ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 24 (2005) 928–935

Distinct neural substrates for the perception of real and virtual visual worlds

Shihui Han,[a,b,*] Yi Jiang,[a,1] Glyn W. Humphreys,[c] Tiangang Zhou,[d] and Peng Cai[d]

[a] *Department of Psychology, Peking University, Beijing, PR China*
[b] *Learning and Cognition Laboratory, Capital Normal University, Beijing, PR China*
[c] *Behavioral Brain Sciences, School of Psychology, University of Birmingham, Birmingham, UK*
[d] *Key Laboratory of Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, PR China*

Received 4 August 2004; revised 23 September 2004; accepted 29 September 2004
Available online 18 November 2004

Virtual environments have been frequently used for training and skill improvement. However, do real and virtual worlds engage the same brain states in human perceivers? We measured brain activity using functional magnetic resonance imaging (fMRI) while adults watched movie and cartoon clips, simulating real and virtual visual worlds, respectively. Relative to baselines using random static images, the medial prefrontal cortex (MPFC) and the cerebellum were activated only by movie clips of other humans. In contrast, cartoon clips of human and non-human agents activated the superior parietal lobes, while movie clips of animals also activated the superior parietal lobes. Our fMRI findings suggest that the perception of real-world humans is characterised by the involvement of MPFC and the cerebellum, most likely for on-line representation of the mental states of others, whereas the perception of virtual-world agents engages the parietal cortex in attention to actions.

Keywords: fMRI; Neural substrates; Real world; Virtual visual worlds

Introduction

Virtual reality is increasingly used for training in a wide range of contexts. For example, virtual human agents simulated using cartoons have been used to help students learn to perform physical, procedural tasks (Rickel and Johnson, 1999). Animated agents in virtual environments have also been used for training skills that require a high level of flexible, interpersonal interactions such as psychotherapy (Beutler and Harwood, 2004). However, whether the human brain differentially perceives and interacts with agents in the real and virtual worlds has been poorly understood. Recent functional magnetic resonance imaging (fMRI) studies have shown that, when we deal with actions assumed to come from real human agents, specific brain regions, such as the medial prefrontal cortex (MPFC), show stronger activation compared with when we assume the actions come from animated agents simulated by computers (Gallagher et al., 2002; Ramnani and Miall, 2004), suggesting that specific neural substrates may be involved in discrimination between human and non-human agents.

The current study assessed whether, when we simply perceive human agents in the real world, different brain regions are engaged compared with when we perceive agents in virtual worlds. To investigate this, we used fMRI to measure brain activations when participants observed movie and cartoon clips, which presented brief sequences of actions involving humans in real-life situations (movie clips) or actions involving either human or non-human agents in virtual worlds (cartoon clips). Movies present real images (photographs of a physical environment) whereas cartoons present virtual images (a simulation on physical principles of that environment). Brain activity when watching the clips was compared with random order static images from the movie and cartoon clips to control for any differences in low level visual feature processing. Relative to the baseline with static random images, movie or cartoon clips presented continuous and coherent visual events that induced explanatory predictions of behaviour. We aimed to identify if there are neural substrates differentiating the perception of human agents in the real visual world (in movie clips) from the perception of human or non-human characters in a virtual world (in cartoons).

Material and methods

Subjects

Twelve adults (6 male; 21–41 years of age, mean 25.5) with no neurological or psychiatric history participated in this study. All

---

* Corresponding author. Department of Psychology, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China. Fax: +86 10 6276 1081.
E-mail address: shan@pku.edu.cn (S. Han).

[1] Present address: Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455.

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.09.046

participants were right-handed, had normal or corrected-to-normal vision, and were not colour blind. Informed consent was obtained from all participants prior to scanning. This study was approved by the Academic Committee of Department of Psychology, Peking University.

Stimuli and procedure

The stimuli were presented through a LCD projector onto a rear-projection screen located at a subject's head. The screen was viewed with an angled mirror positioned on the head-coil. The stimuli were movie and cartoon clips without accompanying sound and static images extracted from them. As illustrated in Fig. 1, one set of movies (Movie I, Fig. 1a) and one set of cartoons (Cartoon II, Fig. 1c) depicted human beings in real visual scenes. There were two clips of Movie I. One clip showed human activities at a subway station (meeting and walking away). Another clip showed students' activities in a classroom (raising hand and talking). There were two clips of Cartoon II, which also showed human activities at a subway station and students' activities in a classroom, similar to those of clips of Movie I. There were two clips of Cartoon I (Fig. 1b) depicting artificial characters in virtual visual scenes (e.g., robots or machine dinosaurs walking and fighting with transformed shapes) and two clips of Movie II (Fig. 1d) depicting animals in real visual scenes (e.g., gorillas walking and playing in jungles). Both the movies and cartoons were made by showing 29 frames of images per second. At a viewing distance of 70 cm, Movie I, Movie II, Cartoon I, and Cartoon II subtended visual angles of $28 \times 16°$ (width × height), $27 \times 17°$, $27 \times 20°$, and $30 \times 16°$, respectively. The sizes of static images were the same as the corresponding movies and cartoons. Four scans of 420 s were first obtained from each subject. Each scan consisted of six 1-min epochs, alternating pseudo-randomly between movie clips, cartoon clips, and the corresponding static images across subjects. There was a 10-s black screen before each epoch to set up a baseline for each epoch of stimuli. Thirty images were extracted at every 2 s from the corresponding movie and cartoon clips and were presented (each with a duration of 2 s) in a random order during the epochs of static images. Fig. 2 illustrates the order of presentation of different stimulation conditions in one scan. Movie I, Cartoons I and II, and

