

# Spontaneous mentalizing captures variability in the cortical thickness of social brain regions

Katherine Rice and Elizabeth Redcay

Department of Psychology, University of Maryland, College Park, MD 20742, USA

**Theory of mind (ToM)—or thinking about the mental states of others—is a cornerstone of successful everyday social interaction. However, the brain bases of ToM are most frequently measured via explicit laboratory tasks that pose direct questions about mental states (e.g. “In this story, what does Steve think Julia believes?”). Neuroanatomical measures may provide a way to explore the brain bases of individual differences in more naturalistic everyday mentalizing. In the current study, we examined the relation between cortical thickness and spontaneous ToM using the novel Spontaneous Theory of Mind Protocol (STOMP), which measures participants’ spontaneous descriptions of the beliefs, emotions and goals of characters in naturalistic videos. We administered standard ToM tasks and the STOMP to young adults (aged 18–26 years) and collected structural magnetic resonance imaging data from a subset of these participants. The STOMP produced robust individual variability and was correlated with performance on traditional ToM tasks. Further, unlike the traditional ToM tasks, STOMP performance was related to cortical thickness for a set of brain regions that have been functionally linked to ToM processing. These findings offer novel insight into the brain bases of variability in naturalistic mentalizing performance, with implications for both typical and atypical populations.**

**Keywords:** individual differences; social cognition; structural MRI; theory of mind

## INTRODUCTION

Typical adults vary in how, and how much, they attend to the thoughts, feelings and intentions of others in daily life. Two people on a bus may sit across from the same couple and yet come away with different narratives; one observer only notices two people talking, while the other observer detects the romantic tension, the nervousness and one partner’s desire to impress. However, if given a list of explanations for the interaction (e.g. first date, argument, two strangers), both observers could identify the correct scenario. That is, when considering how individuals think about others’ internal states—also known as theory of mind (ToM) or mentalizing—there may be a difference between the capacity to answer explicit questions and patterns of spontaneous attention in naturalistic situations (Champagne-Lavau and Moreau, 2013; Koster-Hale and Saxe, 2013). Evidence from atypical development supports a distinction between offline belief reasoning and real-time mentalizing (Senju *et al.*, 2009), but, among typical adults, the brain and behavioral bases for individual differences in spontaneous ToM remain underexplored (Achim *et al.*, 2013).

Most functional magnetic resonance imaging (fMRI) studies of mentalizing employ explicit measures on which typical adults are at ceiling (Koster-Hale and Saxe, 2013). For example, participants may be asked about beliefs of characters in a story (e.g. Dodell-Feder *et al.*, 2011), the emotions of a face (e.g. Adams *et al.*, 2010) or the progression of a cartoon story (e.g. Sommer *et al.*, 2007). Such studies have revealed a robust network of regions that activate during mentalizing, including the temporal parietal junction (TPJ), superior temporal sulcus (STS) and medial prefrontal cortex (mPFC; Saxe, 2009). Similar regions are also activated in tasks employing ambiguous

animations (Castelli *et al.*, 2000; Tavares *et al.*, 2008). However, research on the neural correlates of mentalizing grows sparser as the measures grow more naturalistic, likely because of the difficulty of imaging participants during everyday mentalizing (although see Spiers and Maguire, 2006; Wagner *et al.*, 2011; and Spunt & Lieberman, 2012 for exceptions).

Structural measures may provide a way to probe the neural bases of individual differences in spontaneous ToM. Cortical thinning—an index of maturation—continues through the early twenties in the social brain regions involved in ToM processing (Gogtay *et al.*, 2004; Mills *et al.*, 2014). In an age window where cortical thinning is developmentally appropriate, thinner cortex has been linked to increased selectivity in brain activation (Wendelken *et al.*, 2011), improved creativity (Jung *et al.*, 2010), improved executive function (Lu *et al.*, 2009; Churchwell and Yurgelun-Todd, 2013; Kharitonova *et al.*, 2013), lower rates of hyperactive behaviors (Shaw *et al.*, 2011) and diminished depression and anxiety (Ducharme *et al.*, forthcoming). Although individual differences in typical adults’ ToM performance have been related to amygdala volume (Dziobek *et al.*, 2006b; Rice *et al.*, 2013), and mPFC volume (Lewis *et al.*, 2011), the relation between spontaneous ToM and cortical thinning has not been explored.

Extant behavioral ToM measures, however, are not well suited to examine the relation between spontaneous ToM ability and cortical thickness in typical adults. Typical adults perform at ceiling on classic developmental ToM tasks (e.g. false belief; Koster-Hale and Saxe, 2013), and extensions of these tasks to adults often involve increases in memory load and syntactic complexity (Kinderman *et al.*, 1998). Adult ToM measures that separate clinical and typical populations also often produce ceiling effects for typical adults (reviewed in Dodell-Feder *et al.*, 2013). Although recent ToM measures have employed continuous measurements, including reaction time (Keysar *et al.*, 2003; Samson *et al.*, 2010) and eye movements (Schneider *et al.*, 2012), it is unclear if resultant variability is meaningful or how these laboratory tasks relate to everyday mentalizing.

Recent measures of adult ToM have aimed for stronger ecological validity, by posing questions about mental states based on videos of social interaction (McDonald *et al.*, 2003; Dziobek *et al.*, 2006a).

Received 24 February 2014; Revised 2 May 2014; Accepted 14 May 2014

We thank Jeremiah Baker, Sarah Blankenship, Toby Hamovitz, Eric Meyer, Brogan Murphy, Daniel O’Young, Kayla Velnoskey and Briana Viscomi for assistance with data collection and analysis; Vivi Bauman, Jessica Golding, Jessica Lee and Nina Lichtenberg for assistance with data coding; Nina Lichtenberg and Ruth Ludlum for assistance with stimuli creation; and the Maryland Neuroimaging Center and staff for project assistance.

This research was partially supported by an National Science Foundation Graduate Research Fellowship to K.R., and a University of Maryland Dean’s Research Initiative grant awarded to E.R. and Dr. Tracy Riggins.

Correspondence should be addressed to Katherine Rice, Department of Psychology, University of Maryland, College Park, MD 20742, USA. E-mail: krice@umd.edu.

These measures produce differences between clinical and typical populations (McDonald *et al.*, 2004; Dziobek *et al.*, 2006a; McDonald *et al.*, 2006; Montag *et al.*, 2010; Preißler *et al.*, 2010) but often produce a limited range for typical adults. The recent Short Story Task (SST) produced variability by posing questions about mental states in a naturalistic short story (Dodell-Feder *et al.*, 2013). The SST also assessed whether participants' descriptions of the story contained a mental state, but the resultant binary measure had limited variability. Thus, these tasks are still limited by their explicit nature; participants are directly asked questions about mental states.

