



The Dynamic Differences between Self- and Other-Oriented Mental Inferences: An ERP Study on a False-Belief Task

Xieshun Wang¹ · Yanjie Su¹ · Min Hong¹

Accepted: 4 April 2021 / Published online: 27 April 2021
© The Psychonomic Society, Inc. 2021, corrected publication 2021

Abstract

Theory of mind (ToM) refers to the ability to infer others' mental states. In our everyday lives, we need to interact constantly and appropriately with others. Not only is ToM involved in understanding others' mental states (other-oriented mental inferences), but it also helps to keep our own mental states (self-oriented mental inferences). In this study, we designed a false-belief task using event-related potential (ERP) measures to investigate the dynamic differences between the two types of mental inferences. In the false-belief task, participants were prompted with a cue to attribute a belief to either themselves (self-oriented) or another person (other-oriented). Results showed that other-oriented false-belief attribution elicited a larger **late positive component** (LPC; 320~440 ms post-cue onset) than the one elicited by self-oriented attribution at bilateral parieto-occipital electrodes, and also a larger **late negative component** (LNC; 380~500 ms post-cue onset) at frontal-central electrodes. In addition, the difference in amplitude of LPC between the two types of false-belief attribution was positively correlated with self-reported autistic traits measured by the Autism Spectrum Quotient (AQ). Based on prior literature, we propose that the LPC and LNC reflect the processes of self-other distinction and conflict control, respectively. Moreover, the difficulty in distinguishing other from the self is related to one's degree of autistic traits.

Keywords Theory of mind · Self-oriented · Other-oriented · Autistic traits · Event-related potential

Introduction

Theory of mind (ToM) or mentalizing refers to the ability to understand mental states, such as desires, beliefs, or intentions in social interactions (Premack & Woodruff, 1978; Wimmer & Perner, 1983). ToM often is measured via a false-belief paradigm. Converging evidence across various cultures has demonstrated that the ability to understand another person's false-belief emerges at approximately 4 to 5 years old (Wellman et al., 2001). The complexity of ToM and its associated brain networks then continue to develop throughout adolescence and into early adulthood (Kilford et al., 2016). Impairment in ToM is related to difficulties in social communication, as evident from studies in individuals with autism spectrum disorder (ASD) (Baron-Cohen et al., 1985; Baron-

Cohen et al., 2001a; Brent et al., 2004; Hillier & Allinson, 2002; Rajendran & Mitchell, 2007; White et al., 2009).

In social interactions, individuals are required to process concurrently the constantly changing social information from both themselves and others (de Guzman et al., 2016). Not only does ToM matter for the process of inferring others' mental states, but it also matters for the process of keeping our own mental states intact (Bradford et al., 2019). In other words, other- or self-oriented attribution of mental states should be made according to the needs in social interactions (Crone & Fuligni, 2020). However, most prior studies have mainly focused on other-oriented mental attribution. Therefore, results have been ambiguous about the differences between self- and other-oriented mental attribution in terms of underlying processing mechanisms.

Using a computerized false-belief paradigm (self/other differentiation task), Bradford and colleagues made some preliminary explorations on this question (Bradford et al., 2015; Bradford et al., 2019; Bradford, Hukker, et al., 2018a; Bradford, Jentsch, et al., 2018b). The self/other differentiation task was designed mainly **to investigate whether a perspective-shifting would influence the processing of belief**

✉ Yanjie Su
yjsu@pku.edu.cn

¹ School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, China

attribution, whereas the orientation of belief attribution was included as a secondary variable. Specifically, the self/other differentiation task was modified based on a classic unexpected contents paradigm (Perner et al., 1987), which includes three stages. In the Dilemma stage, participants were asked to determine where they or someone else would look for a specific object (e.g., pencil). Then, participants were required to select an appropriate container (e.g., pencil-box) from three alternatives. In the Contents Revelation stage, the contents in the selected container (e.g., pencil-box) were shown to participants, in which the contents could be either expected (e.g., pencil) or unexpected (e.g., radish). This manipulation was used to create a true- or false-belief scenario, respectively. Finally, in the Probe stage, participants were asked to determine what they thought would be in the container (self-oriented attribution) or what someone else thought would be in the container (other-oriented attribution) before the revelation. Conditions in which participants made judgments for different agents (e.g., from self to other) in the Dilemma and the Probe stages were considered involving a perspective-shifting; conditions of judgments for the same agents (e.g., from self to self) were considered no perspective-shifting involved. By using the self/other differentiation task, these studies showed that participants in other-oriented attribution performed slower and less accurately than those in self-oriented attribution, but this effect was much stronger in the perspective-shifting condition than in the no-perspective-shifting condition.

Although aspects of ToM processing have been considered in the self/other differentiation task, the manipulation of perspective-shifting contributes to the task's complexity, because it requires participants 1) to anchor to one perspective, and 2) to make efforts to switch into other's perspective. Therefore, the manipulation of perspective-shifting might introduce even more domain-general processes, such as attention and/or set-shifting into mental inferences, which would constitute confounding variables to the comparative relationship between self- and other-oriented mental inferences. Such an assumption could be supported by a study using the self/other differentiation task to examine the differences in ToM between neurotypical individuals and individuals with ASD (Bradford, Hukker, et al., 2018a). Of note, results did not show any group difference in either self-oriented attribution or other-oriented belief attribution; however, the overall responses across conditions were slower and less accurate in the ASD group than those in the neurotypical group. Critically, individuals with ASD often suffer from deficits in attention (Louzolo et al., 2017; Mannion & Leader, 2013) and/or executive function (Cremone-Caira et al., 2019; Friedman & Sterling, 2019; Johnston et al., 2019). Hence, these findings warrant further empirical evidence to clarify the differences between self- and other-oriented mental inferences and to identify underlying mechanisms.