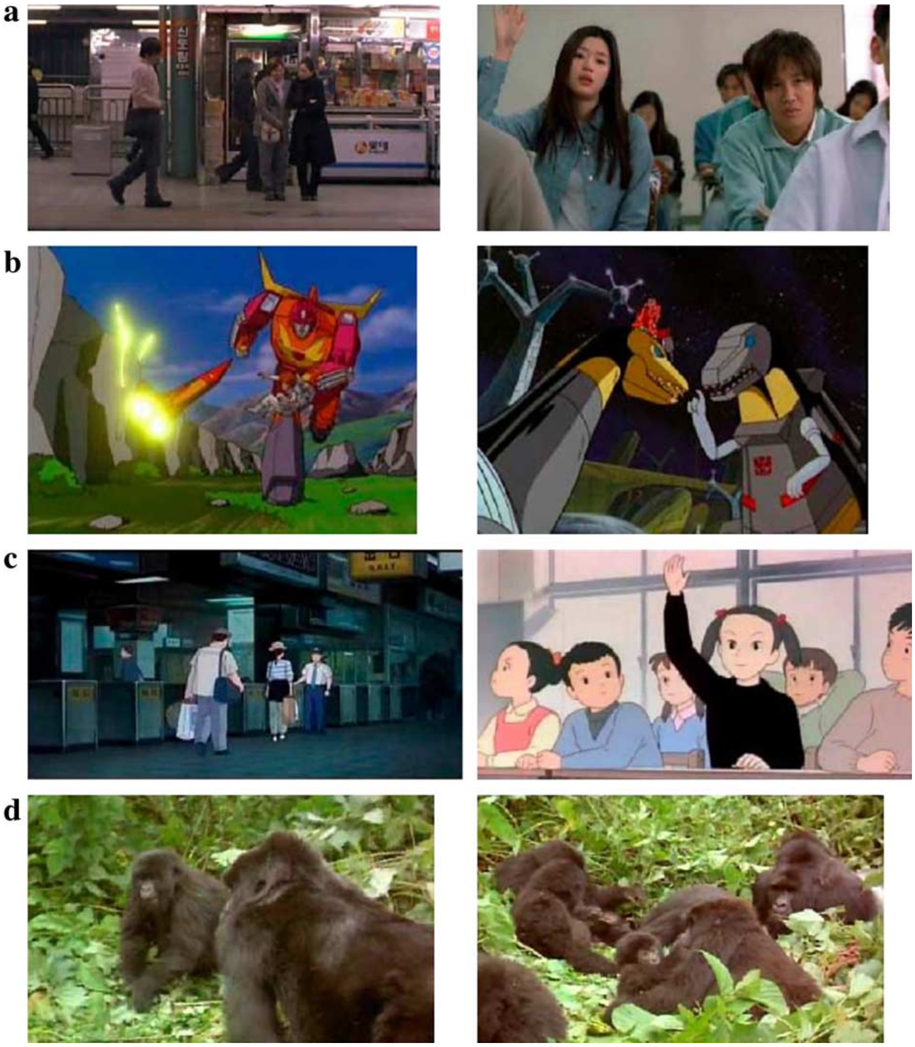

Fig. 1. Illustrations of the stimuli used in the current experiment. (a) Images extracted from Movie I depicting humans in real visual scenes. (b) Images extracted from Cartoon I depicting robots or machine dinosaurs in virtual visual scenes. (c) Images extracted from Cartoon II depicting humans in real visual scenes. (d) Images extracted from Movie II depicting gorillas in real visual scenes.

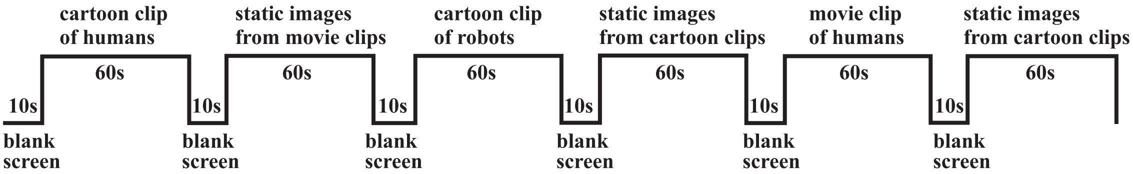

Fig. 2. Ilustrations of the order of presentation of different stimulation conditions in one scan. Movie or cartoon clips alternated with static images.

the corresponding static images were upright in two scans whereas upside down in the other two scans. The sequence of the four scans was counterbalanced across subjects. Two scans of 140 s were obtained from the same group of subjects who was shown with Movie II on a separate day.[2] Each scan consisted of two 1-min epochs of movie and 30 corresponding static images (each lasted for 2 s). Each epoch was preceded by 10-s black screen. Subjects were asked to view freely the movies, cartoons, or static images while keeping their heads still.

fMRI measurement

Scanning was performed on a 3-T Siemens Trio system using a standard head coil at Beijing MRI Center for Brain Research. Thirty-two transversal slices of functional images that covered the whole brain were acquired using a gradient-echo echo-planar pulse sequence ($64 \times 64 \times 32$ matrix with $3.4 \times 3.4 \times 4.4$-mm spatial resolution, TR = 2000 ms, TE = 30 ms, FOV = 220 mm, flip angle = 90°). Anatomical images were obtained using a standard 3D T1-weighted sequence ($256 \times 256 \times 176$ matrix with $0.938 \times 0.938 \times 1.3$-mm spatial resolution, TR = 1600 ms, TE = 3.93 ms). Subjects' heads were immobilised during the scanning sessions using pieces of foam.