One classic task that directly measures spontaneous mental state attribution is Heider and Simmel's (1944) presentation of moving geometric shapes. When asked to describe the animation, almost all typical adults attribute intention to the shapes, despite the lack of explicit social content. Although there are some individual differences in these attributions (Scholl and Tremoulet, 2000), much of the work on animacy attribution has focused on differences between clinical and typical populations (Klin, 2000; Castelli *et al.*, 2002; Klin and Jones, 2006), and not on typical variability. Further, everyday situations involve deciphering the complex behaviors of people, not shapes.

In the current study, we overcome previous limitations by examining the neural and behavioral correlates of a novel ToM task: the Spontaneous Theory of Mind Protocol (STOMP). During the STOMP, participants watch silent film clips depicting complex social scenes and then describe what they saw. Participant responses are coded to determine their propensity to report the internal states of others.

The primary aim of the current study was to determine whether STOMP performance was correlated with cortical thinning in the late-developing brain regions associated with mentalizing. Thus, we first present validation of the STOMP and then analyze the structural correlates of STOMP performance. We hypothesize that individual differences in STOMP performance will be related to cortical thinning in regions functionally linked to ToM processing.

## METHODS

### Participants

Eighty adults (28 males) participated in exchange for course credit or payment. All adults were aged 18–26 years (mean = 20.34, *s.d.* = 1.92). Of those 80 adults, 59 participants (19 males, mean age = 19.96, *s.d.* = 1.75) completed the STOMP and a battery of other behavioral assessments (described in the following section). Structural MRI data were collected from 18 of these participants. An additional 21 participants were recruited to complete the STOMP and a structural MRI. One participant from this group was not scanned due to claustrophobia. Thus, 38 participants contributed STOMP and MRI data (18 males, mean age = 21.05, *s.d.* = 1.68). All scanned participants were native English speakers, had normal hearing and normal or corrected-to-normal vision, had no first-degree relatives with autism or schizophrenia and had no history of neurological impairments or psychological disorders, as determined by self-report screening questionnaire. All protocols were approved by the University of Maryland Institutional Review Board.

### Behavioral assessments

#### The STOMP

The STOMP stimuli were two clips (178 s and 130 s) from two feature films. The clips were freely available at <http://www.movieclips.com>. Both clips had several characters and depicted emotions, deception and changes in mental states. In the first clip, a woman is breaking into a man's apartment, while two others watch, when the man returns home. In the second clip, some friends are watching a romantic date

when they risk being discovered. A third clip was initially used, but discarded, because preliminary coding revealed that participants did not make enough mental state attributions to produce sufficient variability. All clips were muted to prevent reliance on language as a mechanism for imputing internal states (cf. Klin, 2000).

Participants completed the task alone on a computer and were told only that they would watch silent movies. The order of the clips was randomized. After each clip, participants answered two questions: first, whether they had seen the movie previously, and second, participants were asked to describe the scene, using 7–10 lines of typed text (~140–170 words).

Participant responses were divided up into clauses by the first author (Chua *et al.*, 2005; Figure 1). Four trained coders then placed each clause into one of five categories: simple physical, physical inference, emotion, goal/intention and belief state. For later analyses, the two external categories (simple physical and physical inference) were collapsed together, and the three internal categories (emotion, goal/intention and belief state) were collapsed together. The STOMP index was computed by taking the number of internal statements (across both clips), dividing by the number of total statements (internal plus external) and multiplying by 100.

Interrater reliability was calculated using Krippendorff's alpha (Hayes and Krippendorff, 2007). Overall reliability was 0.85 across the two broad categories (internal and external) and was 0.78 across the five subcategories, which are excellent levels of agreement (Fleiss, 1981). If two coders disagreed or if an individual coder believed the clause boundary to be incorrect, which occurred on <2% of clauses, a decision was reached after discussion with the first author.

