Brain & Language 141 (2015) 1–10

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

The theory-of-mind network in support of action verb comprehension: Evidence from an fMRI study

Nan Lin [a,*], Yanchao Bi [b,c], Ying Zhao [d], Chunming Luo [a], Xingshan Li [a]

[a] Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[b] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China
[c] Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China
[d] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

ARTICLE INFO

Article history:
Accepted 9 November 2014
Available online 11 December 2014

Keywords:
Semantic
Theory of mind
Verb
Action
Biological motion
fMRI

ABSTRACT

The theory-of-mind (ToM) network refers to a specific group of brain regions implicated in the thinking of people's mental states. It remains unclear how this network contributes to verb comprehension. In the present study, we compared brain activations evoked by verbs that refer to social actions, private actions, and nonhuman events. All classic regions of the ToM network, including the posterior superior temporal sulcus (pSTS) whose activation during word comprehension is typically interpreted as the processing of motion properties, showed stronger activations to social action verbs than the others. These findings indicate that the ToM network is involved in the processing of social/mental knowledge of verb meanings. Furthermore, the activation of the pSTS during word comprehension mainly reflects the processing of social/mental properties but not that of biological-motion properties.

1. Introduction

Theory of mind (ToM) refers to the ability to attribute independent mental states to self and others in order to predict and explain behavior (Premack & Woodruff, 1978). Neuroimaging studies have found that ToM is selectively associated with a group of brain regions called the "ToM network," which includes the medial prefrontal cortex (MPFC), posterior cingulate (PC)/precuneus, bilateral temporo-parietal junctions (TPJs)/posterior superior temporal sulci (pSTS) and anterior temporal lobes (ATLs) (Gallagher et al., 2000; Mar, 2011; Saxe, 2009; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle & Baetens, 2009). Although the involvement of the network in classic ToM tasks (such as the false belief task) has been reliably observed (Mar, 2011; Schurz et al., 2014), the underlying cognitive processes remain unclear. To clarify the specific cognitive functions of the ToM network, we need to determine the necessary and sufficient conditions that modulate its activation. To this end, the activation of the ToM network in tasks with simple and basic underlying cognitive elements should be investigated.

We here focus on the activation of the ToM network in a basic type of language processing – verb comprehension. Although studies using language comprehension tasks have made substantial contributions in the behavioral and neuroimaging literature on ToM (Saxe, 2006), most of these studies adopted stories or long sentences as their stimuli. Until now only very few neuroimaging studies have investigated the processing of social/mental knowledge during word comprehension (Mitchell, Banaji, & Macrae, 2005; Mitchell, Heatherton, & Macrae, 2002; Zahn et al., 2007) and only a recent study has indicated the involvement of the ToM network in verb comprehension (Spunt, Falk, & Lieberman, 2010). In Spunt et al. (2010), participants were presented with verb phrases (e.g., brush teeth) and were asked in different conditions why or how people typically perform those actions. It was found that several brain regions belonging to the ToM network showed stronger activation in the "why" condition than in the "how" condition, indicating that the ToM network supports the access of "why" knowledge of verb semantics.

In the present study, we will focus on two new questions about how the ToM network contributes to verb comprehension. The first question is, besides the why/how dimension, whether another semantic dimension of verb semantics – to what extent social interactions are indicated (hereinafter referred to as "sociality") modulates the engagement of the ToM network in verb

---

[*] Corresponding author.
E-mail address: linn@psych.ac.cn (N. Lin).

http://dx.doi.org/10.1016/j.bandl.2014.11.004

comprehension. Two previous studies showed that, when participants are presented with comic strips describing human activities with social and private intentions, stronger ToM activations can be observed in the social condition than in the private one (Ciaramidaro et al., 2007; Walter et al., 2004). Based on these findings, it is reasonable to expect that sociality of verb meanings modulates the activation of the ToM network during verb comprehension. However, it should be noted that the access to social knowledge in verb comprehension is different from that in cartoon reading: the social knowledge of verb meanings is a part of our lexical semantic knowledge that can be accessed directly from semantic memory while the understanding of social interactions in the comic strips largely relies on online integrating and reasoning processes. Therefore, if we observe sociality effects in the ToM network during a verb comprehension task, we will infer that any online integrating or reasoning process in order to understand a social interaction is not a necessary condition for the engagement of the ToM network.

The second question is whether the pSTS activation during action verb comprehension reflects the processing of biological-motion knowledge or that of social/mental knowledge. Although word comprehension, especially action verb comprehension, frequently activates one component of the ToM network – the pSTS, such a result has never been interpreted as ToM processing. A prevalent explanation for the activation of the pSTS during word comprehension is that the pSTS stores the biological-motion properties (i.e., the articulated flexible motion properties) of concepts (Beauchamp, Lee, Haxby, & Martin, 2002; Chao, Haxby, & Martin, 1999; Han et al., 2013; Kemmerer, Castillo, Talavage, Patterson, & Wiley, 2008; Lin, Lu, Fang, Han, & Bi, 2011; Martin, 2007). This biological motion hypothesis is derived from an influential idea of semantic representation called the embodied cognition hypothesis, which holds that semantic knowledge about the various physical properties of objects and actions is distributed in or near (e.g., anterior to) cortical areas involved in processing corresponding sensory or motor features (Barsalou, 2008; Mahon & Caramazza, 2008; Martin, 2007). With this basic idea, two aspects of experimental evidence support the biological motion hypothesis. First, the pSTS is involved in the perception of biological motion (Beauchamp et al., 2002; Caspers, Zilles, Laird, & Eickhoff, 2010; Grosbras, Beaton, & Eickhoff, 2012). Second, the processing of concepts that contain biological-motion properties (i.e., animals and actions), in comparison with that of other concepts, can elicit strong activation in the pSTS (Chao et al., 1999; Lin et al., 2011; Martin, 2007). However, biological motion differs from other types of motion not only in its sensory properties but also in its social/mental properties because it is typically intentional and in many cases has social significance. Therefore, the experimental evidence equally supports the alternative hypothesis that the pSTS takes part in the processing of the social/mental properties of word meanings.

To address the two above questions, the present study compared brain activations elicited by three classes of verbs: social action verbs (e.g., embrace, massage, and salute), private action verbs (e.g., walk, eat, and hunt), and nonhuman verbs (e.g., drip, burn, and rot). Our prediction is that if the sociality of verbs modulates the activation of the ToM network during verb comprehension, then social action verbs, whose meanings strongly indicate interactions between people, elicit stronger activation of the ToM network than the other two classes of verbs. And if the pSTS activates in verb comprehension as a part of the ToM network, it should show more preference to social action verbs than to private ones; otherwise if the richness of the biological-motion features is the only factor modulating the activation of the pSTS during verb comprehension, then the pSTS should show equal preference to social and private action verbs in comparison with nonhuman verbs as long as the richness of the biological-motion features is matched between social and private action verbs.