fMRI analysis

SPM99 (the Wellcome Department of Cognitive Neurology, UK) was used for data processing and analysis. The functional images were realigned to the first scan to correct for the head movement between scans. The anatomical image was co-registered with the mean functional image produced during the process of realignment. All images were normalised to a $2 \times 2 \times 2$ mm^3 Montreal Neurological Institute (MNI) template in Talairach space (Talairach and Tournoux, 1998) using bilinear interpolation. Functional images were spatially smoothed using a Gaussian filter with a full-width at half maximum (FWHM) parameter set to 8 mm. The image data were modeled using a box-car function. Contrasts were defined to compare the difference between movie or cartoon clips and their corresponding static images presented in a random order. Random effect analyses were then conducted across the group of subjects based on statistical parameter maps from each individual subject to allow population inference. Areas of significant activation were identified at the cluster level for values exceeding a corrected P value of 0.05. The SPM coordinates for standard brain from MNI template were converted to Talairach coordinates (Talairach and Tournoux, 1998) using a nonlinear transform method (http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html). Region-of-interest analyses were conducted to compare relative changes of fMRI signals in the MPFC clusters showing significant activation in the SPM analysis associated with watching movie clips of humans. The mean fMRI signal changes in the MPFC cluster in each condition were compared using one-sample t test.

Results

In Condition A, we recorded fMRI signals from subjects who freely viewed silent movie clips depicting real-life situations, such as human activities at a subway station or in a classroom (Fig. 1a). The contrast of movies–random static images revealed activation in bilateral middle temporal cortex (MT) and the posterior superior temporal sulcus (STS) (centred at −51, −68, 5, $Z = 4.65$, $P < 0.03$, corrected; and 51, −68, 3, $Z = 4.62$, $P < 0.001$, corrected, see Fig. 3a), and the occipital cortex (centered at 12, −87, −2, $Z = 4.82$, $P < 0.001$, corrected). In addition, movies produced activation in the MPFC, corresponding to Brodmann's area 9 (centered at 0, 53, 23, $Z = 3.94$, $P < 0.03$, corrected), and in the left cerebellum (centred at −30, −75, −20, $Z = 4.68$, $P < 0.001$, corrected). Because many cognitive functions show decreased MPFC activity relative to a resting condition (Gusnard and Raichle, 2001), we further examined if movies increased MPFC activity in comparison with watching blank screens using a region-of-interest analysis. Relative to the blank screens, the movies increased fMRI signals in the MPFC ($t = 3.178$, $P < 0.01$) whereas viewing static images decreased fMRI signals in the MPFC ($t = 3.149$, $P < 0.01$, Figs. 3b–d).

To examine the brain areas associated with the perception of artificial characters engaged in coherent events in a virtual visual world, in Condition B, subjects freely viewed either cartoon clips or their corresponding static images. The cartoons displayed robots or machine dinosaurs walking and fighting with transformed shapes (Cartoon I, Fig. 1b). The contrast of cartoons–random static images showed bilateral activation in MT and the posterior STS (centred at −51, −71, 5, $Z = 5.28$, $P < 0.001$, corrected, and 50, −61, 8, $Z = 4.63$, $P < 0.003$, corrected), and in the occipital cortex (centred at −14, −81, 4, $Z = 4.56$, $P < 0.001$, corrected). However, unlike the movies, the cartoons produced bilateral activation in the superior parietal lobules (SPL, centred at −10, −59, 56, $Z = 4.26$, $P < 0.001$, corrected, and 10, −48, 48, $Z = 4.13$, $P < 0.003$, corrected, see Fig. 4) but not in the MPFC or the cerebellum.

To what degree did the engagement of MPFC and the cerebellum, when viewing the real-life scenes, arise from different movement patterns in the movies and the cartoons? This was examined in Condition C by recording fMRI signals associated with viewing inverted presentations of the movie clips and inverted static images used in Condition A. Since visual features

---

[2] The MPFC activation associated with movie clips of humans took us a few weeks to think out the hypothesis that the MPFC activation reflects on-line representation of the mental states of other humans. To test if the MPFC activation is specific for viewing other humans in the real visual world, we designed the movie clips of gorillas and scanned the same group of subjects when they watched the movie clips of animals. This, however, was done 8 weeks after we had scanned the subjects using the movie clips of humans. This is why the scans involving the movie clips of animals always came after the scans that showed the other three conditions.