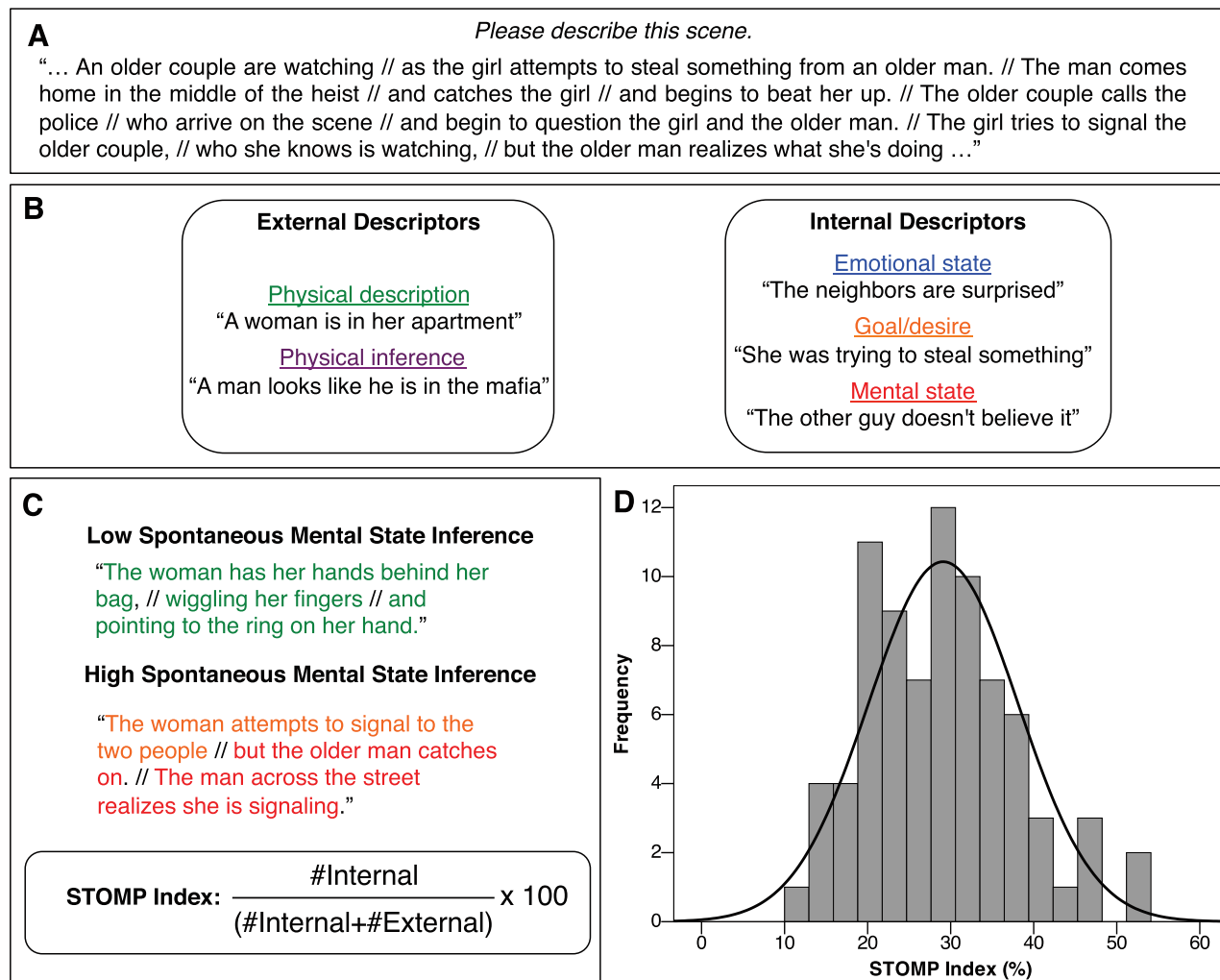
#### Explicit ToM measures

To assess story-based mentalizing, participants completed a task based on the stories of Kinderman and colleagues (1998), although a new set of stories were created because the initial list was too small. Participants listened to 21 stories, and, at the end of each story, evaluated a series of statements as true or false. Statements were of two types: first, statements about mental states of story characters (i.e. ToM), and second, statements based on other facts in the story (i.e. memory). Both sets of statements varied in complexity. First-order questions contained a single clause (e.g. "The story was set in the summer" or "Bryan wanted to marry Amber"). Higher-order questions had up to four levels (e.g. a third-order ToM statement was "Kristin thought that Cody wanted Kent to know the score"). Given the simplicity of the first-order questions, analysis was restricted to higher-order questions (45 memory; 48 ToM). To isolate ToM ability from general attention and comprehension, a composite score was calculated by taking standardized ToM performance and subtracting standardized memory performance.

To assess face-based mentalizing, participants were presented with 36 black-and-white photographs, each depicting the eye region of a face (Mind in the Eyes—Revised Eyes Test; Baron-Cohen *et al.*, 2001a). Participants chose which of four words (e.g. jealous, arrogant, panicked, hateful) best described the mental state of the person in the photograph, receiving one point for each correct answer.

#### Other behavioral assessments

Autistic-like personality traits were assessed using the 50-item autism spectrum quotient (AQ), which is a self-report measure designed for typical adults (Baron-Cohen *et al.*, 2001b). For each item (e.g. "I find social situations easy"), participants answered from 1 (definitely agree) to 4 (definitely disagree), with higher scores indicating more



**Fig. 1** STOMP. (A) An excerpt from a participant's response to a video clip, with the clause breaks inserted by the coder (//). (B) The coding scheme with examples from each coding category. (C) Two participant descriptions of the same scene, varying in number of internal descriptors. These codes are converted into the STOMP index, which is the ratio of internal descriptors to total descriptors. (D) Histogram of STOMP index values for  $N = 80$  typical adult participants. The variable is normally distributed ( $P > 0.1$ ).

autistic-like traits. All analyses with the AQ examined the full-scale measure. Full-scale IQ (FSIQ), verbal IQ (VIQ) and nonverbal IQ (NVIQ) were assessed using the Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman and Kaufman, 2004).

### Procedure

Participants completed all behavioral tasks on a computer in a single visit and completed the STOMP first to prevent bias. Behavioral battery participants who also completed an MRI scan completed the scan within one month of the battery. Participants who did not complete the full ToM battery were scanned immediately after finishing the STOMP. All participants completed the AQ. The participants who contributed ToM battery data and MRI data also completed the KBIT-2.

### Neuroanatomical measures

#### Data acquisition

High-resolution T1-weighted images were acquired using a 12-channel head coil on a single Siemens 3.0-T scanner (MAGNETOM Trio Tim System, Siemens Medical Solutions). The scanning protocol for each participant included one three-dimensional

T1 magnetization-prepared rapid gradient-echo sequence (176 contiguous sagittal slices, voxel size = 1.0 mm × 1.0 mm × 1.0 mm; repetition time/echo time/inversion time = 1900/2.52/900 ms; flip angle = 9°; pixel matrix = 256 × 256). All MRI data ( $n = 38$ ) was free from motion artifact based on visual inspection.

Structural data were analyzed using FreeSurfer's standard automatic volumetric segmentation (surfer.nmr.mgh.harvard.edu). This method is extensively described elsewhere (Dale, et al., 1999; Fischl et al., 2002; Desikan et al., 2006; Fischl, 2012), and its calculations are comparable with manual segmentation (Morey et al., 2009; Lehmann et al., 2010). Analyses for all subjects were completed using FreeSurfer version 5.1.0 using a Linux terminal (RedHat 6.3) on a Macintosh computer (Gronenschild et al., 2012). The T1-weighted images of each participant were compared with a probabilistic atlas, generating new surface maps of gray matter, white matter and pial boundaries. Although reconstruction and thickness calculations are automatized, two independent trained coders compared the new surface maps with the original T1-weighted images, editing that surface's segmentation where necessary. After coder agreement, automatic segmentation was rerun. Cortical thickness was calculated from this finalized data by a FreeSurfer algorithm, which automatically computed the distance between pial and white matter surfaces at each vertex.

### Regions of interest analysis

We identified six regions of interest (ROIs) that have been functionally linked to various facets of mentalizing (Figure 2A): bilateral TPJ, for its role in belief representation (Saxe and Kanwisher, 2003; Young *et al.*, 2010), mPFC, for its involvement in person perception and mental state attribution (Saxe and Powell, 2006; Van Overwalle, 2011), bilateral posterior STS (pSTS), because of its role in processing goals and intentions (Pelphrey *et al.*, 2004; Shultz *et al.*, 2011; Gao *et al.*, 2012), and right inferior frontal gyrus (IFG), because of its involvement in emotional processing and pragmatics (Nakamura *et al.*, 1999; Ishai *et al.*, 2005; Wang *et al.*, 2006; Johnston *et al.*, 2013) and its activation while free-viewing social interactions (Iacoboni *et al.*, 2004). Given evidence that right TPJ and right pSTS may be more selective for belief and intention processing than their left homologues (e.g. Pelphrey *et al.*, 2004; Liu *et al.*, 2009; Donhel *et al.*, 2012), left and right pSTS and TPJ were examined separately.