2. Materials and methods

2.1. Participants

Seventeen healthy undergraduate and graduate students (13 females) participated in the experiment. The average age of the participants was 21.3 years (SD: 2.1 years). All participants were right handed and were native speakers of Chinese. No participant suffered from psychiatric or neurological disorders or had ever sustained a head injury. Before the experiment, each participant read and signed an informed consent issued by the Institutional Review Board of the Beijing Normal University (BNU) Imaging Center for Brain Research.

2.2. Design, task, and stimuli

Three verb conditions, namely, social action verbs (e.g., embrace, massage, and salute), private action verbs (e.g., walk, eat, and hunt), and nonhuman verbs (e.g., drip, burn, and rot), were included in the experiment. Each condition included 70 verbs, all of which were two-character, disyllabic Chinese words. The word frequency was obtained from the Language Corpus System of Modern Chinese Studies (Sun, Huang, Sun, Li, & Xing, 1997) and was matched between conditions [mean frequency count per million (SD): social action verbs, 5.7 (6.1); private action verbs, 6.6 (15.3); and nonhuman verbs, 5.5 (5.6); $ts\ (138) < 1$].

We manipulated two factors among the three verb conditions – the sociality and the richness of biological-motion features. The manipulations of both factors were confirmed by independent rating experiments with additional participants who were blind to the study objectives. In the sociality rating experiment, 16 participants (10 females) were asked to classify verbs on a five-point scale according to how many people are involved in an event that a verb refers to: response 5 corresponded to an event that necessarily involves two or more people; response 4 corresponded to an event that typically, but not necessarily, involves two or more people and their interactions; response 3 corresponded to an event that sometimes involves two or more people and their interactions; response 2 corresponded to an event that typically involves only one person; and response 1 corresponded to an event that needs no human agent. The social action verbs scored 3.87 points on average (SD: 0.77), the private action verbs scored 2.45 points on average (SD: 0.36), and the nonhuman verbs scored 1.16 points on average (SD: 0.35). The rating differences between each two of the three conditions were all significant [$ts\ (138) > 13$; $ps < 0.001$] (Fig. 1) and the inter-rater reliability was high ($ICC\ (2,16) = 0.983$; Shrout & Fleiss, 1979). An interesting finding is that a considerable percentage of people comprehend action verbs whose meanings per se do not contain any social interaction as social events. For example, for the verb "run," half of our subjects thought that it refers to an event that sometimes contains social interactions. For the verb "walk," more than half of our subjects thought that it refers to an event that sometimes or even typically contains social interactions. Therefore, studies of action verb comprehension should consider the sociality effect even when no typical social action verb is included in the stimuli. Analogously, our classification of human and nonhuman verbs is relative but not absolute. Given that the goal of the experiment is to examine whether verbs containing richer social semantic features can evoke stronger activation in the ToM network, we included the nonhuman verb condition as a baseline that contains very few but not necessarily zero social semantic features. The significant difference

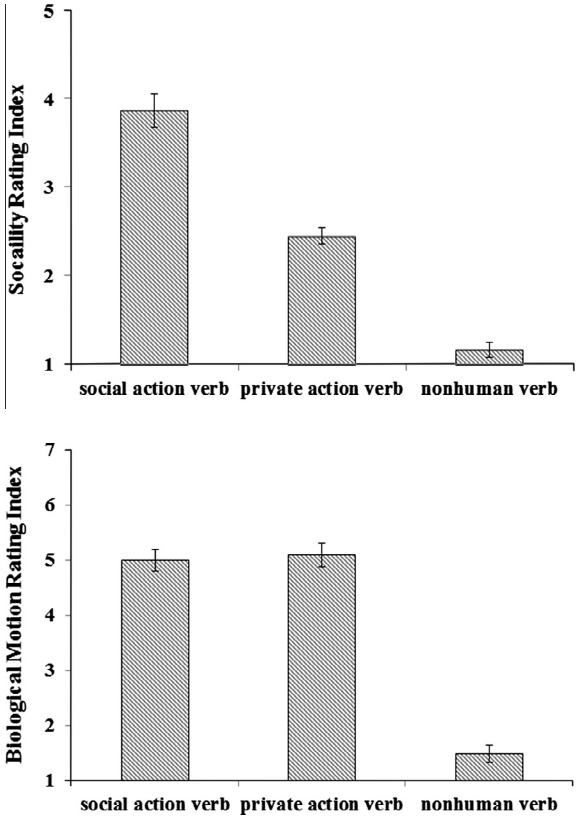


Fig. 1. Results of sociality and biological-motion rating experiments. The error bars depict the unbiased standard errors of the ratings of different verb conditions.

of sociality between our human and nonhuman verb conditions was confirmed by the rating results.

In the biological-motion rating experiment, 16 participants (11 females) were asked to rate on a seven-point scale (“1”: very low; “7”: very high) the extent to which the meaning of a verb brought to mind biological motion. The detailed procedure of the biological-motion rating experiment is identical to that in Lin et al. (2011). The social action verbs, private action verbs, and nonhuman verbs scored 5.00 (SD: 0.77), 5.09 (SD: 0.86), and 1.48 (SD: 0.60) points, respectively. Both the social and private action verbs scored significantly higher than the nonhuman verbs [*ts* (138) > 28; *ps* < 0.001], and the rating difference between them was not significant [*t* (138) < 1] (Fig. 1). The inter-rater reliability was high (*ICC* (2,16) = 0.976).

The fMRI experiment used a semantic judgment task. Each test trial included three verbs arranged in a triangular array on the display: one at the top and two at the bottom. The task was to quickly and accurately indicate which of the two bottom verbs was more closely related in meaning to the one on top. All three verbs for each trial came from the same verb class (social action verbs, private action verbs, or nonhuman verbs). For example,

拥抱(embrace)
握手(shake hands) 扒窃(steal)

The experiment involved 120 trials, 40 for each of the three conditions (for the 70 verbs in each condition, 50 verbs were repeated once, and 20 verbs were not repeated).

2.3. Procedure

We used an event-related design with two runs of 8 min 20 s each. Each run included 60 trials, 20 for each of the three conditions. In each trial, the word stimuli appeared for 3 s, followed by a jitter interval of at least 1 s. No trial was conducted during the first and last 10 s of each run. The order of test trials and the length of jitter intervals were optimized using optseq software (http://surfer.nmr.mgh.harvard.edu/optseq/).

In the scanner, the stimuli were back-projected via a video projector (refresh rate: 60 Hz; spatial resolution: 800 × 600) onto a translucent screen placed inside the scanner bore. Participants viewed the stimuli through a mirror located above their eyes. The background was always black with a red fixation point (dot) at the center of the screen, and all the words were white. Participants were instructed to make their choices by pressing a button with either their right index or middle finger. Before the formal experiment, each participant completed a practice run outside the scanner with additional stimuli; the procedure of this run was identical to that of the formal experiment.