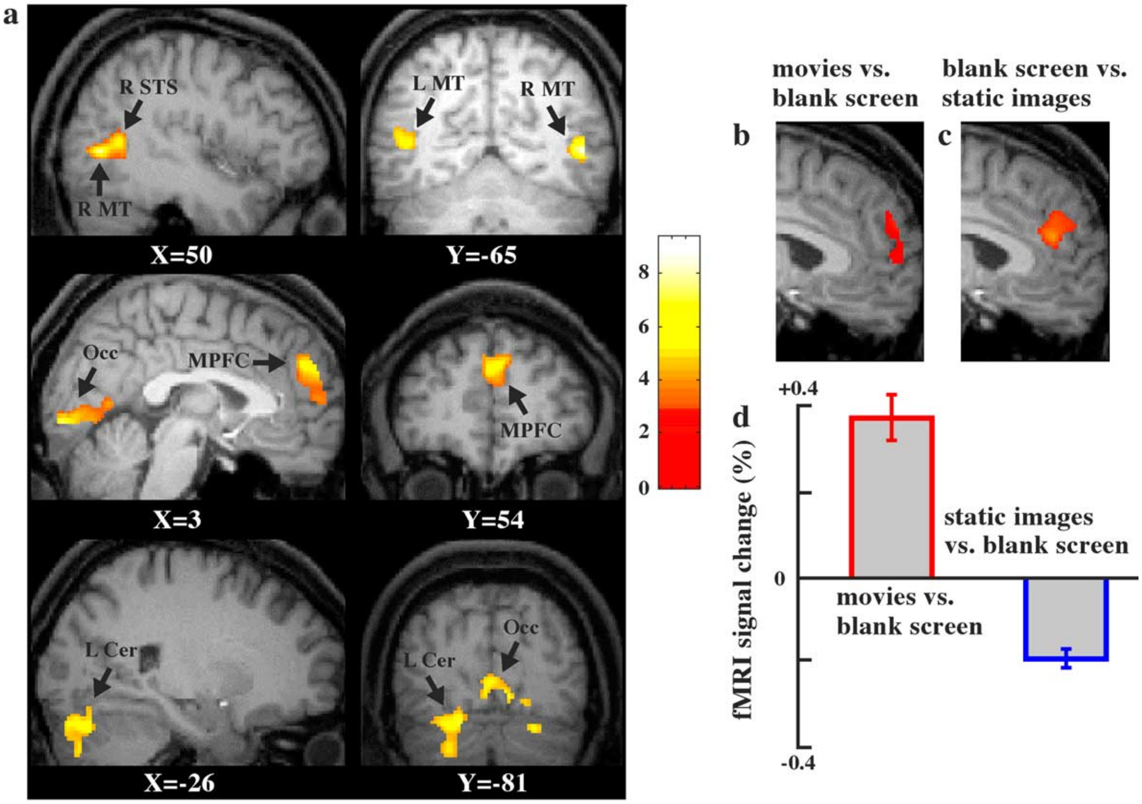

Fig. 3. Brain activations associated with viewing upright movie clips of humans in real visual scenes. (a) The contrast between upright movie clips and random static images. The results of random effect analysis from 12 subjects were plotted on MR images of a representative subject. Coordinates are described with reference to the Tailarach system. The colour bar indicates the scale of *z* values. Activations were observed in bilateral MT and the posterior STS, the medial occipital cortex (Occ), the MPFC, and the left cerebellum (Cer). (b) The MPFC activation in the contrast between upright movie clips and the blank screen (centered at 6, 60, 23, $Z = 3.35$, $P < 0.001$, uncorrected). (c) The MPFC activation in the contrast between the blank screen and random static images from the movie clips (centered at 8, 40, 13, $Z = 5.04$, $P < 0.001$, corrected). (d) Mean fMRI signal changes with standard errors in the MPFC cluster. The fMRI signals were obtained by averaging voxel changes in the MPFC cluster in Fig. 2a in the conditions of watching movie clips, static images, and blank screens. Relative to watching blank screens, watching movie clips increased fMRI signals whereas watching static images decreased fMRI signals in the MPFC clusters.

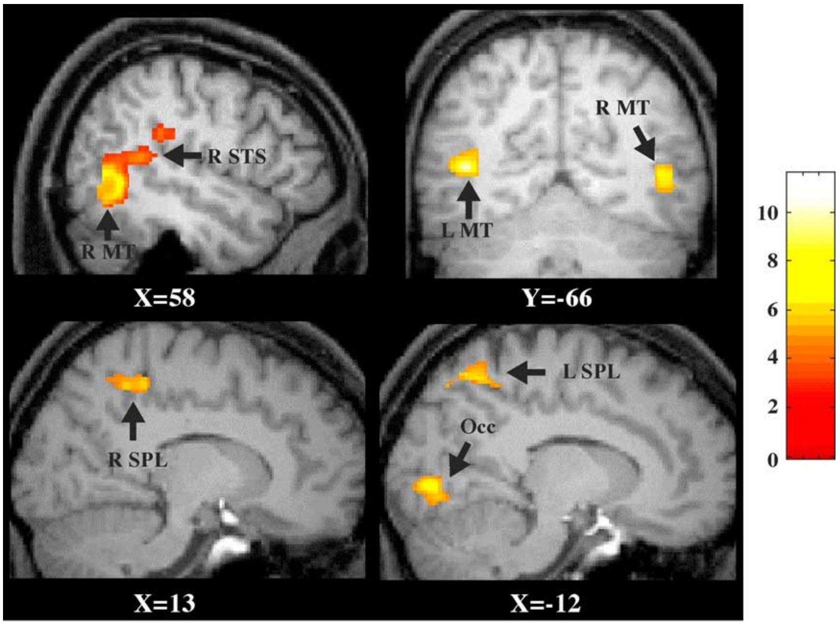

Fig. 4. Illustration of brain activations in the contrast between cartoons of non-human agents and random static images. Activations were observed in bilateral MT and the posterior STS, the medial occipital cortex (Occ), and the left and right superior parietal lobule (SPL).

such as motion, shape, and colour were identical in the upright and inverted stimuli, we would expect effects associated with low-level features to be the same in Conditions C and A. The contrast of inverted movies–random static images showed bilateral activation of MT and the posterior STS (centred at −50, −66, 7, $Z = 4.90$, $P < 0.001$, corrected, and 55, −62, 8, $Z = 4.62$, $P < 0.001$, corrected), the occipital cortex (centred at −22, −93, 8, $Z = 4.11$, $P < 0.001$, corrected), and the left fusiform gyrus (centred at and 39, −60, −13, $Z = 4.57$, $P < 0.02$, corrected, see Fig. 5). Now there was also bilateral activation of the SPL (centred at −10, −55, 60, $Z = 4.11$, $P < 0.01$, corrected, and 28, −61, 49, $Z = 3.92$, $P < 0.01$, corrected), similar to that associated with upright cartoons. However, no activation in the MPFC or the cerebellum was observed in Condition C, indicating that activations observed in these areas in Condition A could not be attributed to specific low-level feature variations in upright movies. The increased activation of the fusiform gyrus for inverted movies may be due to increased attention to the visual features of the scenes or the difficulty of recognising inverted faces, so that extra neural activation is required (though see Kanwisher et al., 1998).