The TPJ, mPFC and pSTS were defined based on the boundaries from Mills and colleagues (2014) for two reasons: first, the group's methodology for defining boundaries, and second, their extensive characterization of age-related changes in these regions. The boundaries for each region were drawn based on structural and functional priors, and the boundaries for one region took the others into account (e.g. TPJ coordinates were selected because of the region's functional connectivity to the selected mPFC region). Further, the plotted cortical thinning trajectories indicated that, for the age range of the current study's sample, the cortex is still thinning, which confirmed prior work (Gogtay *et al.*, 2004). Mills and colleagues (2014) did not define IFG, so the right IFG ROI was created by combining three regions from the FreeSurfer parcellation atlas (pars triangularis, pars opercularis, pars orbitalis; Desikan *et al.*, 2006). Finally, we selected V1 as a control

region, to ensure that effects were specific to the hypothesized regions. V1 was defined based on a built-in FreeSurfer ROI.

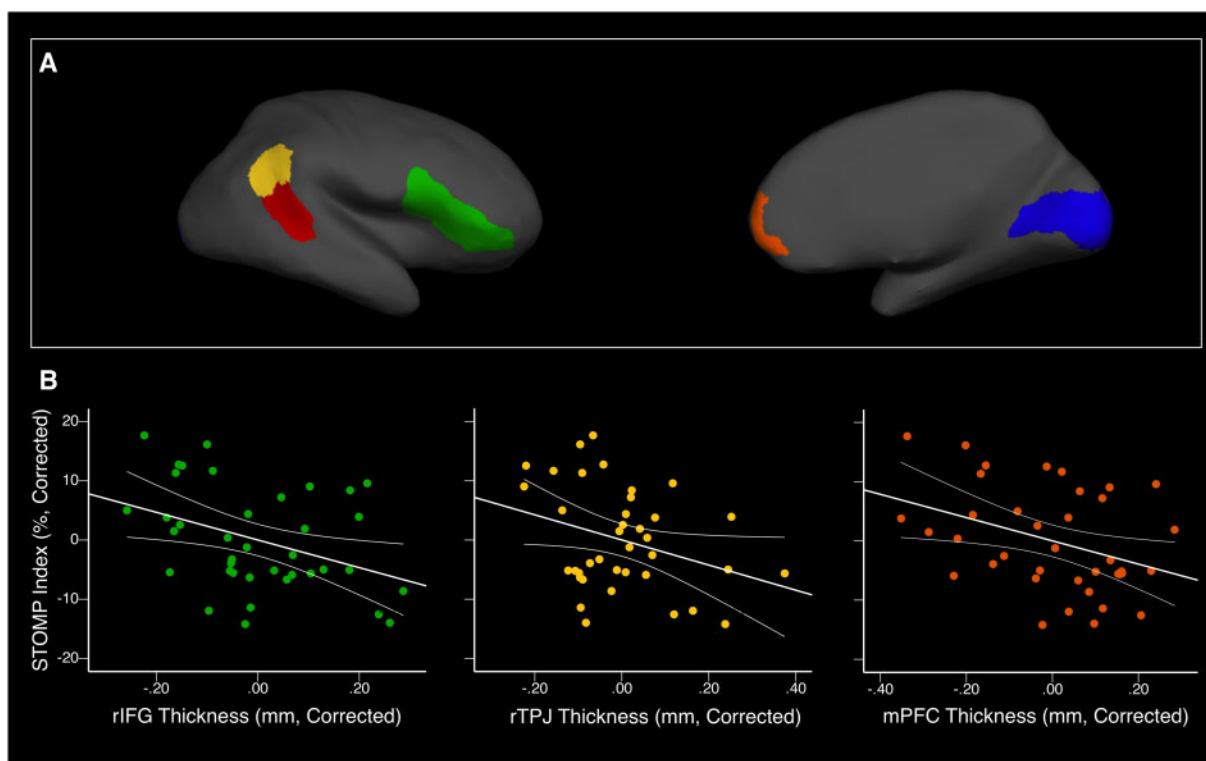
Past research has employed several ways to control for the relation between other neuroanatomical measures and cortical thickness. An extensive meta-analysis indicated the use of ratios can introduce bias and suggested instead controlling for total gray matter volume via regression (Van Petten, 2004). Controlling for total gray matter volume accounts for variability due to a variety of neuroanatomical factors, including surface area, cortical thickness and head size, and we thus controlled for total gray matter volume in our regressions. Further, given potential relationships between age and cortical thickness, all correlations also controlled for age.

## RESULTS

### Behavioral results

For the full STOMP sample, all five subscales and the composite STOMP index showed variability in performance (Table 1). The STOMP index was normally distributed (Shapiro-Wilk = 0.976,  $P = 0.13$ ; Figure 1D). Whether the participant had previously viewed the full movie from which the clip was excerpted had no effect on any of the five subscales or on the STOMP index for that particular clip [ $t(78)s < 1$ ].

For the higher-order ToM story task, scores for the ToM portion (mean = 74.4, s.d. = 10.73) and memory portion (mean = 79.6, s.d. = 10.53) were significantly correlated [ $r(52) = 0.616$ ,  $P < 0.001$ ]. Given these correlations, subsequent analyses of higher-order ToM scores controlled for memory performance. Scores on the Mind in the Eyes ranged from 44% to 92% (mean = 72.9, s.d. = 11.30). For the full STOMP sample, AQ scores ranged from 4 to 32 (mean = 16.16, s.d. = 5.27), with one participant, who completed the



**Fig. 2** ROIs and their relation to STOMP performance. (A) Lateral and medial views of the four right hemisphere ROIs (yellow = TPJ; red = pSTS; green = IFG; orange = mPFC) and the one control region (blue = V1). The ROIs are depicted on the inflated brain of a single participant. The left hemisphere ROIs were homologous to the right, except that left IFG was not an ROI. For statistical analysis, medial regions (mPFC and V1) were collapsed across hemispheres. (B) Partial correlations (mPFC and V1) were collapsed across hemispheres. (B) Partial correlations between cortical thickness and STOMP index scores. All correlations are corrected for age and total gray matter volume. Partial correlations with right IFG and mPFC were significant ( $P < 0.05$ ), and the correlation with right TPJ was marginally significant ( $P < 0.06$ ).