2.4. Acquisition and analysis of magnetic resonance imaging data

Data on structural and functional magnetic resonance imaging (MRI) were collected with a 3T Siemens Trio Tim scanner at the BNU MRI Center. A $T2^*$-weighted gradient-echo planar imaging sequence was applied to acquire blood oxygen level-dependent signals (TR: 2000 ms; TE: 30 ms; flip angle: 90°; matrix size: 64 × 64; 33 slices; voxel size: 3.125 mm × 3.125 mm × 4 mm). A high-resolution 3D structural data set was acquired with a 3DMPRAGE sequence in the sagittal plane (TR: 2530 ms; TE: 3.39 ms; flip angle: 7°; matrix size: 256 × 256; 128 slices; voxel size: 1.33 mm × 1 mm × 1 mm).

The MRI data were analyzed using BrainVoyager QX software (Brain Innovation, Maastricht, The Netherlands). The first five volumes of functional data in each run were discarded. The functional data underwent preliminary processing, which included slice scan time correction (temporal interpolation by cubic spline function), 3D motion correction (co-registration of each volume of the time series to the first remaining volume of the second run by three-plane mechanical body transformations and trilinear spatial interpolation), spatial smoothing (Gaussian filter, 6 mm full width at half-maximum), and temporal filtering [high-pass (GLM-Fourier): 5 sines/cosines]. The functional data of each participant were registered to his/her anatomical data. Functional and anatomical volumes were transformed into a standardized space (Talairach & Tournoux, 1988). All functional data were subsequently analyzed using the general linear model (GLM). A random effect (RFX) GLM analysis was conducted to analyze the group data.

In the whole brain analyses, each two of the conditions were contrasted. In addition, an “RFX of conjunction” analysis of the contrasts “social action verb > private action verb” and “social action verb > nonhuman verb” was conducted to explore the brain regions showing stronger activation to the social verb condition than to the other two conditions. Although the above condition-based analysis can separate the effects of sociality and biological motion clearly and control the confounding effects of RT and word frequency, a parametric modelling analysis should detect the sociality effect more sensitively because we define the sociality as a continuous variable. Therefore, we conducted a parametric modelling analysis on the basis of the mean sociality rating of the three verbs in each trial. The false positive rate for all whole brain analyses was controlled at $\alpha < 0.05$ using the cluster-level statistical threshold estimator of BrainVoyager ($p < 0.01$, cluster size ⩾ 30 voxels).

To illustrate the relationship between the results of the present study and the findings of previous related studies, we conducted further ROI-based analyses. A first set of ROIs was set on the basis of two meta-analyses of previous ToM studies (Mar, 2011; Schurz

Table 1
Previous meta-analyses' coordinates included in our ROI-based analyses.

ROI class	ROI source	Anatomical region	Coordinates (*x*, *y*, *z*)			Space
ROIs based on activations in the ToM network	Mar (2011) (story-based)	Left TPJ/pSTS	−52	−58	26	MNI
		Right TPJ/pSTS	54	−54	26	MNI
		MPFC	4	58	28	MNI
		Precuneus	−10	−50	36	MNI
	Mar (2011) (nonstory-based)	Left TPJ/pSTS	−56	−48	4	MNI
		Right TPJ/pSTS	50	−46	18	MNI
		MPFC	−10	50	34	MNI
		Precuneus	2	−56	38	MNI
	Schurz et al. (2014)	Left IPL	−46	−63	41	MNI
		Left TPJp	−53	−59	20	MNI
		Left TPJa	−53	−30	10	MNI
		Left pMTG	−66	−52	2	MNI
		Right IPL	47	−61	39	MNI
		Right TPJp	56	−56	18	MNI
		Right TPJa	47	−35	12	MNI
		Right pMTG	64	−52	−3	MNI
		MPFC1	3	51	−7	MNI
		MPFC2	−1	54	24	MNI
		MPFC3	6	26	55	MNI
		Precuneus	4	−55	34	MNI
		Left middle temporal	−59	−23	−8	MNI
		Left anterior temporal	−51	0	−19	MNI
		Left inferior frontal	−46	22	8	MNI
		Right middle temporal	62	−22	−10	MNI
		Right anterior temporal	53	0	−21	MNI
		Right inferior frontal	44	20	12	MNI
	Van Overwalle and Baetens (2009)	Left TPJ	−50	55	25	Talairach
		Right TPJ	50	55	25	Talairach
ROIs based on activations in pSTS during action/biological motion observation and comprehension	Caspers et al. (2010)	Left pSTS	−54	−50	8	MNI
		Right pSTS	56	−40	4	MNI
	Grosbras et al. (2012)	Right pSTS	54	−54	10	MNI
	Van Overwalle and Baetens (2009)	Left pSTS	−50	−55	10	Talairach
		Right pSTS	50	−55	10	Talairach
	Watson et al. (2013)	Left pSTS	−58	−50	6	MNI

et al., 2014). Mar (2011) presented two separate meta-analyses for ToM studies using sentential (story-based) and nonverbal stimuli (nonstory-based). The results of both analyses included four most classic regions of the ToM network, i.e., the MPFC, PC/precuneus, and bilateral TPJs/pSTS, resulting in eight critical ROIs for our analyses. Schurz et al. (2014) conducted more fine-grained meta-analyses by sorting the ToM studies into six task groups and by dividing each classic region of the ToM network into subregions. They found a core network for all task groups and task-related activation differences surrounding this core-network. We included all ROIs defined by Schurz et al. (2014) so that we can compare the activation patterns evoked by our task to those of previous ToM tasks. A second set of ROIs was set on the basis of three meta-analyses of studies on action/biological motion observation and comprehension to examine whether the activation of the pSTS during action/biological perception and comprehension can be, at least partially, explained by the processing of social/mental features (Caspers et al., 2010; Grosbras et al., 2012; Watson, Cardillo, Ianni, & Chatterjee, 2013). A third set of ROIs was set on the basis of Van Overwalle and Baetens (2009). We paid special attention to that study because it considered the pSTS and TPJ to be two separate regions that belong to the sensory-motor and ToM systems respectively (but see Hein & Knight, 2008). To clarify whether the sociality effect does exist in the pSTS, we set separate pSTS and TPJ ROIs on the basis of the coordinates provided by Van Overwalle and Baetens (2009). All of these three sets of ROIs were defined on the basis of the coordinates reported in the literature (see Table 1 for all included coordinates). Coordinates reported in the MNI space were converted into the Talairach space (Talairach & Tournoux, 1988) by using the GingerALE software (BrainMap GingerALE 2.3; Research Imaging Center, University of Texas Health Science Center at San Antonio). The ROIs were defined as spheres with a 6-mm diameter centered in the reported or converted coordinates in the Talairach space.

