Results

Maxima of clusters were found to show significantly greater BOLD response during immersion in either of the two virtual worlds presented, as compared to fixation (Table 1). The results indicate increased activation throughout the cortex, with large clusters observed in the occipital cortex and fusiform gyri, regions associated with visual processing (see Fig. 1).

Maxima of clusters show whose response significantly differs between immersion in world 1 (icy) or world 2 (beach). The largest cluster for this contrast was found in the left fusiform gyrus (Table 2).

Table 3 shows the results of an exclusive mask showing regions of contrast between immersion and fixation found when audio is presented but not when either was presented without audio. Both contrasts are thresholded at 0.001. Clusters were observed throughout the occipital and temporal cortices. Of particular interest is a bilateral cluster of

Table 2. World 1 versus World 2

Region		*x*	*y*	*z*	*Z*	*Voxels*	p
1–2:							
Inferior parietal lobule	R	58	−52	46	4.45	13	<0.001
Paracentral lobule	L	−8	−40	74	4.05	10	<0.001
Inferior parietal lobule	R	−44	−44	56	4.03	15	<0.001
Angular gyrus	R	28	−60	40	3.95	6	<0.001
Superior parietal lobule	L	−30	−66	50	3.81	10	<0.001
Postcentral gyrus	L	−60	−18	28	3.79	13	<0.001
Inferior occipital gyrus	R	28	−94	−12	3.75	9	<0.001
Middle cingulate cortex	L	−4	−6	32	3.75	9	<0.001
Lingual gyrus	R	20	−100	−12	3.59	5	<0.001
Precuneus	L	−10	−48	70	3.41	12	<0.001
Postcentral gyrus	R	36	−32	54	3.33	5	<0.001
2–1:							
Fusiform gyrus	L	−22	−42	−14	4.04	24	<0.001

Table 3. Immersion-Fix (with Audio) Masked Exclusively with Immersion-Fix (without Audio)

Region		*x*	*y*	*z*	*Z*	*Voxels*	p
Superior occipital gyrus	L	−28	−94	24	4.83	515	<0.001
Lingual gyrus (area 18)	R	22	−90	−4	4.54	174	<0.001
Superior parietal lobule	L	−22	−66	64	4.22	20	<0.001
Inferior temporal gyrus	R	42	−58	−6	4.05	19	<0.001
Angular gyrus	R	40	−64	−48	4	9	<0.001
Inferior occipital gyrus	R	42	−82	−10	3.99	7	<0.001
Superior temporal gyrus	R	50	−12	−6	3.94	16	<0.001
Area 18	R	−30	−86	−20	3.69	5	<0.001
Superior parietal lobule	R	22	−76	52	3.68	17	<0.001
Hippocampus	R	14	−28	−16	3.61	19	<0.001
Middle temporal gyrus	R	68	−16	−8	3.6	6	<0.001
Precuneus	R	6	−62	68	3.58	5	<0.001
Hippocampus	L	−26	−30	−10	3.55	9	<0.001
Supramarginal gyrus	R	40	−32	38	3.55	14	<0.001
Postcentral gyrus	L	−50	−24	58	3.49	9	<0.001
Middle occipital gyrus	R	48	−80	4	3.43	10	<0.001

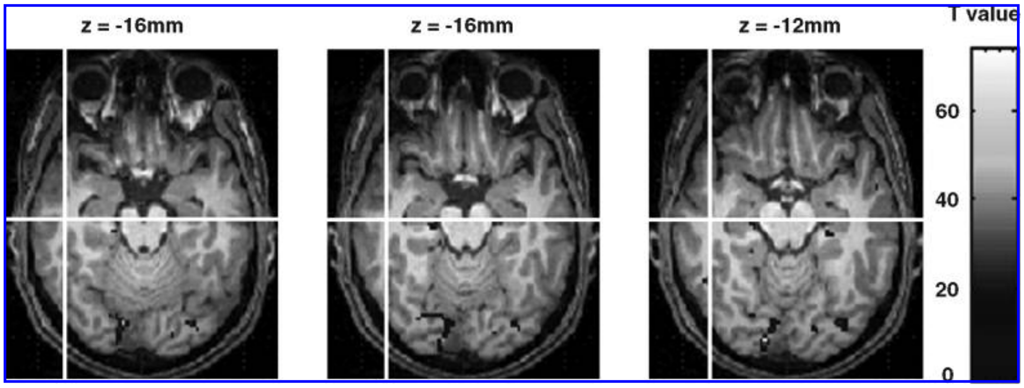

FIG. 2. Cluster of selective activation to presentation with audio as compared to without. Activations in visual cortex and hippocampus are visible.

selective activation in the medial temporal lobe corresponding to the hippocampi (see Fig. 2).