In Condition D, we assessed whether perception of human agents in a virtual visual world activated similar brain regions associated with the perception of human agents in the real visual world. Subjects viewed cartoon clips of human agents in visual scenes similar to those in the movies (Fig. 1c, Cartoon II). Performance was again contrasted relative to static images extracted from the clips. The contrast of human cartoons–random static images showed bilateral activation in MT and the posterior STS (centred at −51, −66, 7, $Z = 4.66$, $P < 0.001$, corrected, and 55, −60, 8, $Z = 5.23$, $P < 0.001$, corrected), the STS (centred at −63, −43, 15, $Z = 4.85$, $P < 0.001$, corrected, and 64, −38, 19, $Z = 5.00$, $P < 0.001$, corrected), and the SPL (centred at −4, −50, 52, $Z = 3.18$, $P < 0.04$, corrected, and 20, −49, 61, $Z = 4.35$, $P < 0.03$, corrected, see Fig. 6), similar to those observed for the cartoon clips of artificial characters. However, there was no activation in the MPFC or the cerebellum. The results indicate that SPL is recruited when coherent actions are viewed in cartoons, regardless of whether the cartoons depict artificial characters or human agents; in contrast, the MPFC and the cerebellum were not automatically engaged in perception of virtual reality. A further region-of-interest analysis confirmed that there was stronger MPFC activation associated with movies clips of human beings relative to cartoon clips of human or non-human agents ($t = 6.00$ and 4.75, respectively, $P < 0.001$), whereas no differential activation in the MPFC was observed between cartoon clips of human and non-human agents.

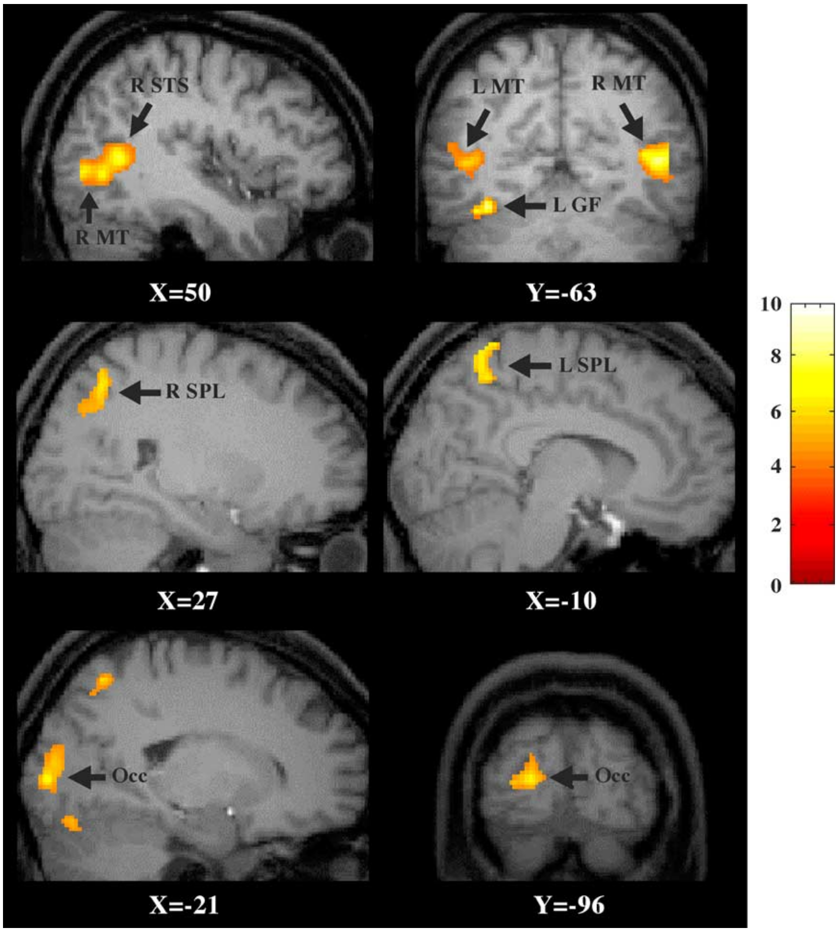

Fig. 5. Brain activations associated with viewing inverted movies of humans. Activations were observed in bilateral MT and the posterior STS, the left occipital cortex (Occ), the left fusiform gyrus (GF), and the left and right superior parietal lobule (SPL).