A last set of ROIs were set on the basis of our previous study (Lin et al., 2011). In that study, we found that verbs referring to biological-motion events, in comparison with verbs referring to mechanical-motion events and low-motion events, evoke strong activation in the pSTS. We pay special attention to that study because it is the only study that clearly differentiates verbs that denote different types of motions as separate experimental conditions and demonstrates brain activation specific to biological-motion verbs in the pSTS. Although the observations in that study are consistent with the prevalent biological motion hypothesis, they can also be explained as the processing of social/mental features. The biological-motion verbs in our previous study include social and private action verbs, and its mechanical-motion and low-motion verbs correspond to nonhuman verbs. Therefore, we extracted the pSTS clusters showing stronger activation to biological-motion verbs than other verbs in Lin et al. (2011) as our ROIs (threshold: $p < 0.01$, cluster size $\geqslant 30$ voxels) to examine whether sociality can better explain the activation in the pSTS regions in Lin et al. (2011).

3. Results

The data of two subjects (both females) were discarded because of excessive head movement (>3.0 mm in any direction). Thus, all subsequent analyses were based on the remaining 15 participants.

3.1. Behavioral data

Response time (RT) and accuracy data were collected when the participants performed their instructed tasks in the scanner. We contrasted different conditions by using the within-subject paired *t*-test. The RT of the nonhuman verb condition was significantly longer than that of the other two conditions [mean RT (SD): social action verb, 1564 ms (227 ms); private action verb, 1561 ms (234 ms); and nonhuman verb, 1638 ms (207 ms); *t*-test results: social action verb vs. nonhuman verb, $t\,(14) = 3.08$, $p = 0.008$; social action verb vs. private action verb, $t\,(14) < 1$; and private action verb vs. nonhuman verb, $t\,(14) = 4.05$, $p = 0.001$]. No significant difference in accuracy was observed between the conditions [mean accuracy (SD): social action verb, 95.5% (3.6%); private action verb, 95.3% (4.5%); and nonhuman verb, 96.0% (4.1%); *t*-test results: $ts\,(14) < 1$].

3.2. FMRI data

3.2.1. Whole-brain analysis

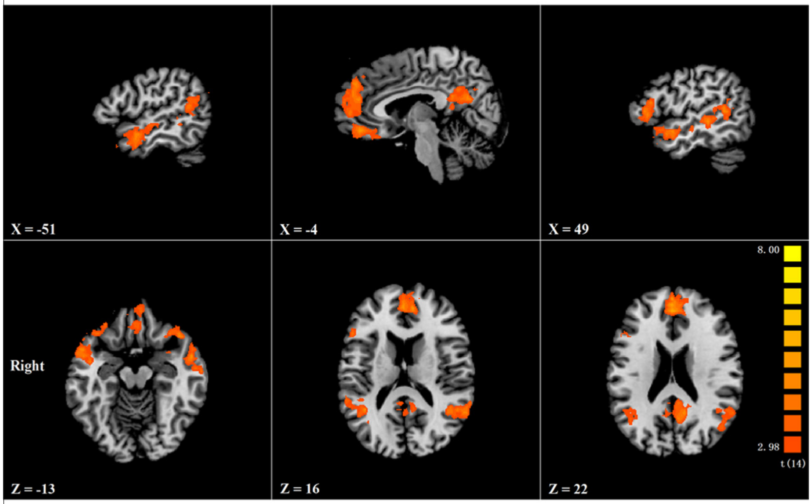
The results of the whole-brain contrasts are reported in Table 2. For the contrasts "social action verb > private action verb" and "social action verb > nonhuman verb," significant activations were observed in all classic regions of the ToM network (MPFC, PC/precuneus, bilateral TPJs/pSTS and ATLs). Activations were also observed in other brain regions observed in previous studies of social or emotional knowledge processing, including the ventral medial prefrontal cortex and inferior frontal gyrus (Mar, 2011; Vigliocco et al., 2014; Zahn et al., 2007). See Fig. 2A for the result

Table 2
Results of whole-brain analyses.