In the present study, we designed a computerized false-belief task, which was inspired by the classic unexpected location paradigm (Baron-Cohen et al., 1985). Compared with prior studies, the present task focused more on the differences between self- and other-oriented belief attribution. In our false-belief task, both participants and an on-screen character were allowed to see two boxes at the very beginning. A red ball was situated inside one of the boxes. The red ball could be transferred to another box later during the experiment. To introduce a false-belief scenario, in half of the experimental trials, the on-screen character would be blind-folded with a black rectangle “cloth” during which the ball was being transferred to another box. In this case, only participants (but not the on-screen character) would be able to see the final location of the red ball, whereas the on-screen character would believe that the ball was still in the initial box. **Similar to the self/other differentiation task, we presented a cue question at the end of the task to prompt participants to identify their own belief while ignoring other's (i.e., self-oriented attribution) or to identify other's belief while ignoring theirs (i.e., other-oriented attribution).** In this way, we could exclude, as much as possible, the potential confounding variables related to attention and/or set-shifting and then provide a joint validation for the conclusions made by past studies using the self/other differentiation task. The present study also employed event-related potential (ERP) measures to detect subtle differences in time course between self- and other-oriented belief attribution.

To our knowledge, **little or no study has employed ERPs to compare the processing of self- and other-oriented belief attribution**, except for Bradford et al.'s (2019), which discussed this question under the context of perspective-shifting. Using the self/other differentiation task, Bradford et al. (2019) found an interaction between perspective-shifting and the orientation of belief attribution on both the right parieto-occipital scalp and the frontal-central scalp. These results were consistent with the conclusions drawn from prior quantitative meta-reviews that identified a core-network for ToM across multiple tasks and stimulus formats, i.e., the bilateral temporoparietal junction (TPJ) and the medial prefrontal cortex (mPFC) (Mar, 2011; Schurz et al., 2014). However, Bradford et al. (2019) did not delve into the specific processing mechanism revealed by these findings due to the complicated task design that brought about several multivariable interactions. **Based on prior literature, the TPJ, especially in the right hemisphere, is a cortical area associated with self-other distinction in mentalizing** (Decety & Lamm, 2007; Steinbeis, 2016), while **the mPFC reflects an integration of self- and other-oriented processes** (Crone & Fuligni, 2020), such as controlling conflicts between self and others (Brass et al., 2009; Ramsey, 2018). Therefore, **the present study focused on ERP components on both the bilateral posterior scalp and the frontal-central scalp.**

Previous studies demonstrated that people would usually adopt egocentric anchoring-and-adjustment during mentalizing. This implies that people would adopt their own experience as an “anchor” to infer the mental states of other people and then perform a series of adjustments to eliminate the dissimilarities between self and others (Tamir et al., 2010; Tamir & Mitchell, 2013). These findings suggest that engaging in other-oriented belief attribution would be more challenging than engaging in self-oriented attribution. Yet, it is still unclear whether the difficulty occurs during the self-other distinction phase or the controlling phase. In this study, we investigated this issue by examining the differences in ERPs between self- and other-oriented belief attribution on the bilateral posterior scalp and the frontal-central scalp. In exploratory analyses, we also evaluated whether and how the differences between the two types of belief attribution are correlated with autistic traits measured by the Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001b). As suggested that individuals with ASD are characterized by egocentricity in mentalizing (Frith & de Vignemont, 2005; Lombardo & Baron-Cohen, 2011; White et al., 2006; White & White, 2013), we predicted a relationship between differences of self-oriented belief attribution versus other-oriented belief attribution and the degree of one’s autistic traits.

Method

Participants

Using G*Power 3.1.9.2 software (Faul et al., 2007), we determined a priori minimum sample size of 30 for a 2×2 repeated-measures ANOVA. The alpha level was set at 0.05, power at 90%, with an anticipated medium effect size ($f = 0.25$; Cohen, 1988). A total of 37 neurotypical adults were first recruited. All participants had normal or corrected-to-normal vision and provided informed consent before the experiment. Data from four participants were excluded from the final analysis due to excessive artifacts in EEG signals (>50% trials in any one condition). There were 33 participants included in our final analyses (18–25 years; $M = 20.9$, $SD = 1.6$; 20 females).

False-belief Task

Figure 1 illustrates the procedure of our false-belief task. Each trial began with a scene whereby a cartoon character was facing two boxes (labeled “F” and “J,” respectively), and a red ball was situated inside one of the boxes (inside box F for half of the trials and inside box J for the other half). The character’s eyes were then blind-folded with a black rectangle during which the ball was being transferred to another box. Participants were informed in advance that the character

would not be able to see the ball transferring when its eyes were blind-folded, and therefore the character would believe that the ball was still in the initial box. Next, two different scenes were designed to create a true- and a false-belief condition, respectively. In the true-belief condition, the black rectangle was removed from the character’s eyes after the ball was being transferred to another box, so the character would be able to see the final position of the ball. In the false-belief condition, the black rectangle was not being removed from the character’s eyes until the box was covered by a lid. In this case, the character would believe that the ball was still in the initial box. In other words, in the true-belief condition, participants and the character would hold a congruent belief, whereas in the false-belief condition, they would hold an incongruent belief. Finally, we presented a cue question and asked participants to identify either their own thought (“Do you think the ball is in box F or box J?”) or the character’s thought (“Does he think the ball is in box F or box J?”) about the final position of the ball. The cue question would remain on the screen for 2,000 ms before the next trial began. All participants would be required to respond within this time period. Apart from this, as Fig. 1 shows, all the other scenes remained a duration of 1,000 ms before the next scene appeared.