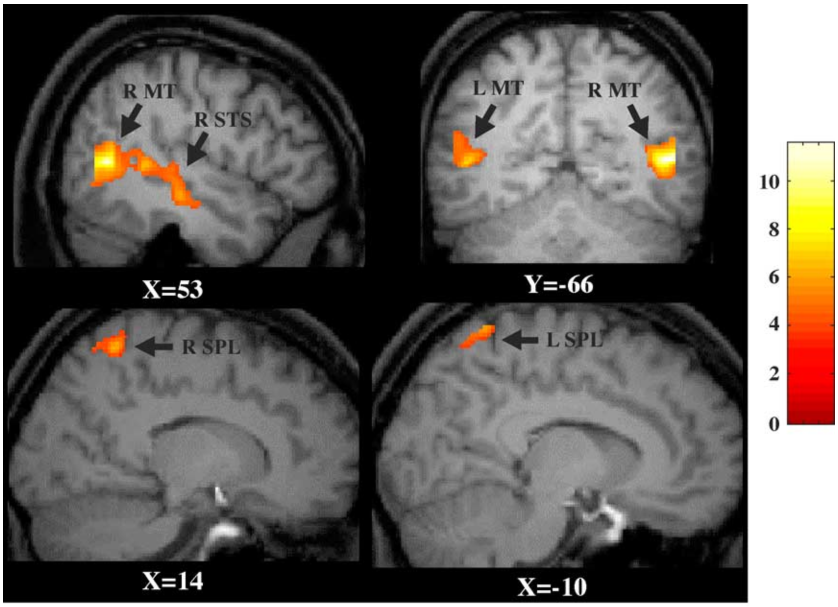

Fig. 6. Illustration of brain activations in the contrast between cartoons of human agents and random static images. Activations were observed in bilateral MT and the posterior STS and the left and right superior parietal lobule (SPL).

To assess if the MPFC and cerebellum activities are specifically involved in the perception of other humans but not animals in the real world, in Condition E, we had the same group of subjects watch upright movie clips that depicted gorillas walking and playing in jungles or the corresponding static images presented in a random order (Fig. 1d). The contrast between the movie clips of gorillas and static images showed bilateral activation in MT and the posterior STS (centred at −51, −66, 3, $Z = 5.70$, $P < 0.001$, corrected, and 51, −73, 2, $Z = 5.39$, $P < 0.001$, corrected), the medial occipital cortex (centred at −2, −91, −2, $Z = 5.74$, $P < 0.001$, corrected), and the SPL (centred at −14, −53, 61, $Z = 4.31$, $P < 0.001$, corrected, and 6, −44, 57, $Z = 4.39$, $P < 0.03$, corrected, see Fig. 7). However, there was no differential activation in the MPFC or the cerebellum.

Discussion

Our functional neuroimaging findings provide important clues about the way we perceive characters within coherent successive events in real and virtual worlds. A number of common areas were activated by all the conditions with movie and cartoon clips, relative

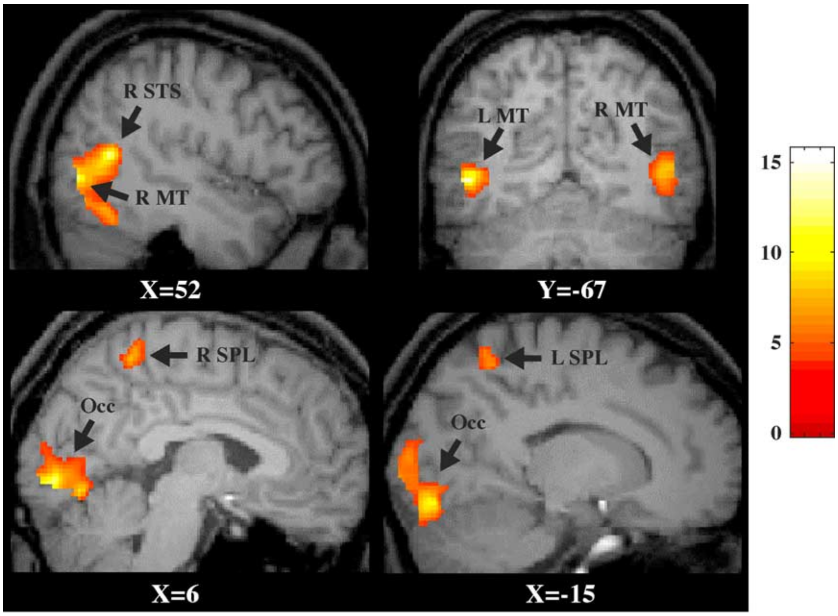

Fig. 7. Illustration of brain activations in the contrast between movies of animals in real visual scenes and random static images. Activations were observed in bilateral MT and the posterior STS, the medial occipital cortex (Occ), and the left and right superior parietal lobule (SPL).

to their static image baselines. The medial occipital cortex and MT are likely engaged by the processing of low-level visual features of the moving images, such as changes in the shape, colour (Livingstone and Hubel, 1998), and motion direction (Tootell et al., 1995) of the stimuli, and the posterior STS is involved in the processing of biological motion (Puce et al., 1998) even when inverted images were perceived. The fMRI results suggest that the processing of low-level visual features of the real and virtual visual worlds is mediated by similar neural mechanisms in the posterior brain areas.