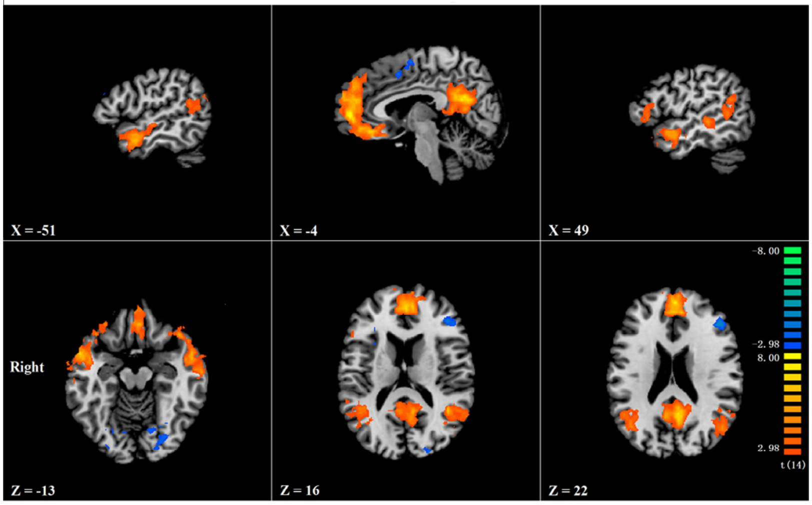
Contrast	Anatomical region	Cluster size (mm^3)	Talairach coordinates (x, y, z)			Peak t value
Condition-based analysis						
Social action verb > private action verb	Left Middle Temporal Gyrus	11,162	−54	−10	−5	8.87
	Right Middle Frontal Gyrus	7612	45	17	28	11.33
	Right Middle Temporal Gyrus	6236	54	−31	1	7.36
	Left Inferior Frontal Gyrus	5797	−42	20	−14	7.09
	Left Medial Frontal Gyrus	3964	−6	47	37	7.73
	Left Precuneus	1651	−12	−46	31	6.26
	Right Medial Frontal Gyrus	988	3	44	−8	4.96
	Left Lingual Gyrus	918	−18	−79	−5	6.86
Social action verb > non-human verb	Right Superior Frontal Gyrus	30,124	12	32	46	10.38
	Right Middle Temporal Gyrus	19,863	54	5	−11	8.88
	Right Cingulate Gyrus	11,119	6	−55	28	6.71
	Left Middle Temporal Gyrus	8263	−48	2	−17	7.91
	Left Superior Temporal Gyrus	7848	−39	−49	16	7.17
	Left Inferior Frontal Gyrus	4377	−39	23	−11	8.18
Private action verb > social action verb	No significant cluster was found					
Private action verb > non-human verb	Right Posterior Cingulate	5435	9	−49	13	7.70
	Left Medial Frontal Gyrus	4806	0	56	10	6.41
	Left Middle Temporal Gyrus	1429	−36	−70	28	4.36
	Left Superior Frontal Gyrus	1283	−15	41	37	5.80
	Right Middle Temporal Gyrus	841	45	−67	25	4.80
Non-human verb > social action verb	Right Parahippocampal Gyrus	27,806	24	−58	−5	9.51
	Left Middle Frontal Gyrus	2382	−48	35	28	6.59
	Left Postcentral Gyrus	1884	−30	−34	67	5.28
	Right Superior Frontal Gyrus	1201	3	8	49	10.24
Non-human verb > private action verb	Left Lingual Gyrus	63,183	−15	−85	−2	15.81
	Left Middle Frontal Gyrus	4115	−39	20	31	8.03
	Left Superior Frontal Gyrus	1639	−3	11	49	6.23
	Right Insula	1426	30	26	4	10.35
	Left Precuneus	1339	−24	−61	31	6.58
	Right Middle Frontal Gyrus	1327	42	29	31	6.15
The RFX of conjunction analysis: social action verb > private action verb and social action verb > non-human verb	Right Medial Frontal Gyrus	9270	3	47	19	7.08
	Right Superior Temporal Gyrus	6526	39	−49	16	6.52
	Right Inferior Frontal Gyrus	6357	48	29	10	6.22
	Left Superior Temporal Gyrus	5740	−48	−49	16	5.60
	Left Middle Temporal Gyrus	5065	−54	−10	−5	5.78
	Right Posterior Cingulate	5007	6	−40	13	5.58
	Left Inferior Frontal Gyrus	3463	−45	23	1	6.79
	Left Medial Frontal Gyrus	2538	−3	44	−8	5.71
Parametric analysis of sociality						
Activation	Left Medial Frontal Gyrus	26,215	−7	49	6	10.51
	Left Posterior Cingulate	11,601	−4	−56	24	7.99
	Left Superior Temporal Gyrus	8903	−43	22	−24	7.91
	Right Middle Temporal Gyrus	6666	47	−32	0	5.92
	Right Middle Temporal Gyrus	6647	53	4	−12	8.40
	Left Superior Temporal Gyrus	5393	−40	−50	18	6.87
Deactivation	Left Middle Occipital Gyrus	3727	−22	−89	12	−7.38
	Left Middle Frontal Gyrus	1996	−40	31	27	−6.40
	Right Fusiform Gyrus	1793	23	−59	−6	−6.12
	Right Superior Frontal Gyrus	1341	2	7	48	−10.89
	Right Lingual Gyrus	1152	14	−83	−9	−6.44
	Right Claustrum	890	29	22	6	−6.51

Note: The anatomical regions were identified by inputting the peak Talairach coordinates into the Talairach Client (http://www.talairach.org).

A. RFX of conjunction: SV > NV & SV > PV



B. Parametric modulation of sociality ratings



C. FFX results of females and males

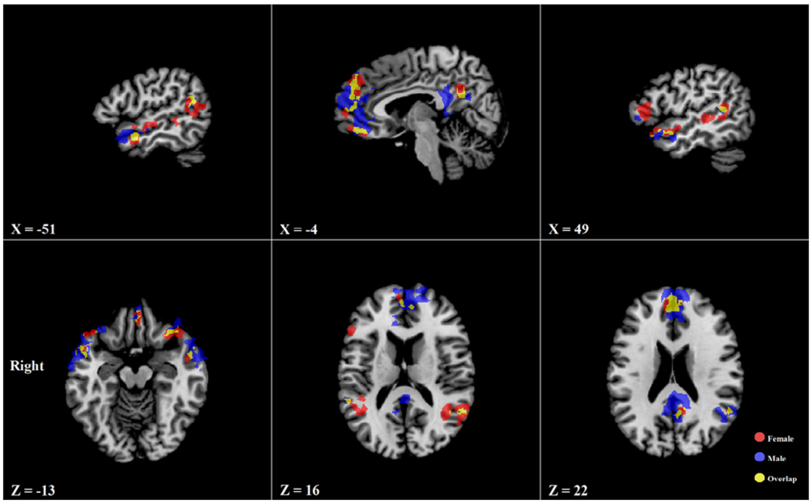


Fig. 2. Representative slices ($x = -51$, $x = -4$, $x = 49$, $z = -13$, $z = 16$, and $z = 22$) of the results of whole-brain analyses (corrected $\alpha < 0.05$: $p < 0.01$, cluster size $\geqslant$ 30 voxels): (A) results of the "RFX of conjunction" analysis of the contrasts "social action verb > private action verb" and "social action verb > nonhuman verb"; (B) results of the parametric analysis of sociality; and (C) results of the FFX analyses in female and male participants.

of the "RFX of conjunction" analysis, which reflects the consistent results of the two contrasts.

The private action verbs evoked stronger activations than the nonhuman verbs in the MPFC, left superior frontal gyrus, PC, and bilateral TPJs. These brain regions overlap with the brain network we found in the RFX of conjunction analysis of the contrasts "social action verb > private action verb" and "social action verb > nonhuman verb," indicating that social/mental knowledge can be activated even in the processing of action verbs with low sociality. No brain region showed stronger activation for the private action verbs than that for the social action verbs.

The nonhuman verbs evoked stronger activation than the other two classes of verbs in large areas of the occipital cortex, posterior temporal cortex, and frontal gyrus. Considering the longer RT for nonhuman verbs than that for the other two kinds of verbs, the stronger activations may reflect the item difficulty of the nonhuman verb condition.

The results of the parametric modelling analysis are very similar to those of the condition-based analysis. Strong correlations between brain activation and sociality ratings were observed in MPFC, PC/precuneus, bilateral TPJs/pSTS and ATLs. The results of the parametric modelling analysis are shown in Fig. 2B.

Given that our study sample largely comprises female participants, we further conducted fixed effect GLM analyses in our female and male participants respectively to examine whether our results reflect the data of our female participants only. We adopted the conjunction of "social action verb > private action verb" and "social action verb > nonhuman verb" as the contrast of interest. The results of female and male participants are very similar and they overlap in all critical regions of the ToM network, including MPFC, precuneus, bilateral TPJs/pSTS and ATLs. The results are shown in Fig. 2C.

3.2.2. ROI-based analyses

We conducted a series of ROI-based analyses to further illustrate the relationship between the result of the present study and the findings of previous related studies of ToM processing, action/biological motion observation, and action semantic processing.