In the false-belief task, there were a total of 8 ($2 \times 2 \times 2$) conditions: Initial position of the ball (box F or box J); Congruency (Congruent or Incongruent) of belief between participants and the character; Agent (Self or Other), the person whom participants would refer to when determining the belief for the cue question. In the experiment, there were a total of 288 trials by 3 blocks. Each type of trial was presented 12 times in a randomized order in each block. Before the experiment, participants were required to perform 24 practice trials to ensure that they fully understand the procedure. The whole experiment lasted approximately 45 min. As the initial position of the ball was not the main concern in this study, in the final analyses, the initial eight conditions were collapsed into four (72 trials per condition) and formed a 2×2 design: Congruency (Congruent vs. Incongruent) \times Agent (Self vs. Other). Therefore, the Self-Incongruent and Other-Incongruent conditions would involve a false-belief scenario, whereas the Self-Congruent and Other-Congruent conditions would involve a true-belief scenario.

Autism Spectrum Quotient

The AQ (Baron-Cohen et al., 2001b) is a self-report screening questionnaire assessing autistic traits for adults. It consists of 50 items in 5 domains with 10 items in each: social skill, attention switching, attention to detail, communication, and imagination. Participants rated each item on a scale of 1 to 4: 1 (definitely agree), 2 (slightly agree), 3 (slightly disagree), and 4 (definitely disagree). The rating scores were converted to a dichotomous outcome (0 or 1). The sum of dichotomous

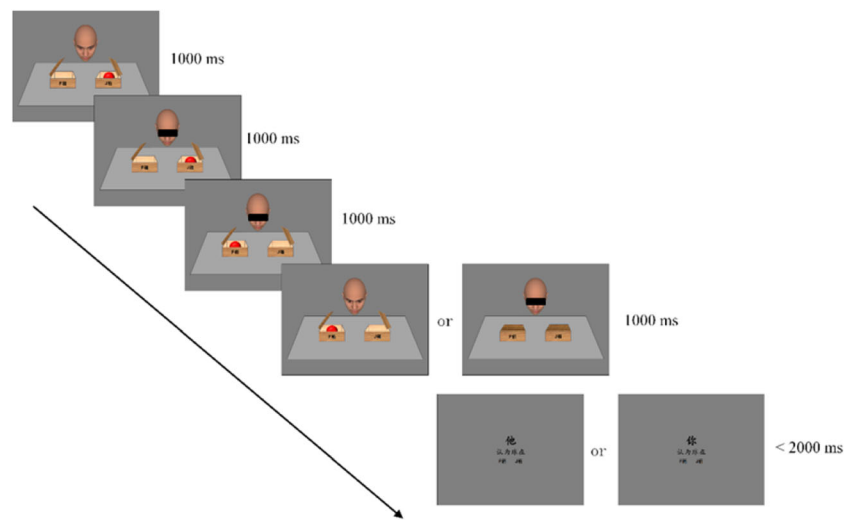


Fig. 1 Schematic for the procedure of false-belief task

scores for each item served as the index for the degree of one's autistic traits. The present study employed a Chinese version of AQ (Poon et al., 2020). The internal consistency reliability (Cronbach's alpha) was 0.73.

EEG Recording and Analyses

The EEG was recorded using a 128-channel Geodesic Sensor Net (Electrical Geodesics Inc., EGI), which is a lightweight elastic thread structure containing Ag/AgCl electrodes housed in a synthetic sponge on a pedestal. The Geodesic Sensor Net was soaked in a KCl solution for approximately 20 min to render the electrodes conductive. The online reference was the Cz electrode placed on the scalp vertex and re-referenced to the average of the two mastoids in offline analyses. The impedance of each electrode was kept below 50 k Ω . This impedance level has been suggested sufficient for a quality recording of EEG as introduced in the Technical Manual of EGI. The recorded EEG signals were amplified by the EGI Net Amps amplifier and continuously digitized at 1,000 Hz. The offline analyses were conducted with the EEGLAB toolbox (version 14.1.2) based on MATLAB 2016b software (The MathWorks). The EEG data were first downsampled to 500 Hz and filtered with a bandpass of 0.1–30 Hz. Then, an independent component analysis (ICA; Stone, 2002) with a threshold individualized to each participant was performed to correct the ocular artifacts from eye blinks or horizontal movements. The corrected EEG data were segmented into epochs beginning 200 ms before the onset of the cue question in the false-belief task and continuing to 1000 ms after the onset. The 200-ms pre-onset interval was chosen for baseline correction. Finally, the epochs with amplitudes exceeding ± 100 μ V as well as incorrect responses were excluded (the overall rejection rate was 17.6%). The remaining epochs for each condition were averaged to produce the grand mean waveforms.