**Table 1** Descriptive statistics for the number of statements in each category ( $N = 80$ )

Statement type	Minimum	Maximum	Mean	s.d.
Simple physical	2	29	14.70	6.03
Physical inference	4	19	10.16	3.75
<b>External</b>	<b>10</b>	<b>42</b>	<b>24.85</b>	<b>6.58</b>
Emotion	0	6	1.60	1.46
Goal/intention	2	18	8.40	2.97
Belief	0	8	2.80	1.60
<b>Internal</b>	<b>4</b>	<b>19</b>	<b>9.98</b>	<b>3.25</b>
<b>STOMP index (% internal)</b>	<b>12.50</b>	<b>52.38</b>	<b>29.11</b>	<b>9.00</b>

Note. Values are summed across both film clips. Bold rows indicate aggregation across statement types. External statement totals are calculated by summing simple physical and physical inference statements, and internal statement totals are calculated by summing emotion, goal/intention and belief statements. The STOMP index represents the number of internal statements divided by the number of internal plus external statements, multiplied by 100.

full behavioral battery, but was not scanned, scoring at the conventional autism cutoff of 32 (Baron-Cohen *et al.*, 2001b).

There was no significant difference between males and females on the STOMP index [ $M_{\text{male}} = 29.51$ ,  $s.d. = 8.80$  vs  $M_{\text{female}} = 28.90$ ,  $s.d. = 9.22$ ,  $t(78) = 0.29$ ,  $P = 0.77$ ]. There was no significant relation between the STOMP index and age [ $r(78) = -0.058$ ,  $P = 0.61$ ], VIQ [ $r(16) = -0.15$ ], NVIQ [ $r(16) = 0.057$ ] or FSIQ [ $r(16) = -0.08$ ]. There was no significant effect of gender on any of the two explicit ToM measures or AQ scores ( $t_s < 1$ ). None of these measures' correlations with age or IQ scales were significant.

For the group that completed the full behavioral battery, STOMP performance was positively correlated with higher-order ToM performance [ $r(52) = 0.28$ ,  $P = 0.038$ ]. STOMP performance was not significantly related to Mind in the Eyes scores [ $r(56) = -0.16$ ,  $P = 0.24$ ]. Mind in the Eyes was not significantly correlated with higher-order ToM [ $r(53) = 0.001$ ,  $P = 0.99$ ]. The AQ was not significantly related to higher-order ToM, Mind in the Eyes or STOMP performance ( $P_s > 0.1$ ).

## Neuroanatomical results

### Cortical thickness correlates of STOMP and AQ scores

For the six social brain ROIs, controlling for age and total gray matter volume, the relations between cortical thickness and the STOMP index were negative and reached significance for two regions (Table 2; Figure 2B). We also examined correlations with AQ scores—a well-established measure of real-world autistic-like personality traits—to confirm that, for this age-group, thinner cortex is associated with improved social skills. For all six ROIs, partial correlations with AQ score were positive, indicating that fewer autistic-like personality traits were associated with thinner cortex.

To determine the specificity of the relation between STOMP performance and cortical thickness, we also examined a control region (V1). The partial correlation with V1 thickness was not significant for either the STOMP index [ $r(34) = -0.11$ ,  $P = 0.51$ ] or for AQ scores [ $r(32) = -0.028$ ,  $P = 0.88$ ].

### Comparison of spontaneous and explicit ToM measures.

For the subsample of participants who completed an MRI scan and the full behavioral battery, none of the partial correlations between ROI thickness and performance on either explicit ToM task (i.e. Mind in the Eyes, higher-order ToM stories) were significant. For this subsample, controlling for age and total gray matter volume, the partial correlations between the STOMP index and the cortical thickness of mPFC and right IFG remained significant [ $r(14) = -0.76$ ,  $P < 0.01$ ;  $r(14) = -0.81$ ,  $P < 0.001$ ]. The partial correlation between mPFC

**Table 2** Relations between cortical thickness and two behavioral measures

Region of interest	STOMP index ( $N = 38$ )	Reverse scored AQ ( $N = 35$ )
mPFC	<b>-0.377*</b> [-0.63, -0.06]	-0.165 [-0.48, 0.19]
Right IFG	<b>-0.387*</b> [-0.63, -0.07]	-0.012 [-0.35, 0.33]
Right TPJ	-0.321 <sup>#</sup> [-0.59, 0.01]	-0.341 <sup>#</sup> [-0.61, 0.002]
Left TPJ	-0.053 [-0.38, 0.28]	-0.071 [-0.40, 0.28]
Right pSTS	-0.084 [-0.40, 0.25]	<b>-0.486**</b> [-0.71, -0.17]
Left pSTS	-0.118 [-0.43, 0.22]	-0.216 [-0.52, 0.14]

Note. Correlations control for age and total gray matter volume. To facilitate comparison with STOMP scores, the AQ is presented as reverse scored so that lower scores indicate more autistic-like traits. Three participants did not complete the AQ. Numbers in brackets represent 95% confidence intervals around the correlation coefficient estimates. Bold values indicate significant correlations. <sup>#</sup> $P < 0.06$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ .

thickness and the STOMP score was significantly greater than corresponding relation between mPFC thickness and Mind in the Eyes score [ $r(14) = -0.004$ ;  $z = -2.30$ ,  $P = 0.019$ ] and was marginally greater than the partial correlation between mPFC thickness and higher-order ToM [ $r(13) = -0.25$ ;  $z = -1.82$ ,  $P = 0.068$ ]. The relation between right IFG thickness and the STOMP index was significantly greater than the corresponding relation between right IFG thickness and higher-order ToM [ $r(13) = 0.040$ ;  $z = -3.26$ ,  $P = 0.001$ ], and was marginally greater than the partial correlation between right IFG thickness and the Mind in the Eyes [ $r(14) = -0.28$ ;  $z = 1.85$ ,  $P = 0.06$ ]. For both regions, full regression models that included all three ToM measures, age and total gray matter volume also indicated that the STOMP significantly predicted cortical thickness [mPFC:  $\beta_{\text{STOMP}} = -0.947$ ,  $t(11) = -4.15$ ,  $P = 0.002$ ; right IFG:  $\beta_{\text{STOMP}} = -0.944$ ,  $t(11) = -4.82$ ,  $P < 0.001$ ], but that the other two ToM measures did not ( $P_s > 0.05$ ). That is, for mPFC and right IFG, the relation between cortical thickness and the STOMP was greater than the correlation between cortical thickness and the explicit ToM tasks.