A first question to be answered is whether the brain network activated by social action verbs overlaps with the classic ToM network obtained by previous ToM studies, which includes at least four key regions: MPFC, PC/precuneus, and bilateral TPJs/pSTS. This classic network corresponds to the eight ROIs defined on the basis of Mar (2011) and four of the 18 ROIs defined on the basis of Schurz et al. (2014): MPFC2, Precuneus, left TPJp and right TPJp. As shown in Table 3, within-subject paired *t*-tests demonstrated that all those ROIs showed "social action verb > nonhuman verb" effects

Table 3
Results of ROI-based analyses.

ROI class	ROI source	Anatomical region	Social action verb – nonhuman verb		Social action verb – private action verb		Private action verb – nonhuman verb		Parametric modulation of sociality	
			t	*p*	*t*	*p*	*t*	*p*	*t*	*p*
ROIs based on activations in the ToM network	Mar (2011) (story-based)	Left TPJ/pSTS	3.48^{**}	0.004	3.18^{**}	0.007	1.72	0.107	3.09^{**}	0.008
		Right TPJ/pSTS	3.94^{**}	0.001	1.57	0.140	2.37^{*}	0.033	3.92^{**}	0.002
		MPFC	3.50^{**}	0.004	2.86^{*}	0.013	1.69	0.113	3.10^{**}	0.008
		Precuneus	3.76^{**}	0.002	2.82^{*}	0.014	2.64^{*}	0.019	3.75^{**}	0.002
	Mar (2011) (nonstory-based)	Left TPJ/pSTS	3.01^{**}	0.009	4.70^{***}	0.000	−1.57	0.139	2.12^{+}	0.052
		Right TPJ/pSTS	4.94^{***}	0.000	4.43^{**}	0.001	1.37	0.192	4.07^{**}	0.001
		MPFC	3.76^{**}	0.002	2.09^{+}	0.056	2.53^{*}	0.024	3.97^{**}	0.001
		Precuneus	4.43^{**}	0.001	3.03^{**}	0.009	1.28	0.222	5.54^{***}	0.000
	Schurz et al. (2014)	Left IPL	2.26^{*}	0.040	2.13^{+}	0.051	1.26	0.229	1.96^{+}	0.071
		Left TPJp	3.67^{**}	0.003	2.77^{*}	0.015	1.58	0.137	3.34^{**}	0.005
		Left TPJa	0.76	0.460	0.06	0.955	0.63	0.539	0.53	0.603
		Left pMTG	1.63	0.125	1.65	0.122	0.99	0.340	2.15^{*}	0.049
		Right IPL	3.33^{**}	0.005	1.27	0.224	2.64^{*}	0.020	2.46^{*}	0.027
		Right TPJp	3.10^{**}	0.008	2.09^{+}	0.055	2.59^{*}	0.021	3.07^{**}	0.008
		Right TPJa	1.70	0.111	4.01^{**}	0.001	−0.56	0.581	1.84^{+}	0.087
		Right pMTG	2.44^{*}	0.029	1.70	0.111	1.64	0.123	1.94^{+}	0.073
		MPFC1	3.86^{**}	0.002	2.00^{+}	0.065	3.15^{**}	0.007	3.65^{**}	0.003
		MPFC2	5.54^{***}	0.000	3.55^{**}	0.003	2.48^{*}	0.026	5.73^{***}	0.000
		MPFC3	1.00	0.336	1.84^{+}	0.087	−0.40	0.695	0.72	0.483
		Precuneus	5.67^{***}	0.000	3.37^{**}	0.005	2.57^{*}	0.022	6.01^{***}	0.000
		Left middle temporal	3.21^{**}	0.006	2.57^{*}	0.022	1.02	0.325	1.87^{+}	0.082
		Left anterior temporal	5.64^{***}	0.000	7.28^{***}	0.000	1.63	0.126	5.41^{***}	0.000
		Left inferior frontal	0.13	0.900	1.45	0.170	−1.54	0.145	−0.48	0.637
		Right middle temporal	3.17^{**}	0.007	3.19^{**}	0.007	−0.37	0.714	1.88^{+}	0.082
		Right anterior temporal	5.18^{***}	0.000	3.96^{**}	0.001	2.16^{*}	0.049	5.71^{***}	0.000
		Right inferior frontal	−0.33	0.749	1.31	0.212	-2.03^{+}	0.062	0.29	0.774
	Van Overwalle and Baetens (2009)	Left TPJ	3.23^{**}	0.006	2.59^{*}	0.021	2.27^{*}	0.039	2.44^{*}	0.028
		Right TPJ	3.32^{**}	0.005	1.59	0.134	2.48^{*}	0.026	2.77^{*}	0.015
ROIs based on activations in pSTS during action/biological motion observation and comprehension	Caspers et al. (2010)	Left pSTS	2.46^{*}	0.027	4.53^{***}	0.000	−0.92	0.373	2.19^{*}	0.046
		Right pSTS	2.78^{*}	0.015	2.95^{*}	0.010	−0.03	0.979	2.73^{*}	0.016
	Grosbras et al. (2012)	Right pSTS	3.09^{**}	0.008	2.24^{*}	0.042	3.52^{**}	0.003	3.41^{**}	0.004
	Lin et al. (2011)	Left pSTS	5.39^{***}	0.000	2.70^{*}	0.017	3.83^{**}	0.002	5.83^{***}	0.000
		Right pSTS	6.52^{***}	0.000	4.72^{***}	0.000	3.20^{**}	0.006	4.79^{***}	0.000
	Van Overwalle and Baetens (2009)	Left pSTS	3.42^{**}	0.004	2.84^{*}	0.013	1.76^{+}	0.100	3.57^{**}	0.003
		Right pSTS	2.91^{*}	0.011	2.22^{*}	0.044	2.96^{*}	0.010	3.16^{**}	0.007
	Watson et al. (2013)	Left pSTS	2.80^{*}	0.014	4.83^{***}	0.000	−1.04	0.315	2.56^{*}	0.023

$^{+}$ $p < 0.1$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

and that all but one ROI (Mar, 2011: right TPJ/pSTS, story-based) showed "social action verb > private action verb" effects. All of the ROIs showed sociality effects in the parametric modelling analysis. These results demonstrated that the brain network showing specific activation for social action verbs in the present study corresponds to the classic ToM network obtained by previous studies.

A second question we are interested in is functional subdivisions in the brain network observed in the present study. Schurz et al. (2014) explored the functional subdivisions in the ToM network by sorting the ToM studies into six task groups and comparing the results of different task groups in 18 subregions of the ToM network. We conducted ROI-based analyses following Schurz et al. (2014). As shown in Table 3, among the 18 ROIs defined on the basis of Schurz et al. (2014), nine ROIs showed consistent sociality effects in both of the condition-based and parametric analyses, including left TPJp, right TPJp, MPFC1, MPFC2, precuneus, left middle temporal, left anterior temporal, right middle temporal and right anterior temporal. According to Schurz et al. (2014), MPFC and bilateral TPJs consist of a core network for ToM that is consistently engaged in all sorts of ToM tasks; the rest of these regions engage only in some particular types of ToM tasks. We will discuss the functions of these regions in ToM processing by comprehensively considering the results of the present study and those of Schurz et al. (2014).

