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.
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter
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 $28 \times 16^\circ$ (width $\times$ height), $27 \times 17^\circ$, $27 \times 20^\circ$, and $30 \times 16^\circ$, 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 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]

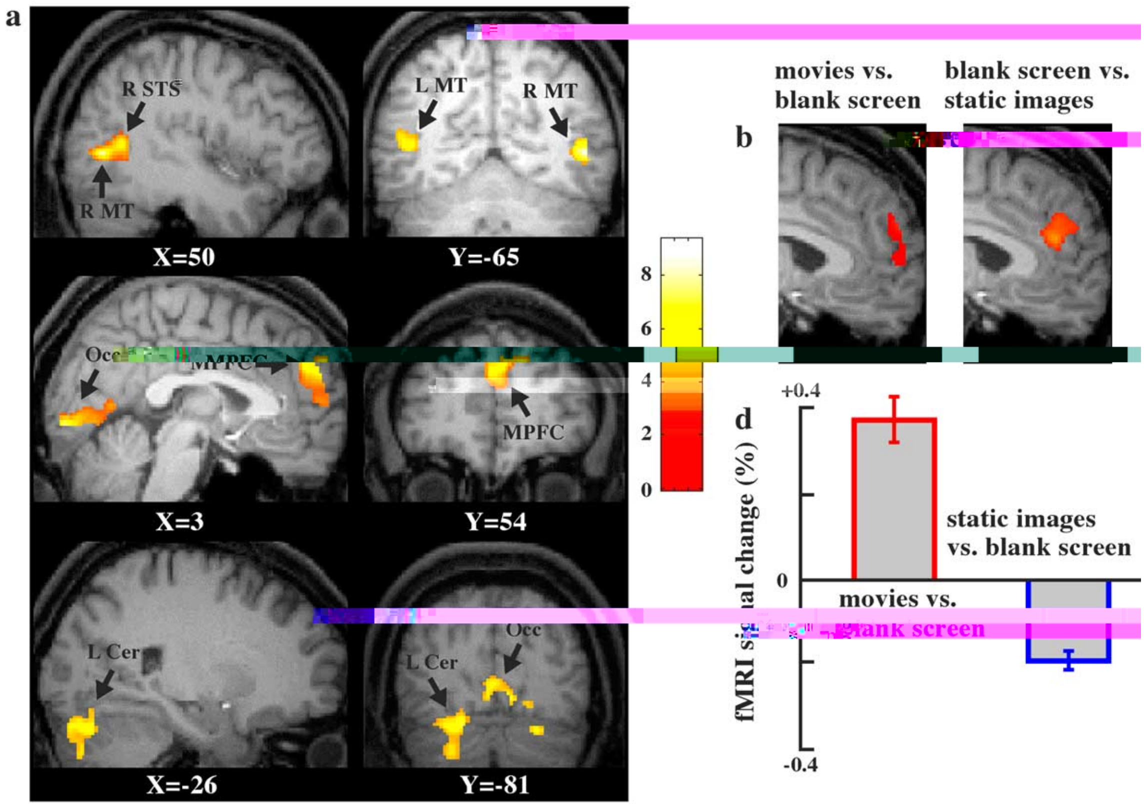

Fig. 3. Brain activations associated with viewing upright movie clips of humans in real visual scenes. (a) The contrast between upright movie clips and random static images. The results of random effect analysis from 12 subjects were plotted on MR images of a representative subject. Coordinates are described 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, 1/F21Tf0.8251D[(mah)-247.4i(centeredActivS86.8ct-257.*P*