## DISCUSSION

We examined the relation between brain structural maturation and spontaneous mentalizing abilities using a novel task—the STOMP—that overcomes limitations in extant ToM measures for typical adults. In the STOMP, participants watched silent film clips and gave a spontaneous written description of each clip's events. A composite measure—the STOMP index—was created based on the proportion of a participant's spontaneous statements about characters' internal states. As hypothesized, the STOMP index produced robust individual differences, was normally distributed, produced no ceiling effects and was correlated with conventional ToM measures. Higher scores on the STOMP were significantly correlated with thinner cortex in mPFC and right IFG and marginally correlated with thinner cortex in right TPJ, controlling for age and total gray matter volume. These correlations were significantly stronger for the STOMP than for conventional ToM measures and were not significant between the STOMP and a control region (i.e. V1 of visual cortex).

### Behavioral characterization of spontaneous ToM

The STOMP index was significantly correlated with understanding higher-order ToM stories. The Mind in the Eyes was not related to either task. These three ToM tasks examined different facets of ToM and had different task demands beyond mentalizing (e.g. syntactic skills, memory). Given such disparate features, the lack of correlations

between tasks is perhaps unsurprising. However, the correlation between the STOMP and an explicit story-based mentalizing measure may indicate that the STOMP captures a common element among mentalizing assessments. Further studies with more varied ToM assessments could test this hypothesis.

The lack of correlation between any of the ToM measures and the AQ is surprising (Baron-Cohen *et al.*, 2001a; Voracek and Dressler, 2006). However, although individuals with autism have both high AQ scores and impaired ToM, this relation may not be as robust for typical samples. To better characterize the link between STOMP performance and real-world social functioning, future research could examine larger community samples using more varied measures of social outcomes. Extending the STOMP to atypical populations would also be informative, given evidence from autism that implicit spontaneous ToM differs from explicit belief reasoning and that ToM is relatively less impaired during explicit tasks (Klin, 2000; Senju *et al.*, 2009).

Although this study represents an important first step in more richly characterizing spontaneous mentalizing, there are some limitations. The STOMP score is based on participant report. A participant could richly consider the mental states of the clip's characters but not report this in their written description. However, the assumption behind this task, and similar measures (e.g. Heider and Simmel, 1944), is that participants report on what they perceived to be most important. Thus, the STOMP index also captures how much participants believe internal states are crucial in describing social interactions, and perhaps this tendency, more than the raw capacity for mentalizing, can explain success or failure in social situations (Klin, 2000). Further, in spite of the STOMP's dynamic and naturalistic stimuli, participants were not engaged in real social interaction, which may alter mentalizing (Redcay *et al.*, 2010; Champagne-Lavau and Moreau, 2013; Schilbach *et al.*, 2013).

### Spontaneous ToM and cortical thickness

Our a priori ROIs were selected based on their functional involvement in various aspects of mentalizing. For example, right TPJ has been identified as a core region for processing beliefs (Dohnel *et al.*, 2012; Koster-Hale and Saxe, 2013), mPFC has been implicated in more general social perception (Van Overwalle, 2011), pSTS has been linked to action perception (Vander Wyk *et al.*, 2009) and right IFG has been implicated in processing emotional expressions (Nakamura *et al.*, 1999; Ishai *et al.*, 2005) and free-viewing social interaction (Iacoboni *et al.*, 2004).

In this study, correlations between cortical thickness and STOMP scores were stronger for right *vs* left hemisphere ROIs, supporting functional work suggesting that right regions play a greater role in social processing (e.g. Pelphrey *et al.*, 2004; Dohnel *et al.*, 2012). Specifically, three regions showed a structural relationship to spontaneous mentalizing—mPFC, right IFG and right TPJ—which suggests that this study's naturalistic mentalizing task, the STOMP, captures multiple components of ToM that are reflected in brain structure. This is consistent with the hypothesis that everyday mentalizing relies on processing several simultaneous types of information (e.g. faces, body language, cognitive representations).

Right pSTS thickness was not related to STOMP performance. Although the pSTS is involved in intention understanding (e.g. Vander Wyk *et al.*, 2009), it is less strongly linked to belief processing (e.g. Saxe, 2009) and is also involved in human action perception (e.g. biological motion; Thompson *et al.*, 2005; Saygin, 2007). However, the STOMP index considers descriptions of individuals' physical actions to be external, not internal, statements.

In contrast to the STOMP, the tasks designed to assess specific subcomponents of mentalizing (i.e. inferring mental states from the eyes and parsing stories about higher-order belief states) did not show structural relations with any ROIs. Although the sample size was smaller for this comparison ( $n=18$ ), when restricted to this sample, the STOMP captured significantly more structural variability than the traditional explicit ToM tasks. Thus, although these regions may be functionally involved during ToM processing, such a role is not reflected in cortical structure.

The present study does not establish the exact mechanism linking cortical thickness to spontaneous ToM. One possibility is that thinner cortex and higher STOMP scores develop independently because of separate genetic influences (cf. Kremen *et al.*, 2010). Another possibility is that a thinner cortex, perhaps due to a genetic propensity (Panizzon *et al.*, 2009), directly promotes efficient social processing, and thus individuals with this advantage have higher STOMP scores. A stronger hypothesis would be that cortical thickness both affects, and is also affected by, social experience. The causal relation between environmental experiences and cortical thickness has been demonstrated in animals (e.g. Pappas *et al.*, 1979; Kolb and Elliot, 1987) and humans (Engvig *et al.*, 2010). For the present study, perhaps extensive exposure to general social stimuli, or mentalizing specifically, affects the cellular processes, such as synaptic pruning (Huttenlocher and Dabholkar, 1997) or myelination (Sowell *et al.*, 2004), that result in developmental cortical thinning (Johnson, 2011). With improved social processing, there may even be a cascading effect; individuals with extensive social experiences develop better social skills and then find social experiences more rewarding and seek them out further (Klin *et al.*, 2003). Thus, participants with more mature cortical thinning may have years of everyday mentalizing experiences that are reflected in their propensity to focus on internal states during the STOMP. A longitudinal study could parse apart questions of causality and permit the examination of rates of change in cortical thickness (Shaw *et al.*, 2006). Modifications to the STOMP to make the task appropriate for younger age-groups would also allow for examination of the relation between spontaneous ToM and cortical thickness during periods of robust development in ToM (e.g. early and middle childhood; Wellman and Liu, 2004; Gweon *et al.*, 2012).