Based on a visual inspection and the prior relevant ERP study (Bradford et al., 2019), we identified a late positive component (LPC) on the bilateral parieto-occipital scalp and a late negative component (LNC) on the frontal-central scalp that revealed the differences between the two false-belief conditions (Other-Incongruent vs. Self-Incongruent). The time windows and electrodes of interest for each ERP component were determined by the *pop_comperp* function in the EEGLAB toolbox. The *pop_comperp* function helped to compare ERP amplitudes of the two conditions in each electrode and highlight the time windows in which the differences between the two conditions are significant ($p < 0.05$). In this way, the time window of LPC was set between 320–440 ms, and the time window of LNC was 380–500 ms. We selected eight representative electrodes for LNC (E16, E11, E18, E19, E12, E10, E4, and E5) and eight representative electrodes for LPC (E60, E66, E67, and E71 in the left hemisphere; E85, E84, E77, and E76 in the right hemisphere). These representative electrodes indicate the loci of maximal difference effects and are marked with white dots in Fig. 2. Furthermore, broadly consistent with a prior ERP study (Wang et al., 2008), we also found two early components (N100: 80–140 ms; N250: 200–370 ms), revealing the differences between true- and false-belief conditions in the electrodes of interest for LNC.

Results

Behavioral results

Trials in the false-belief task with an incorrect response or with a response time (RT) exceeding ± 3 SDs from the grand mean were all removed from the final analyses. The means of RT and accuracy (ACC) in each condition are shown in Table 1. Two Congruency (Congruent vs. Incongruent) \times

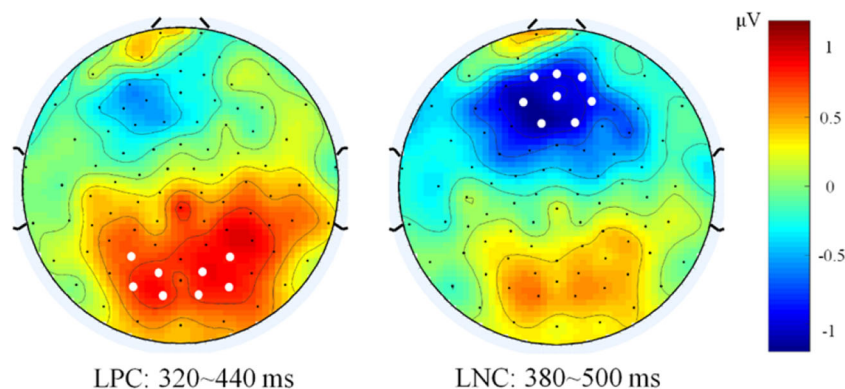


Fig. 2 Topographical maps of scalp voltage showing the differences between the two conditions involving a false-belief scenario (Other-Incongruent minus Self-Incongruent). The electrodes of interest for each ERP component are marked with white dots

Agent (Self vs. Other) repeated-measures ANOVAs were performed for RT and ACC, respectively.

The ANOVA on RT revealed a significant effect of Congruency, $F(1, 32) = 213.51, p < 0.001, \eta_p^2 = 0.87$, and a marginal effect of Agent, $F(1, 32) = 3.81, p = 0.06, \eta_p^2 = 0.11$. However, the two-way interaction was not significant, $F(1, 32) = 0.41, p = 0.53, \eta_p^2 = 0.01$. The ANOVA on ACC showed that both the effect of Congruency, $F(1, 32) = 20.77, p < 0.001, \eta_p^2 = 0.39$, and the effect of Agent, $F(1, 32) = 4.79, p = 0.036, \eta_p^2 = 0.13$, were significant. Similarly, the two-way interaction was not significant, $F(1, 32) = 0.60, p = 0.44, \eta_p^2 = 0.02$. These results suggested that participants made more efforts when the other person's belief was incongruent with their own, regardless of the orientation of attribution. Moreover, making other-oriented belief attribution was more difficult than making self-oriented belief attribution across the true- and false-belief scenarios. The present results were in line with the findings from previous studies using the self/other differentiation task (Bradford et al., 2015; Bradford et al., 2019; Bradford, Hukker, et al., 2018a; Bradford, Jentzsch, et al., 2018b).

ERP results

Figure 3 presents the pooled activities of representative electrodes by condition for each ERP component. We performed Congruency (Congruent vs. Incongruent) \times Agent (Self vs.

Other) repeated-measures ANOVAs on N100, N250, and LNC. For the LPC, a Congruency (Congruent vs. Incongruent) \times Agent (Self vs. Other) \times Hemisphere (Left vs. Right) repeated-measures ANOVA was conducted.

For N100, the ANOVA only revealed an effect of Congruency, $F(1, 32) = 17.09, p < 0.001, \eta_p^2 = 0.35$, but neither the effect of Agent, $F(1, 32) = 0.19, p = 0.67, \eta_p^2 = 0.01$, nor the two-way interaction, $F(1, 32) = 2.65, p = 0.113, \eta_p^2 = 0.08$, was significant. Similarly, for N250, there was only an effect of Congruency, $F(1, 32) = 6.78, p = 0.014, \eta_p^2 = 0.18$, but no significant effect of either Agent, $F(1, 32) = 6.53, p = 0.16, \eta_p^2 = 0.17$, or interaction, $F(1, 32) = 0.87, p = 0.35, \eta_p^2 = 0.03$.