A third question to be answered is whether the processing of social/mental features can at least partially explain the activation of the pSTS in previous studies of action/biological perception and action concept processing. To answer this question, six pSTS ROIs were defined on the basis of two meta-analyses of action/biological motion observation (Caspers et al., 2010; Grosbras et al., 2012), one meta-analysis of action semantic processing (Watson et al., 2013), and our prior study on specific brain activations for verbs denoting biological motion (Lin et al., 2011). As shown in Table 3, all six pSTS ROIs showed significant "social action verb > nonhuman verb" and "social action verb > private action verb" effects. Considering that the biological-motion ratings and RTs of the social and private action verbs were well matched, the significant "social action verb > private action verb" effects can only be attributed to the difference in the social/mental aspects of their meanings. The "private action verb > nonhuman verb" effects, strongly predicted by the prevalent biological motion hypothesis, were observed in only half of these ROIs. Considering that some researchers consider the pSTS and TPJ to be two separate regions with distinct functions, we conducted further analyses in separate pSTS and TPJ ROIs following Van Overwalle and Baetens (2009). As shown in Table 3, the pSTS ROIs showed strong and reliable sociality effects in both of the condition-based and parametric analyses, indicating their role in social/mental semantic processing.

4. Discussion

The present study focused on determining whether the ToM network can be selectively activated by the thinking of verbs that denote social actions. Brain activation profiles elicited by social action verbs, private action verbs, and nonhuman verbs were compared. In both condition-based and parametric analyses, all classic regions of the ToM network showed stronger activation to the social action verb condition than to the other conditions. In addition, the ROI-based analyses confirmed that the brain regions showing specific activation to the social action verb condition overlap with the classic ToM network obtained by previous ToM studies and the pSTS region obtained by previous studies of action/biological motion observation and comprehension.

The key result of the present study is the co-occurrence of the "social action verb > nonhuman verb" and "social action verb > private action verb" effects in the ToM network (including the pSTS). When explaining the "social action verb > nonhuman verb" effect, the significantly low biological-motion rating and long RT of the nonhuman verb condition should be considered: the effect can be explained as a reflection of the processing of biological-motion properties or that of the item difficulty reflected by the RT. However, the biological-motion rating and RT were well matched between the social and private action verb conditions, and the contrast "social action verb > private action verb" significantly activated brain regions that clearly overlapped with those showing the "social action verb > nonhuman verb" effect. Therefore, the only factor that can easily explain the stronger activation for the social action condition than that for the other two conditions is the social knowledge of verb meanings. In addition to the key result, a "private action verb > nonhuman verb" effect was also observed in the ToM network. However, this effect should be carefully interpreted because the two conditions are significantly different not only in the sociality ratings but also in the biological-motion ratings and RTs.

Our findings provide novel insights into the cognitive function of the ToM network by demonstrating that it supports the processing of social/mental semantic knowledge of verb meanings. The engagement of the ToM network in the semantic processing of single verbs has rarely been investigated or considered in the literature. In the present study, we proposed a novel dimension of verb meaning – to what extent social interactions are indicated, and demonstrated it to be a factor modulating ToM activity during verb comprehension. Although previous studies have showed similar sociality effects in cartoon reading tasks (Ciaramidaro et al., 2007; Walter et al., 2004), the underlying cognitive processes could be different. The social/mental knowledge of verb meanings is a part of our lexical semantic knowledge so that it can be accessed directly from semantic memory. Therefore, the results of the present study show that any online integrating or reasoning process in order to understand a social interaction is not a necessary condition for the engagement of the ToM network.

In addition, by conducting ROI analyses following Schurz et al. (2014), we provided new evidence about task effects in ToM activity. We observed sociality effects in all of the three regions (i.e., MPFC and bilateral TPJs) that were consistently found in all types of ToM tasks in Schurz et al. (2014), in accord with the idea that these regions consist of a core network for ToM (Mar, 2011; Schurz et al., 2014). We also observed sociality effects in several ROIs that have shown task-related activation differences in Schurz et al. (2014), which are distributed in the precuneus and bilateral ATLs. According to Schurz et al. (2014), the precuneus is activated in the ToM tasks requiring mental imagery but not those showing visual action directly. Consistent with this idea, the present study used action verbs as its stimuli that do not show visual action directly and can easily evoke mental imageries about human actions. The engagements of the ATLs in ToM and in semantic processing have been suggested by both fMRI and lesion studies (Duval et al., 2012; Irish, Hodges, & Piguet, 2014; Patterson, Nestor, & Rogers, 2007; Zahn et al., 2007). As proposed by Schurz et al. (2014), the activation of these regions in previous ToM studies can be explained by a hypothesis that the ATLs represent social concepts (Zahn et al., 2007). In accord with this hypothesis, we observed activation of bilateral ATLs in a semantic judgment task and found sociality of verb meanings as a factor modulating that activation.

Another important finding of the present study is that the activation of the pSTS during verb comprehension reflects the processing of social/mental properties of verb meanings. In previous studies of word comprehension, the activation of the pSTS was frequently explained as the processing of motion properties or biological motion properties (Chao et al., 1999; Kemmerer et al., 2008;

Lin et al., 2011; Martin, 2007). However, this prevalent biological motion hypothesis cannot explain the significant "social action verb > private action verb" effect we observed in the pSTS region, especially considering that the richness of biological properties was well matched between the social and private action verb conditions. Therefore, we propose that the activation of the pSTS during word comprehension as well as that during action/biological motion observation reflects the processing of the social/mental properties of action concepts but not, or not only, that of biological-motion properties. Two sets of evidence accord with our hypothesis. First, two studies have found that the activation of the pSTS during action/biological motion observation is modulated by the processing of intentions. De Lange, Spronk, Willems, Toni, and Bekkering (2008) observed an increase in activation in the right pSTS when participants paid attention to the intentionality of the action compared with when they paid attention to the means of action. Morris, Pelphrey, and McCarthy (2008) found that the left pSTS showed stronger activation in the observation of intended hand motions than that in the observation of unintended hand motions. These findings, together with the results of the present study, indicate that the activation of the pSTS during action/biological motion observation and comprehension is at least partially due to the processing of the social/mental properties of actions. The second set of evidence is that three studies have manipulated the richness of the motion features of verbs and observed no motion preference, or even a reverse effect, in the posterior lateral temporal region (Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008; Bedny, Caramazza, Pascual-Leone, & Saxe, 2012; Grossman et al., 2002). Although these studies did not consider the possibility that words can activate the pSTS as a part of the ToM network (for detailed comments see Lin et al., 2011), their findings indicated that factors other than the richness of motion properties modulate the activation of the pSTS during word comprehension.