The potential link between individual differences in experiences, brain structure and spontaneous ToM may also account for the lack of a correlation between the explicit ToM tasks and cortical thickness. If cortical thinning reflects a history of real-world social experiences, then cortical thickness will best correlate with measures that capture naturalistic mentalizing constructs. That is, more spontaneous measures better capture a history of social processing, and thus are more strongly linked to neuroanatomy. This theory is lent support by the present study's correlations between AQ scores and cortical thickness as the AQ attempts to tap into real-world behaviors and functioning.

The current work is limited by the fact that, although the structural ROIs were based on prior functional neuroimaging studies, functional data were not collected for these participants. Thus, on an individual level, the boundaries of the selected ROIs may not map exactly on to the regions engaged by these participants during mentalizing. Establishing a link between function, structure and behavioral performance is an important future direction, building on this study's important contribution to understanding the intersection of brain development, laboratory-based ToM measures and everyday mentalizing.

### Conflict of Interest

None declared.

## REFERENCES

- Achim, A.M., Guitton, M., Jackson, P.L., Boutin, A., Monetta, L. (2013). On what ground do we mentalize? Characteristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment*, 25(1), 117.
- Adams, R.B., Jr, Rule, N.O., Franklin, R.G., Jr, et al. (2010). Cross-cultural reading the mind in the eyes: an fMRI investigation. *Journal of Cognitive Neuroscience*, 22(1), 97–108.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I. (2001a). The “Reading the Mind in the Eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241–51.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. (2001b). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17.
- Castelli, F., Frith, C., Happé, F., Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, 125(8), 1839–49.
- Castelli, 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(3), 314–25.
- Champagne-Lavau, M., Moreau, N. (2013). Characteristics of current tasks that contribute to mentalizing judgments: does the engagement of the participants in the social interaction matter? Comment on Achim et al. (2013). *Psychological Assessment*, 25(4), 1404–6.
- Chua, H.F., Leu, J., Nisbett, R.E. (2005). Culture and diverging views of social events. *Personality and Social Psychology Bulletin*, 31(7), 925–34.
- Churchwell, J.C., Yurgelun-Todd, D.A. (2013). Age-related changes in insula cortical thickness and impulsivity: significance for emotional development and decision-making. *Developmental Cognitive Neuroscience*, 6, 80–6.
- Dale, A.M., Fischl, B., Sereno, M.I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9, 179–94.
- Desikan, R.S., Segonne, F., Fischl, B., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, 31, 968–80.
- Dodell-Feder, D., Koster-Hale, J., Bedny, M., Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, 55(2), 705–12.
- Dodell-Feder, D., Lincoln, S.H., Coulson, J.P., Hooker, C.I. (2013). Using fiction to assess mental state understanding: a new task for assessing theory of mind in adults. *PLoS One*, 8(11), e81279.
- Donhel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., Sommer, M. (2012). Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *Neuroimage*, 60, 1652–61.
- Ducharme, S., Albaugh, M.D., Hudziak, J.J., et al. (forthcoming). Anxious/depressed symptoms are linked to right ventromedial prefrontal cortical thickness maturation in healthy children and young adults. *Cerebral Cortex*.
- Dziobek, I., Fleck, S., Kalbe, E., et al. (2006a). Introducing MASC: a movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, 36(5), 623–36.
- Dziobek, I., Fleck, S., Rogers, K., Wolf, O.T., Convit, A. (2006b). The ‘amygdala theory of autism’ revisited: linking structure to behavior. *Neuropsychologia*, 44, 1891–9.
- Engvig, A., Fjell, A.M., Westlye, L.T., et al. (2010). Effects of memory training on cortical thickness in the elderly. *Neuroimage*, 52(4), 1667–76.
- Fischl, B. (2012). FreeSurfer. *Neuroimage*, 62, 774–81.
- Fischl, B., Salat, D.H., Busa, E., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33, 341–55.
- Fleiss, J.L. (1981). *Statistical Methods for Rates and Proportions* 2nd edn. New York: Wiley.
- Gao, T., Scholl, B.J., McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *The Journal of Neuroscience*, 32(41), 14276–80.
- Gogtay, N., Giedd, J.N., Lusk, L., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 8174–9.
- Gronenschild, E.H., Habets, P., Jacobs, H.I., et al. (2012). The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. *PLoS One*, 7(6), e38234.
- Gweon, H., Dodell-Feder, D., Bedny, M., Saxe, R. (2012). Theory of Mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Development*, 83(6), 1853–68.
- Hayes, A.F., Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.
- Heider, F., Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243–59.
- Huttenlocher, P.R., Dabholkar, A.S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, 387(2), 167–78.
- Jacoboni, M., Lieberman, M.D., Knowlton, B.J., et al. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage*, 21(3), 1167–73.
- Ishai, A., Schmidt, C.F., Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin*, 67(1), 87–93.
- Johnson, M.H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1(1), 7–21.
- Johnston, P., Mayes, A., Hughes, M., Young, A.W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex*, 49(9), 2462–72.
- Jung, R.E., Segall, J.M., Jeremy Bockholt, H., et al. (2010). Neuroanatomy of creativity. *Human Brain Mapping*, 31(3), 398–409.
- Kaufman, A.S., Kaufman, N.I. (2004). *Kaufman Brief Intelligence Test* 2nd edn (KBIT-2). Bloomington, MN: Pearson, Inc.
- Keysar, B., Lin, S., Barr, D.J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41.
- Kharitonova, M., Martin, R.E., Gabrieli, J.D., Sheridan, M.A. (2013). Cortical gray-matter thinning is associated with age-related improvements on executive function tasks. *Developmental Cognitive Neuroscience*, 6, 61–71.
- Kinderman, P., Dunbar, R., Bentall, R.P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, 89(2), 191–204.
- Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: the Social Attribution Task. *Journal of Child Psychology and Psychiatry*, 41(7), 831–46.
- Klin, A., Jones, W. (2006). Attributing social and physical meaning to ambiguous visual displays in individuals with higher-functioning autism spectrum disorders. *Brain and Cognition*, 61(1), 40–53.
- Klin, A., Jones, W., Schultz, R., Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences*, 358, 345–60.
- Kolb, B., Elliott, W. (1987). Recovery from early cortical damage in rats. II. Effects of experience on anatomy and behavior following frontal lesions at 1 or 5 days of age. *Behavioural Brain Research*, 26(1), 47–56.
- Koster-Hale, J., Saxe, R. (2013). Functional neuroimaging of theory of mind. In: Baron-Cohen, S., Lombardo, M., Tager-Flusberg, H., Cohen, D., editors. *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*. New York: Oxford University Press, pp. 132–63.
- Kremen, W.S., O’Brien, R.C., Panizzon, M.S., et al. (2010). Salivary cortisol and prefrontal cortical thickness in middle-aged men: a twin study. *Neuroimage*, 53(3), 1093–102.
- Lehmann, M., Douiri, A., Kim, L.G., et al. (2010). Atrophy patterns in Alzheimer’s disease and semantic dementia: a comparison of FreeSurfer and manual volumetric measurements. *Neuroimage*, 49, 2264–74.
- Lewis, P.A., Rezaie, R., Brown, R., Roberts, N., Dunbar, R.I. (2011). Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage*, 57(4), 1624–9.
- Liu, D., Meltzoff, A.N., Wellman, H.M. (2009). Neural correlates of belief- and desire-reasoning. *Child Development*, 80, 1162–71.
- Lu, L.H., Dapretto, M., O’Hare, E.D., et al. (2009). Relationships between brain activation and brain structure in normally developing children. *Cerebral Cortex*, 19(11), 2595–604.
- McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): a clinical test of social perception. *Disability and Rehabilitation*, 28(24), 1529–42.
- McDonald, S., Flanagan, S., Martin, I., Saunders, C. (2004). The ecological validity of TASIT: a test of social perception. *Neuropsychological Rehabilitation*, 14(3), 285–302.
- McDonald, S., Flanagan, S., Rollins, J., Kinch, J. (2003). TASIT: a new clinical tool for assessing social perception after traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, 18(3), 219–38.
- Mills, K.L., Lalonde, F., Clasen, L.S., Giedd, J.N., Blakemore, S.J. (2014). Developmental changes in the social brain in late childhood and adolescence. *Social Cognitive and Affective Neuroscience*, 9, 123–31.
- Montag, C., Ehrlich, A., Neuhaus, K., et al. (2010). Theory of mind impairments in euthymic bipolar patients. *Journal of Affective Disorders*, 123(1), 264–9.
- Morey, R.A., Petty, C.M., Xu, Y., et al. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage*, 45, 855–66.
- Nakamura, K., Kawashima, R., Ito, K., et al. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *Journal of Neurophysiology*, 82(3), 1610–4.
- Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., et al. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral Cortex*, 19(11), 2728–35.
- Pappas, C.T., Diamond, M.C., Johnson, R.E. (1979). Morphological changes in the cerebral cortex of rats with altered levels of ovarian hormones. *Behavioral and Neural Biology*, 26(3), 298–310.
- Pelphrey, K.A., Morris, J.P., McCarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, 16(10), 1706–16.
- Preißler, S., Dziobek, I., Ritter, K., Heekeren, H.R., Roepke, S. (2010). Social cognition in borderline personality disorder: evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience*, 4, 182.