For the LNC, we found the effect of Congruency was significant, $F(1, 32) = 6.51, p = 0.016, \eta_p^2 = 0.17$. Although the effect of Agent was not significant, $F(1, 32) = 2.24, p = 0.14, \eta_p^2 = 0.07$, a significant two-way interaction was detected, $F(1, 32) = 6.35, p = 0.017, \eta_p^2 = 0.17$. Simple effect analyses showed that the effect of Agent was significant only in Incongruent condition, $F(1, 32) = 5.71, p = 0.023, \eta_p^2 = 0.15$, but not in Congruent condition, $F(1, 32) = 1.58, p = 0.22, \eta_p^2 = 0.05$.

For the LPC, we detected a significant effect of Congruency, $F(1, 32) = 8.43, p = 0.007, \eta_p^2 = 0.21$, and a marginal effect of Agent, $F(1, 32) = 3.52, p = 0.07, \eta_p^2 = 0.10$. Moreover, we also found a significant effect of Hemisphere in which the amplitude of LPC was higher in the right hemisphere than the one in the left hemisphere, $F(1, 32) = 103.28, p < 0.001, \eta_p^2 = 0.76$. Although neither two-way nor three-way interaction was detected, it is visually obvious in Fig. 3 that differences caused by agent only existed between Other-Incongruent and Self-Incongruent conditions, but not between Other-Congruent and Self-Congruent conditions. We performed an exploratory Agent (Self vs. Other) \times Hemisphere (Left vs. Right) repeated-measures ANOVA separately for Congruent and Incongruent conditions. As expected, a significant effect of Agent was detected only in Incongruent condition, $F(1, 32) = 4.86, p = 0.035, \eta_p^2 =$

Table 1 Mean RT (ms) and ACC (%) of each condition in the false-belief task

| | Congruent | | Incongruent | |
|-----|------------|------------|-------------|------------|
| | Self | Other | Self | Other |
| RT | 416 (156) | 402 (151) | 693 (115) | 671 (130) |
| ACC | 99.0 (2.9) | 98.0 (3.4) | 95.5 (5.1) | 94.0 (8.1) |

Standard deviations in parentheses.

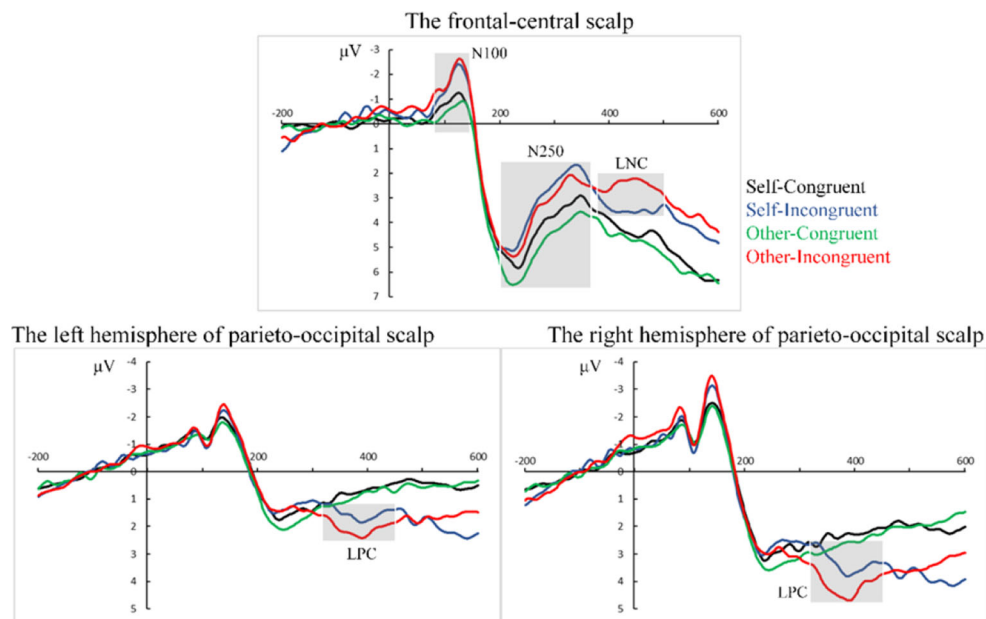


Fig. 3 Grand mean ERPs by the condition in each region of interest. To calculate the grand mean ERPs, activities of representative electrodes in each region of interest were pooled

0.13, but not in Congruent condition, $F(1, 32) = 0.64$, $p = 0.43$, $\eta_p^2 = 0.02$. In addition, a significant effect of Hemisphere was found in both Congruent and Incongruent conditions ($ps < 0.001$). Further analyses suggested that the failure to detect an interaction between Agent and Congruency might be due to a greater variability in Congruent condition than in Incongruent condition (Left hemisphere: Levene's test = 2.81, $p = 0.096$; Right hemisphere: Levene's test = 4.79, $p = 0.03$).¹ These were consistent with the claim that participants might take a cognitive shortcut in a true-belief scenario (Bio et al., 2018).