However, distinct neural substrates at higher level brain structures are involved when we view agents in coherent events in the real and virtual visual worlds. Most important, there were distinct brain regions activated by movie clips of human beings, on the one hand, and movie clips of animals and cartoon clips of human and non-human agents, on the other. The MPFC and the left cerebellum showed enhanced activity when adults watched movies depicting human activities in everyday-life situations. Our preliminary separate analysis of the fMRI data associated with the movie clips showed MPFC activation for both clips regardless of the difference in contents between them. Nevertheless, none of the clips of Cartoon II showed MPFC activation even though Cartoon II showed similar scenes as those of Movie I. Thus, it appears that the way to present the visual scenes rather than the content is critical for the MPFC and cerebellum activations. The MPFC and cerebellum activations could not result from summed amounts of motions in the movies because motion patterns were the same in the upright and inverted movies whereas the MPFC and cerebellum showed enhanced activities only to the upright movies. The movie clips of gorillas were also different from cartoon clips in motion patterns but did not induce the MPFC and cerebellum activations. It is unlikely that the effects are simply due to the perception of social relations between the characters (Iacoboni et al., 2004) or to the presence of coherent actions in the movie condition but not in its baseline (with random, static images). There were social relations between the individuals and coherent actions in the cartoons of people, similar to those in the movie clips, but the MPFC did not show enhanced activity. The effects cannot be attributed to simulation of human agents' actions with a third person perspective, which activated different brain regions (Ruby and Decety, 2001). It is also unlikely that the effects were caused by the stimuli being differentially interesting. We asked the participants to rate the movie and cartoon clips for interest, and found no differences in the rated interest for the critical conditions: movies of humans, cartoons of non-humans, cartoons of human characters, and movies of animals [the mean ratings were 3, 3.25, 2.88, 4, respectively, on a 5-point scale where 5 = very interesting, $F(1,11) = 0.25$, $P > 0.5$]. The cerebellum activity could be argued to arise from enhanced working memory when viewing a familiar scene (Desmond and Fiez, 1998), but there is no reason to expect this only for the movies of people compared with the matched cartoons of humans.

Given that the MPFC and cerebellum activations were eliminated when viewing animals in real visual scenes, it may be proposed that these areas are possibly engaged in 'mentalising' about other humans or forming a theory of mind of other humans (Frith and Frith, 1999; Premack and Woodruff, 1978; Wellman, 1990), consistent with previous neuroimaging findings (Brunet et al., 2000; Fletcher et al., 1995; Gallagher et al., 2000), and brain lesion results (Stone et al., 1998; Stuss et al., 2001). It is possible that aspects of mental state reasoning are engaged only when adults view other people in real-life situations (in movies), whereas this might not occur when viewing both human and non-human agents in virtual scenes (in cartoons). The MPFC and cerebellar regions were also not differentially activated when movies of people were inverted, presumably because the difficulty of encoding the people and objects in such scenes (Farah et al., 1995; Rock, 1986) impaired mental state reasoning. Prior studies have shown that the MPFC and associated regions can be activated when participants are required to carry out mental state reasoning on static cartoon pictures (Gallagher et al., 2000) and even moving shapes (Casteli et al., 2000), suggesting that 'mentalising' associated with the MPFC can be conducted on agents in cartoons (and with non-human agents) when explicitly required. However, the contrast in our data between activation states when viewing humans in everyday behaviours and when viewing the other stimuli indicates that the MPFC is not spontaneously engaged in mental state reasoning unless we observe other humans in the real visual world.

Watching characters in cartoons and animals in movies generated distinct patterns of activation bilaterally in the SPLs. Previous neuroimaging studies have shown the SPLs are engaged in perception of coherent motion (Kononen et al., 2003), motion tracking (Jovicich et al., 2001), and attentional shifts in visual space (Hopfinger et al., 2000; Nobre et al., 2003).However, coherent motion was the same in the upright and inverted movies and existed in both cartoon and movie clips. There is no reason to assume that, relative to upright movies of humans, inverted movies of humans induced enhanced motion tracking. The reverse might be true since faces and objects were difficult to recognize in the inverted movies. In addition, because the contents of the upright movies of humans were matched with those of the cartoons of humans, it is unlikely that cartoons induced more frequent attentional shifts in space. Prior work has indicated that regions of the parietal cortex including the SPL are activated when participants view actions for later memory (Grezes et al., 1999) or pay attention to their own finger movement (Rowe et al., 2002), suggesting that the SPL is activated particularly when attention is engaged on actions. It follows that the SPL activation found when viewing cartoons, movies of animals, and even inverted movies of people is due to participants engaging their attention on the actions taking place in these scenes. Importantly, attention to actions alone was not sufficient to generate differential activation of the MPFC circuit that we link with representing and empathising with the mental states of others.

It is also important to note that, when participants viewed the movies of other people, there was no evidence for increased activation of the SPL relative to the static image baseline. This was despite the fact that salient events still took place in the movies with humans. This null result indicates that the presence of action per se may not be vital, but rather activations may vary with how much attention is engaged on different aspects of the scene. If there is attention to, and empathising with, the mental states of individuals, then there may be correspondingly reduced attention to the physical actions taking place.

Prior neuroimaging (Saxe and Kanwisher, 2003) and brain lesion studies (Samson et al., 2004) have shown evidence that the temporo-parietal junction (TPJ) is also involved in mental states reasoning. The current study, however, did not find activation in the TPJ in any stimulus condition. Viewing cartoons of human agents induced activations in the STS, which are, however, anterior to the TPJ area reported in the Saxe and Kanwisher (2003) study. Saxe and Kanwisher contrasted brain states when reading stories about a character's mental states and when reading stories about other people in physical detail. The current study, however, contrasted brain states when viewing people in coherent visual events and when viewing static images with social cues. It is possible that mental state reasoning from text stories involved the posterior part of the brain whereas mental state reasoning when viewing other people's behaviour in the real visual world depends mainly upon the involvement of the MPFC. This, however, needs to be clarified in the future work.