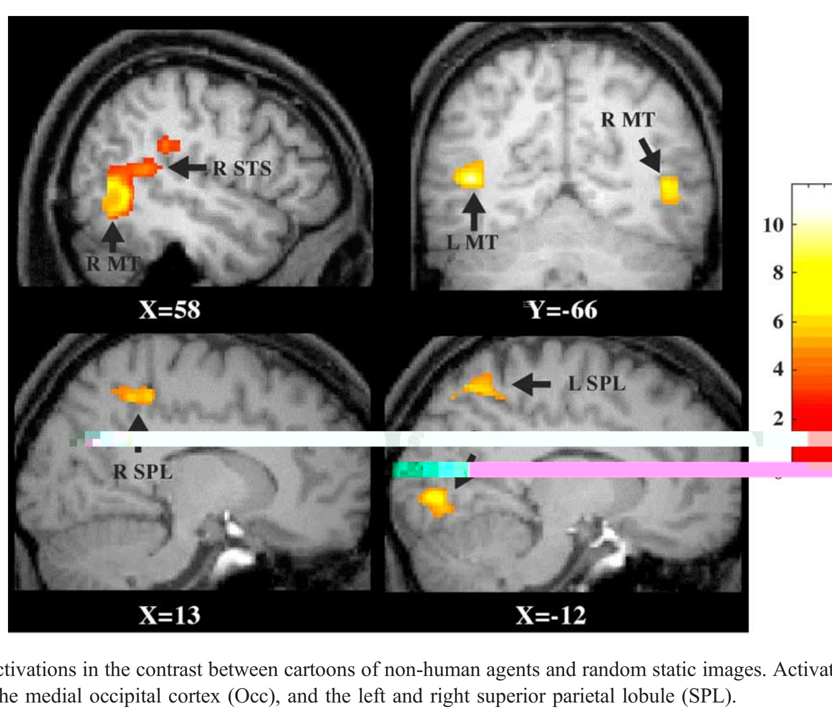

Fig. 4. Illustration of brain activations in the contrast between cartoons of non-human agents and random static images. Activations were observed in bilateral MT and the posterior STS, the medial occipital cortex (Occ), and the left and right superior parietal lobule (SPL).

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.

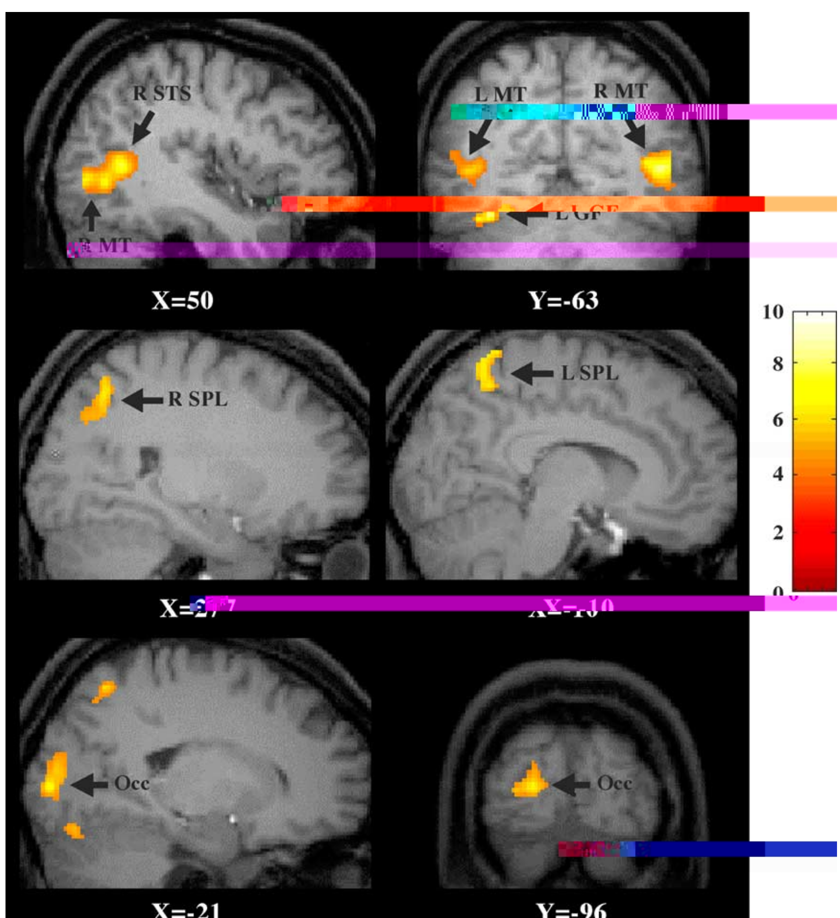

Fig. 5. Brain activations associated with viewing inverted movies of humans. Activations were observed in bilateral MT and the posterior STS, the left occipital cortex (Occ), the left fusiform gyrus (GF), and the left and right superior parietal lobule (SPL).