Correlation with AQ

Given the small sample size, we calculated correlations with a bootstrap of 1,000 iterations (bias-corrected and accelerated) by using SPSS 21 to prevent the observed relationships from being driven by some specific data points. As the scatterplots in Fig. 4 shows, we found that the differences in LPC (Other-Incongruent minus Self-Incongruent) in both hemispheres were positively correlated with the score on AQ (Left hemisphere: $r = 0.43$, $p = 0.013$, 95% bootstrapped confidence interval (CI): 0.18–0.65; Right hemisphere: $r = 0.38$, $p = 0.028$, 95% bootstrapped CI: 0.16–0.61), but the difference in LNC was not significantly related with the score on AQ ($r = 0.18$, $p = 0.32$, 95% bootstrapped CI: -0.13 to 0.47). Moreover, we found that the RT differences between Other-Incongruent and Self-Incongruent conditions and the

differences between Other-Congruent and Self-Congruent conditions were not significantly correlated with the AQ scores ($ps > 0.10$). We also examined the relationship between the AQ score and the Congruency effect (Incongruent minus Congruent) in either self- or other-oriented attribution. However, no significant correlation was detected in either RT or ERP measures ($ps > 0.10$).

Discussion

To communicate efficiently in social interaction, we must concurrently reason about mental states for both ourselves and others (Bradford et al., 2019). Using a false-belief task, the present study examined how the self- and other-oriented belief attributions are distinguished from each other. Both behavioral and ERP data were recorded. The behavioral results showed that participants made more errors and had a slower response tendency in other-oriented belief attribution than in self-oriented attribution. In ERP data, we found that engaging in other-oriented attribution elicited a larger amplitude than engaging in self-oriented attribution in two late ERP components: LPC and LNC. However, these results appeared only in the false-belief scenario, but not in the true-belief scenario. Furthermore, the differences in LPC amplitude between other-oriented attribution and self-oriented attribution were positively correlated with one's autistic traits.

Based on prior literature (Brass et al., 2009; Crone & Fuligni, 2020; Decety & Lamm, 2007; Steinbeis, 2016) and the characteristics of ERP components in the present study, we believed that the LPC and LNC might reflect the self-other

¹ In Levene's tests, the factor Congruency was treated as an independent-sample variable, and the self-oriented and other-oriented trials were merged in both Congruent and Incongruent conditions.

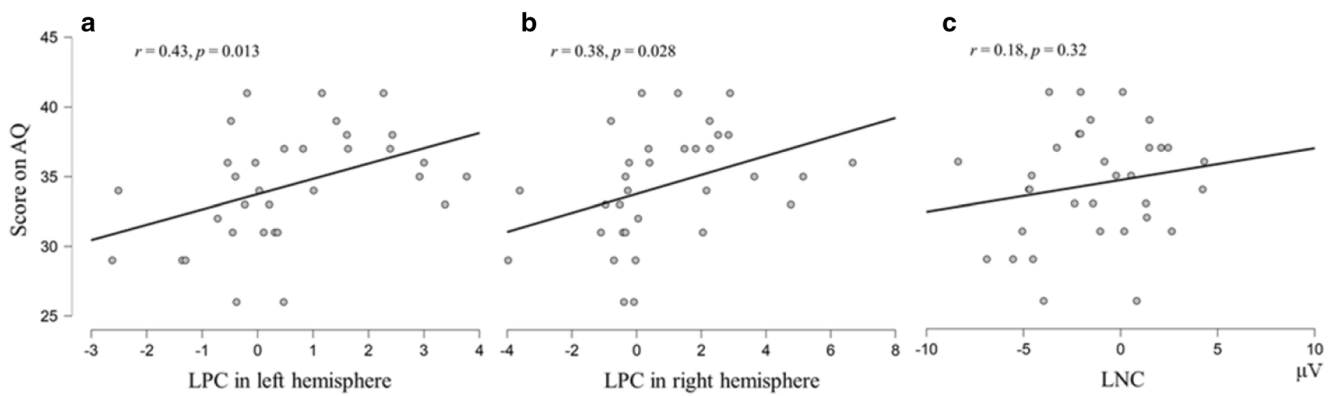


Fig. 4 Correlations between autistic traits (scores on AQ) and differences in LPC and LNC between other- and self-oriented false-belief attribution. **(A)** Scatterplots of correlation between AQ scores and LPC differences in

the left hemisphere. **(B)** Scatterplots of correlation between AQ scores and LPC differences in the right hemisphere. **(C)** Scatterplots of correlation between AQ scores and LNC differences

distinction process and the conflict control process, respectively. Results from the self/other differentiation task have suggested that resisting egocentric intrusion was more effortful than resisting altercentric intrusion during belief attribution (Bradford et al., 2015; Bradford et al., 2019; Bradford, Hukker, et al., 2018a; Bradford, Jentsch, et al., 2018b). The present ERP results provided further understanding of how the efforts came about. Specifically, we found that the efforts were captured by a larger LPC on the bilateral parieto-occipital scalp and a larger LNC on the frontal-central scalp. This indicated that more effort is required in making other-oriented belief attribution than in making self-oriented belief attribution in both the self-other distinction phase and the conflict control phase. Moreover, we found that the onset of LPC was earlier than the onset of LNC, yet there was an overlap in the time window between LPC and LNC (Fig. 3). We would interpret these findings in terms of computational processes underlying ToM. In a false-belief scenario, there would be a conflict between the belief of self and others (Bio et al., 2018; Kovács et al., 2010). Holding two conflicting representations of belief may force people to exert much more effort to suppress the mutual interference between self and others. However, just before controlling interference, we should first distinguish whether these conflicting representations belong to self or anyone else. Otherwise, it might lead to confusion between self and others (Steinbeis, 2016), and thus we may fail to determine whose belief to process. The overlap of time window between LPC and LNC suggests that the self-other distinction in false-belief attribution is an ongoing and iterative process. We assume that, although the self-other distinction starts earlier than the conflict control, it may fail to process if there is a stark conflict between self and others. Therefore, a process of conflict control before the complete separation of the self and others may be especially important. This assumption is consistent with the research on how people control their imitative behaviors, which claimed that there might be information exchange between the self-other

distinction and the conflict control (Brass et al., 2009). Crucially, our interpretation of the functional overlap of LPC and LNC would complement and extend the “anchoring-and-adjustment” framework (Tamir et al., 2010; Tamir & Mitchell, 2013), such that ongoing and iterative interaction between self-other distinction and conflict control forms the basis of the serial adjustments in mentalizing.