- Redcay, E., Dodell-Feder, D., Pearrow, M.J., et al. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage*, 50(4), 1639–47.
- Rice, K., Viscomi, B., Riggins, T., Redcay, E. (2013). Amygdala volume linked to individual differences in mental state inference in early childhood and adulthood. *Developmental Cognitive Neuroscience*, 8, 153–63.
- Samson, D., Apperly, I.A., Braithwaite, J.J., Andrews, B.J., Bodley Scott, S.E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1255.
- Saygin, A.P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, 130(9), 2452–61.
- Saxe, R. (2009). Theory of mind (neural basis). In: Banks, W., editor. *Encyclopedia of Consciousness*. Oxford: Elsevier, pp. 401–9.
- Saxe, R., Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in theory of mind. *Neuroimage*, 19, 1835–42.
- Saxe, R., Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*, 17, 692–9.
- Schneider, D., Lam, R., Bayliss, A.P., Dux, P.E. (2012). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, 23(8), 842–7.
- Schilbach, L., Timmermans, B., Reddy, V., et al. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393–414.
- Scholl, B.J., Tremoulet, P.D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8), 299–309.
- Senju, A., Southgate, V., White, S., Frith, U. (2009). Mindblind Eyes: an absence of spontaneous theory of mind in Asperger Syndrome. *Science*, 325, 883–5.
- Shaw, P., Gilliam, M., Liverpool, M., et al. (2011). Cortical development in typically developing children with symptoms of hyperactivity and impulsivity: support for a dimensional view of attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, 168(2), 143.
- Shaw, P., Greenstein, D., Lerch, J., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, 440(7084), 676–9.
- Shultz, S., Lee, S.M., Pelphrey, K., McCarthy, G. (2011). The posterior superior temporal sulcus is sensitive to the outcome of human and non-human goal-directed actions. *Social Cognitive and Affective Neuroscience*, 6(5), 602–11.
- Sommer, M., Döhl, K., Sodian, B., et al. (2007). Neural correlates of true and false belief reasoning. *Neuroimage*, 35(3), 1378–84.
- Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience*, 24(38), 8223–31.
- Spiers, H.J., Maguire, E.A. (2006). Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage*, 31(4), 1826–40.
- Spunt, R.P., Lieberman, M.D. (2012). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *Neuroimage*, 59(3), 3050–9.
- Tavares, P., Lawrence, A.D., Barnard, P.J. (2008). Paying attention to social meaning: an fMRI study. *Cerebral Cortex*, 18(8), 1876–85.
- Thompson, J.C., Clarke, M., Stewart, T., Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *The Journal of Neuroscience*, 25(39), 9059–66.
- Wagner, D.D., Kelley, W.M., Heatherton, T.F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex*, 21(12), 2788–96.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain*, 129(4), 932–43.
- Wellman, H.M., Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, 75(2), 523–41.
- Wendelken, C., O'Hare, E.D., Whitaker, K.J., Ferrer, E., Bunge, S.A. (2011). Increased functional selectivity over development in rostralateral prefrontal cortex. *The Journal of Neuroscience*, 31(47), 17260–8.
- Van Overwalle, F. (2011). A dissociation between social mentalizing and general reasoning. *Neuroimage*, 54(2), 1589–99.
- Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: Review and meta-analysis. *Neuropsychologia*, 42, 1394–413.
- Vander Wyk, B.C., Hudac, C.M., Carter, E.J., Sobel, D.M., Pelphrey, K.A. (2009). Action understanding in the superior temporal sulcus region. *Psychological Science*, 20(6), 771–7.
- Voracek, M., Dressler, S.G. (2006). Lack of correlation between digit ratio (2D: 4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences*, 41(8), 1481–91.
- Young, L., Dodell-Feder, D., Saxe, R. (2010). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia*, 48(9), 2658–64.