The present study extends in important ways the prior brain imaging studies contrasting effects of actions assigned to humans or to computers (Gallagher et al., 2002; Ramnani and Miall, 2004). We show engagement of a neural loop including the MPFC and the cerebellum even when we simply view humans in real scenes, but not when active agents are seen in virtual environments. Hence, the perception of real and virtual visual worlds can be mediated with distinct neural substrates. The distinct neural and functional processes engaged in perceiving agents in real and virtual worlds may contribute to the limited efficacy of virtual visual environments for training (Seidel and Chatelier, 1999). Our findings suggest that human brains possibly function, when we interact with real people in everyday life, in a way different from when we view or interact with artificial characters or static social stimuli. It may be further speculated that the brain may not attempt to model the behaviour of cartoon characters as it does with real people, which possibly constrains the social impact of cartoons on adult viewers.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Project 30225026 and 30328016), the Ministry of Science and Technology of China (Project 2002CCA01000), the Ministry of Education of China (02170), the Medical Research Council (UK), and the Royal Society (UK).

References

Beutler, L.E., Harwood, T.M., 2004. Virtual reality in psychotherapy training. J. Clin. Psychol. 60, 317–330.

Brunet, E., Sarfati, Y., Hardy-Baylé, M., Decety, J., 2000. A PET investigation of the attribution of intentions with a nonverbal task. NeuroImage 11, 157–166.

Casteli, F., Happé, F., Frith, U., Frith, C., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. NeuroImage 12, 314–325.

Desmond, J.E., Fiez, J.A., 1998. Neuroimaging studies of the cerebellum: language, learning and memory. Trends Cogn. Sci. 2, 355–362.

Farah, M.J., Tanaka, J.W., Drain, H.M., 1995. What causes the face inversion effect? J. Exp. Psychol. Hum. Percep. Perform. 21, 628–634.

Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.F., Frackowiak, R.S., Frith, C.D., 1995. Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57, 109–128.

Frith, C.D., Frith, U., 1999. Interacting minds—a biological basis. Science 286, 1692–1695.

Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 38, 11–21.

Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. NeuroImage 16, 814–821.

Grezes, J., Costes, N., Decety, J., 1999. The effects of learning and intention on the neural network involved in the perception of meaningless actions. Brain 122, 1875–1887.

Gusnard, D.A., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. Nat. Rev., Neurosci. 2, 685–694.

Hopfinger, J.B., Buonocore, M.H., Mangun, G.R., 2000. The neural mechanisms of top-down attentional control. Nat. Neurosci. 3, 284–291.

Iacoboni, M., Lieberman, M.D., Knowlton, B.J., Molnar-Szakacs, I., Moritz, M., Throop, C.J., Fiske, A.P., 2004. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. NeuroImage 21, 1167–1173.

Jovicich, J., Peters, R.J., Koch, C., Braun, J., Chang, L., Ernst, T., 2001. Brain areas specific for attentional load in a motion-tracking task. J. Cogn. Neurosci. 13, 1048–1058.

Kanwisher, N., Tong, F., Nakayama, K., 1998. The effect of face inversion on the human fusiform face area. Cognition 68, B1–B11.

Kononen, M., Paakkonen, A., Pihlajamaki, M., Partanen, K., Karjalainen, P.A., Soimakallio, S., Aronen, H.J., 2003. Visual processing of coherent rotation in the central visual field: an fMRI study. Perception 32, 1247–1257.

Livingstone, M., Hubel, D., 1998. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240, 740–749.

Nobre, A.C., Coull, J.T., Walsh, V., Frith, C.D., 2003. Brain activations during visual search: contributions of search efficiency versus feature binding. NeuroImage 18, 91–103.

Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? Behav. Brain Sci. 1, 515–526.

Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci. 18, 2188–2199.

Ramnani, N., Miall, R.C., 2004. A system in the human brain for predicting the actions of others. Nat. Neurosci. 7, 85–90.

Rickel, J., Johnson, W.L., 1999. Animated agents for procedural training in virtual reality: perception, cognition, and motor control. Appl. Artif. Intell. 13, 343–382.

Rock, I., 1986. The Description and Analysis of Objects and Event Perception. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.). Handbook of Perception and Human Performance vol. 2. Wiley, New York, pp. 1–71. Chap. 33.

Rowe, J., Friston, K., Frackowiak, R., Passingham, R., 2002. Attention to action: specific modulation of corticocortical interactions in humans. NeuroImage 17, 988–998.

Ruby, P., Decety, J., 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. Nat. Neurosci. 4, 546–550.

Samson, D., Apperly, I.A., Chiavarino, C., Humphreys, G.W., 2004. Left temporoparietal junction is necessary for representing someone else's belief. Nat. Neurosci. 7, 499–500.

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". NeuroImage 19, 1825.

Seidel, R.J., Chatelier, P.R., 1999. Virtual Reality, Training's Future? Perspectives on Virtual Reality and Related Emerging Technologies. Plenum, New York.

Stone, V.E., Baron-Cohen, S., Knight, R.T., 1998. Frontal lobe contributions to theory of mind. J. Cogn. Neurosci. 10, 640–656.

Stuss, D.T., Gallup, G.G., Alexander, J.P., Alexander, M.P., 2001. The frontal lobes are necessary for 'theory of mind'. Brain 124, 279–286.

Talairach, J., Tournoux, P., 1998. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Tootell, R.B., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R., Belliveau, J.W., 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J. Neurosci. 15, 3215–3230.

Wellman, H.M., 1990. The Child's Theory of Mind. MIT Press, Cambridge, MA.