Most past studies of ToM focused on how the understanding of others’ false-beliefs differed from the understanding of others’ true-beliefs. Several ERP studies with a true- and false-belief manipulation showed that false-belief attribution often encompassed a late frontal slow-wave (Liu et al., 2009; Meinhardt et al., 2011; Meinhardt et al., 2012; Wang et al., 2008) and some early components, such as N100 and P200 (Wang et al., 2008). Similarly, in the present study, the ERP results revealed that the relative difficulty of belief attribution in the false-belief scenario induced a larger amplitude on the two early components, N100 and N250, as well as a larger longer-lasting late slow wave. Moreover, our behavioral data showed that the belief attribution was slower and less accurate in the false-belief scenario than in the true-belief scenario. The temporal differences in ERP responses between the effects of Congruency and Agent demonstrated that holding conflicting beliefs in mind would lead to cognitive difficulties at an earlier stage, while the attribution of the conflicting beliefs to self or other would be implemented at a later stage.

There has been consensus that the impairments in ToM capacity are associated with deficits in social communication among individuals with ASD. The present study showed that the differences in LPC between other- and self-oriented false-belief attribution were related to the variation of individuals’ autistic traits. This indicates that individuals with a higher degree of autistic traits have more difficulties distinguishing others’ beliefs from one’s own in the false-belief scenario. Furthermore, although more research on the specific role of the LPC and LNC is needed, the correlation results suggest that we should consider distinguishing different components

or processes in the processing of mental attribution, rather than considering only the overall responses such as the standard measures response time and accuracy. This may explain why the present study did not detect any correlation between behavioral data and the self-reported autistic traits. In addition, such a result could be relevant to the lack of group differences in the effect of attribution orientation in the self/other differentiation task between individuals with ASD and neurotypical individuals (Bradford, Hukker, et al., 2018a). However, another possible interpretation is that the control of self-other conflict may not be a bottleneck of mental attribution for healthy adults, because the healthy adults could have all achieved this ability (Qureshi et al., 2020). Hence, the conflicting control in belief attribution will not contribute to individual differences in performance that will correlate to healthy adults' autistic traits. This assumption is in accord with some researchers who stressed the critical role of self-other distinction in social cognition (Quesque & Brass, 2019; Steinbeis, 2016). Of course, given the small sample size in the present study, follow-up research with larger samples is needed to revalidate our correlation findings.

In numerous studies, visual perspective-taking has been regarded as an aspect of ToM. It reflects how individuals consider and calculate the perspectives of self and others (Drayton et al., 2018; Santiesteban et al., 2012; Symeonidou et al., 2016). It has been demonstrated that engaging in other-oriented perspective-taking requires more cognitive control than engaging in self-oriented visual perspective-taking (McCleery et al., 2011; Ramsey et al., 2013). An ERP study has shown that the temporoparietal scalp was activated in the calculation and representation of the perspectives of self and others, and the right frontal cortex was responsible for resolving the conflicts between perspectives (McCleery et al., 2011). Combined with prior findings (Bradford et al., 2019; McCleery et al., 2011), the present study revealed a consistent pattern of neural activity in mentalizing under a context of social interaction, namely a sequential coupling of the temporoparietal scalp and the frontal scalp. This neural circuit has been demonstrated to play a central role in the development of cognitive mentalizing from adolescence to early adulthood (Crone & Fuligni, 2020).

Finally, one last issue requires further discussion. In our false-belief task, the “other’s belief” was inferred based on the task scenes, whereas the “own belief” was always the true state of affairs. In the self/other differentiation task, participants would have to recall their own belief. This process could help to avoid the problem of “own belief = always reality.” However, both studies that employed the self-/other-differentiation task and the present study found that engaging in other-oriented belief attribution would be more difficult than engaging in self-oriented belief attribution. Recently, some researchers argued that participants may not necessarily generate a representation of mental states when the mental states correspond to

reality (Deschrijver & Palmer, 2020). This has been a common confound in the existing false-belief tasks, which could be a possible reason why making self-oriented attribution may be easier than making other-oriented attribution. Future false-belief research could take this potential confound into account.

Conclusions

By using a false-belief task, the current study investigated the dynamic differences between self- and other-oriented ToM processing. Results suggested that, to apply ToM in a context of social interaction, individuals would first have to distinguish between the self and others and then control the potential conflict between representations. Moreover, the difficulty in self-other distinction may relate to the degree of individuals' autistic traits. Based on our findings and the suggestions from prior literature, we propose that the self-other distinction in ToM is an ongoing and iterative process, which involves a functional interaction between the bilateral posterior scalp and the frontal-central scalp.

Funding This research was supported by the National Natural Science Foundation of China to Yanjie Su (32071075 and 31872782).