Discussion

The results of these analyses indicate that there is substantial activation in response to immersion in virtual worlds as compared to rest. In particular, increased activation is seen in primary visual and auditory cortices as well as in other regions known to respond strongly to visual stimulation, such as the fusiform cortex and amygdala. These findings conform to the results of previous studies of visual and auditory perception.[12]

The results also show significant differences in the neural response to the two worlds used. Overall, substantially more activation was found in the icy world than in the beach world. Regions of selective activation seem to be primarily related to the body's somatosensory map, including multiple regions of the parietal lobe and precuneus. There are also areas of distinct activation in the primary visual cortex.

Most interesting are the results of the masked contrast experiments. These findings show that as an experience becomes more immersive (in this case through the use of audio related to events in the video), several brain regions not engaged in the less immersive condition are activated. Unsurprisingly, these regions include the primary auditory cortex (located in the superior temporal gyrus). However, several regions not involved in audition are also activated, including substantially increased activity in primary visual cortex and in at least one higher visual association area (the inferior temporal cortex, part of the ventral visual stream). Parietal somatosensory areas, including the precuneus, are also selectively engaged in this condition, as is the angular gyrus, which is believed to be involved in internal verbalization.[13] Perhaps most intriguing is that bilateral clusters in the hippocampus are increasingly engaged when the more immersive audio condition is used. The hippocampus is a region widely known to be critical in learning and memory and has no known sensory function.[14] This indicates that increased immersion in a virtual environment increasingly engages higher cognitive processes, particularly those related to memory.

Increased hippocampal activity during encoding has been consistently related to better subsequent memory.[9] Thus, these data confirm our hypothesis, that brain regions associated with memory encoding would show increased activation in the more immersive multimodal condition.

Because this was a preliminary study, several caveats should be considered in interpreting these results. First, the sample size is relatively small. Second, both the unimodal and multimodal conditions should be seen as representing relatively low levels of immersion, because participants were only passively observing virtual navigation rather than controlling it. Finally, the participants in this study were not phobic or diagnosed with posttraumatic stress, and the stimuli they observed were substantially less arousing than those that would be used in an exposure therapy session.

Nonetheless, even with a modest sample viewing relatively neutral stimuli, significant differences in several brain regions associated with cognition were observed between the unimodal and multimodal conditions. Further studies are needed to assess the neural effects of immersion in a broader patient population, ideally relating these effects to treatment outcomes.

Disclosure Statement

No competing financial interests exist.

References

1. Wolitsky-Taylor KB, Horowitz JD, Power MB, Telch MJ. Psychological approaches in the treatment of specific phobias: a meta-analysis. Clinical Psychology Review 2008; 28:1021–37.
2. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. Journal of Behavioral Therapy & Experimental Psychiatry 2008; 39:250–61.
3. Wiederhold BK, Jang DP, Gervitz RG, et al. The treatment of fear of flying: a controlled study of imaginal and virtual reality graded exposure therapy. IEEE Transactions on Information Technology in Biomedicine 2002; 6:218–23.
4. Regenbrecht HT, Schubert TW, Friedmann F. Measuring the sense of presence and its relations to fear of heights in virtual environments. International Journal of Human-Computer Interactions 1998; 10:233–49.
5. Robillard G, Bouchard S, Fournier T, Renaud P. Anxiety and presence during VR immersion: A comparative study of the

reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. CyberPsychology & Behavior 2003; 6:467–76.
6. Slater M. Presence and emotions. CyberPsychology & Behavior 2004; 7:121.
7. Price M, Anderson P. The role of presence in virtual reality exposure therapy. Journal of Anxiety Disorders 2007; 21:742–51.
8. De Carvalho MR, Freire RC, Nardi, AE. Virtual reality as a mechanism for exposure therapy. World Journal of Biological Psychiatry 2008; 31:1–11.
9. Wais PE. fMRI signals associated with memory strength in the medial temporal lobes: a meta-analysis. Neuropsychologia 2008; 46:3185–96.
10. Kirwan CB, Wixted JT, Squire, LR. Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. Journal of Neuroscience 2008; 28:10541–8.
11. Friston KJ, Holmes AP, Worsley KJ, et al. Statistical parametric maps in functional imaging: a general linear approach. Human Brain Mapping 1995; 2:189–210.
12. Barton JJ. Higher cortical visual function. Current Opinion in Ophthalmology 1998; 9:40–5.
13. Platel H. Functional neuroimaging of semantic and episodic musical memory. Annals of the NY Academy of Sciences 2005; 1060:136–47.
14. Eichenbaum H, Otto T, Cohen NJ. The hippocampus: What does it do? Behavioral & Neural Biology 1992; 57:2–36.

Address reprint requests to:
Dr. Brenda K. Wiederhold
The Virtual Reality Medical Center
6160 Cornerstone Court East, Ste 155
San Diego, California 92121

E-mail: bwiederhold@vrphobia.com

This article has been cited by:

1. References 261-314. [CrossRef]
2. Dr. Michele D. Estes, Mr. Randell SnowAdult Learning and Virtual Worlds Simulations 1270-1279. [CrossRef]