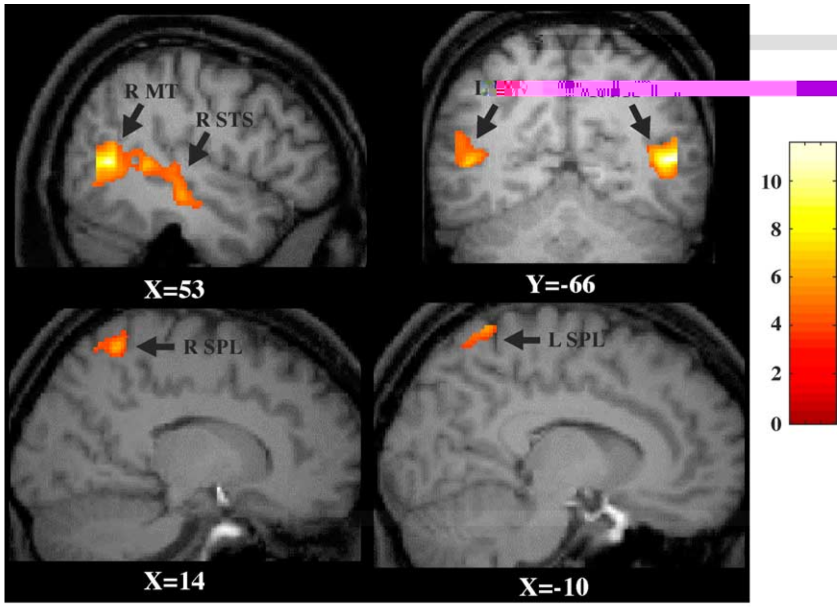

Fig. 6. Illustration of brain activations in the contrast between cartoons of human agents and random static images. Activations were observed in bilateral MT and the posterior STS and the left and right superior parietal lobule (SPL).

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

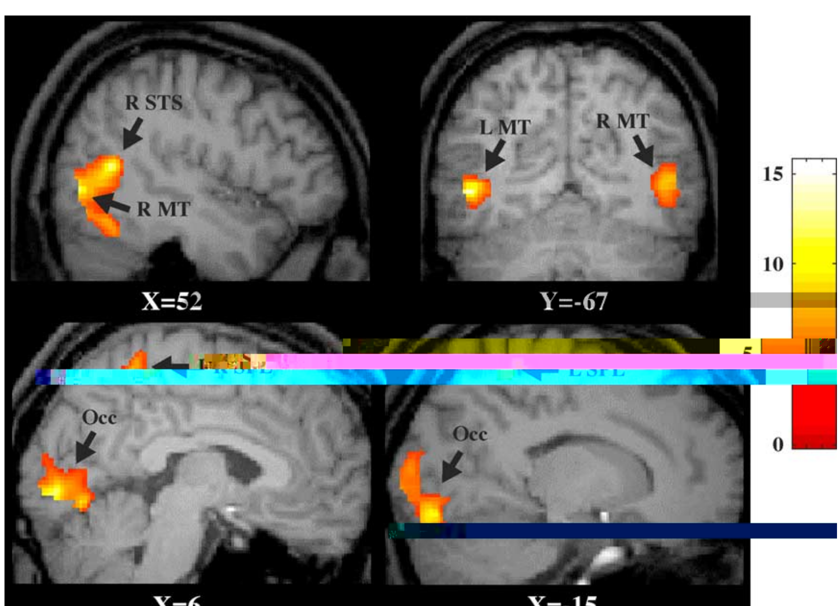

Fig. 7. Illustration of brain activations in the contrast between movies of animals in real visual scenes and random static images. Activations were observed in bilateral MT and the posterior STS, the medial occipital cortex (Occ), and the left and right superior parietal lobule (SPL).

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 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.