The finding that the ToM network activates in the semantic processing of single words opens a new avenue for studying the cognitive functions of the network. In comparison with story/sentential comprehension tasks, word comprehension tasks have several methodological advantages: (1) the cognitive processes underlying word comprehension are fewer and easily controlled than those underlying story/sentence comprehension; (2) the use of word stimuli allows researchers to easily and precisely consider and control the effects of potential confounding factors, such as word frequency and RT; and (3) the use of word stimuli also allows researchers to include many trials and conditions within a limited scanning time because a word is processed faster than a story or sentence. Therefore, ToM studies using word stimuli will considerably promote our understanding of the ToM network by effectively and precisely identifying what types of informational input and processing trigger and modulate the activation of the network.

Finally, the result of the present study sheds new light on how language exposure facilitates the development of ToM by demonstrating a strong relationship between ToM processing and word processing. Previous studies of deaf children have suggested that language exposure is necessary for the development of verbal and non-verbal ToM abilities (Peterson & Siegal, 1999; Woolfe, Want, & Siegal, 2002). One explanation for this phenomenon is that the acquisition of sentential complements is a necessary precursor for the child's ability to represent beliefs (de Villiers, 2007). However, this explanation contradicts the recent finding that the ToM ability of 17- to 26-month-old hearing infants is significantly better than that of their deaf counterparts (Meristo et al., 2012). These infants are too young to acquire sentential complements, so that the development of ToM may benefit from simpler and earlier-mastered linguistic structures. Our findings suggest a possibility that the development of ToM is facilitated by action verbs that imply mental states. Tomasello and Kruger (1992) found that in most cases, mothers use action verbs to refer to impending actions but not to ongoing or completed actions in their conversations with infants. The experience that hearing action verbs precedes seeing corresponding actions may trigger the maturation of ToM by indicating to the infant the causal relationships between inner mental states (reflected by the hearing of verbs) and behavior. Therefore, we suggest that the development of ToM should be investigated in relation to the acquisition of words, especially of action verbs.

In summary, we investigated how the ToM network activates during verb comprehension. In a semantic judgment task, social action verbs evoked stronger activation than private action verbs and nonhuman verbs in all classic regions of the ToM network, including the pSTS. The activation patterns can be easily explained by sociality of verb meanings but not by other possible factors, such as item difficulty and biological motion. We draw two conclusions from our findings that the ToM network supports the processing of social/mental knowledge of verb meanings and that the activation of the pSTS during word comprehension mainly reflects the processing of social/mental properties but not that of biological-motion properties.

Acknowledgments

This study was supported by the 973 Program (2013CB837300), the Natural Science Foundation of China Grant (31300842, 31070904, 31222024, and 31171073), the Scientific Foundation of the Institute of Psychology, Chinese Academy of Sciences (Y2CQ032005), the Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (CNLYB1209), and the Knowledge Innovation Program of the Chinese Academic Sciences (KSCX2-YW-BR-6). We thank Xiping Liu, Jing Li, Xiaoying Wang, and Xiaosha Wang for their helpful discussions.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron, 34*(1), 149–159.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *The Journal of Neuroscience, 28*(44), 11347–11353.

Bedny, M., Caramazza, A., Pascual-Leone, A., & Saxe, R. (2012). Typical neural representations of action verbs develop without vision. *Cerebral Cortex, 22*(2), 286–293.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage, 50*(3), 1148–1167.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*(10), 913–919.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia, 45*(13), 3105–3113.

De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., & Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current Biology, 18*(6), 454–457.

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua, 117*(11), 1858–1878.

Duval, C., Bejanin, A., Piolino, P., Laisney, M., de La Sayette, V., Belliard, S., et al. (2012). Theory of mind impairments in patients with semantic dementia. *Brain, 135*(1), 228–241.

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*(1), 11–21.

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human Brain Mapping, 33*(2), 431–454.

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., et al. (2002). Neural representation of verb meaning: An fMRI study. *Human Brain Mapping, 15*(2), 124–134.

Han, Z., Bi, Y., Chen, J., Chen, Q., He, Y., & Caramazza, A. (2013). Distinct regions of right temporal cortex are associated with biological and human–agent motion: Functional magnetic resonance imaging and neuropsychological evidence. *The Journal of Neuroscience, 33*(39), 15442–15453.

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—It's my area: Or is it? *Journal of Cognitive Neuroscience, 20*(12), 2125–2136.

Irish, M., Hodges, J. R., & Piguet, O. (2014). Right anterior temporal lobe dysfunction underlies theory of mind impairments in semantic dementia. *Brain, 137*(4), 1241–1253.

Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language, 107*(1), 16–43.

Lin, N., Lu, X., Fang, F., Han, Z., & Bi, Y. (2011). Is the semantic category effect in the lateral temporal cortex due to motion property differences? *Neuroimage, 55*(4), 1853–1864.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1–3), 59–70.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology, 62*, 103–134.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25–45.

Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., et al. (2012). Belief attribution in deaf and hearing infants. *Developmental Science, 15*(5), 633–640.

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage, 28*(4), 757–762.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences of the United States of America, 99*(23), 15238–15243.

Morris, J. P., Pelphrey, K. A., & McCarthy, G. (2008). Perceived causality influences brain activity evoked by biological motion. *Social Neuroscience, 3*(1), 16–25.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976–987.

Peterson, C. C., & Siegal, M. (1999). Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological Science, 10*(2), 126–129.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Saxe, R. (2006). Why and how to study Theory of mind with fMRI. *Brain Research, 1079*(1), 57–65.

Saxe, R. (2009). Theory of mind (neural basis). In W. P. Banks (Ed.), *Encyclopedia of consciousness* (pp. 401–409). Oxford: Academic Press.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9–34.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420.

Spunt, R. P., Falk, E. B., & Lieberman, M. D. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Science, 21*(11), 1593–1598.

Sun, H., Huang, J., Sun, D., Li, D., & Xing, H. (1997). Introduction to language corpus system of modern Chinese study. In M. Y. Hu (Ed.), *Paper collection for the fifth world Chinese teaching symposium* (pp. 459–466). Beijing: Peking University Press.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.

Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*(2), 311–333.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage, 48*(3), 564–584.

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex, 24*(7), 1767–1777.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., & Bara, B. G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience, 16*(10), 1854–1863.

Watson, C. E., Cardillo, E. R., Ianni, G. R., & Chatterjee, A. (2013). Action concepts in the brain: An activation likelihood estimation meta-analysis. *Journal of Cognitive Neuroscience, 25*(8), 1191–1205.

Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development, 73*(3), 768–778.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 104*(15), 6430–6435.