Declarations

Conflict of interest none

References

- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- 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. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251. <https://doi.org/10.1017/S0021963001006643>
- 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. <https://doi.org/10.1023/A:1005653411471>
- Bio, B., Webb, T., & Graziano, M. (2018). Projecting one's own spatial bias onto others during a theory-of-mind task. *Proceedings of the National Academy of Sciences of the United States of America*, 115(7), E1684–E1689. <https://doi.org/10.1073/pnas.1718493115>
- Bradford, E. E. F., Gomez, J., & Jentsch, I. (2019). Exploring the role of self/other perspective-shifting in theory of mind with behavioural and EEG measures. *Social Neuroscience*, 14(5), 530–544. <https://doi.org/10.1080/17470919.2018.1514324>
- Bradford, E. E. F., Hukker, V., Smith, L., & Ferguson, H. J. (2018a). Belief-attribution in adults with and without autistic spectrum disorders: Belief-attribution in adults with autism. *Autism Research*, 11(11), 1542–1553. <https://doi.org/10.1002/aur.2032>

- Bradford, E. E. F., Jentsch, I., & Gomez, J. (2015). From self to social cognition: Theory of mind mechanisms and their relation to executive functioning. *Cognition*, 138, 21–34. <https://doi.org/10.1016/j.cognition.2015.02.001>
- Bradford, E. E., Jentsch, I., Gomez, J., Chen, Y., Zhang, D., & Su, Y. (2018b). Cross-cultural differences in adult theory of mind abilities: a comparison of native-English speakers and native-Chinese speakers on the Self/Other differentiation task. *Quarterly Journal of Experimental Psychology*, 71(12), 2665–2676. <https://doi.org/10.1177/1747021818757170>
- Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2359–2367. <https://doi.org/10.1098/rstb.2009.0066>
- Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism*, 8(3), 283–299. <https://doi.org/10.1177/1362361304045217>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J: L. Erlbaum Associates.
- Cremone-Caira, A., Buirkle, J., Gilbert, R., Nayudu, N., & Faja, S. (2019). Relations between caregiver-report of sleep and executive function problems in children with autism spectrum disorder and attention-deficit/hyperactivity disorder. *Research in Developmental Disabilities*, 94, 103464. <https://doi.org/10.1016/j.ridd.2019.103464>
- Crone, E. A., & Fuligni, A. J. (2020). Self and others in adolescence. *Annual Review of Psychology*, 71, 447–469. <https://doi.org/10.1146/annurev-psych-010419-050937>
- Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6), 580–593. <https://doi.org/10.1177/1073858407304654>
- Deschrijver, E., & Palmer, C. (2020). Reframing social cognition: Relational versus representational mentalizing. *Psychological Bulletin*, 146(11), 941–969. <https://doi.org/10.1037/bul0000302>
- Drayton, L. A., Santos, L. R., & Baskin-Sommers, A. (2018). Psychopaths fail to automatically take the perspective of others. *Proceedings of the National Academy of Sciences of the United States of America*, 115(13), 3302–3307. <https://doi.org/10.1073/pnas.1721903115>
- de Guzman, M., Bird, G., Banissy, M., & Catmur, C. (2016). Self-other control processes in social cognition: From imitation to empathy. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 371(1686), 20150079. <https://doi.org/10.1098/rstb.2015.0079>
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Friedman, L., & Sterling, A. (2019). A review of language, executive function, and intervention in autism spectrum disorder. *Seminars in Speech and Language*, 40(4), 291–304. <https://doi.org/10.1055/s-0039-1692964>
- Frith, U., & de Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition*, 14(4), 719–738. <https://doi.org/10.1016/j.concog.2005.04.006>
- Hillier, A., & Allinson, L. (2002). Beyond expectations: autism, understanding embarrassment, and the relationship with theory of mind. *Autism*, 6(3), 299–314. <https://doi.org/10.1177/1362361302006003007>
- Johnston, K., Murray, K., Spain, D., Walker, I., & Russell, A. (2019). Executive function: cognition and behaviour in adults with autism spectrum disorders (ASD). *Journal of Autism and Developmental Disorders*, 49(10), 4181–4192. <https://doi.org/10.1007/s10803-019-04133-7>
- Kilford, E., Garrett, E., & Blakemore, S. (2016). The development of social cognition in adolescence: an integrated perspective. *Neuroscience and Biobehavioral Reviews*, 70, 106–120. <https://doi.org/10.1016/j.neubiorev.2016.08.016>
- Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834. <https://doi.org/10.1126/science.1190792>
- Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2009). Neural correlates of children's theory of mind development. *Child Development*, 80(2), 318–326. <https://doi.org/10.1111/j.1467-8624.2009.01262.x>
- Lombardo, M. V., & Baron-Cohen, S. (2011). The role of the self in mindblindness in autism. *Consciousness and Cognition*, 20(1), 130–140. <https://doi.org/10.1016/j.concog.2010.09.006>
- Louzolo, A., Gustavsson, P., Tigerström, L., Ingvar, M., Olsson, A., & Petrovic, P. (2017). Delusion-proneness displays comorbidity with traits of autistic-spectrum disorders and ADHD. *PloS One*, 12(5), e0177820. <https://doi.org/10.1371/journal.pone.0177820>
- Mannion, A., & Leader, G. (2013). Comorbidity in autism spectrum disorder: a literature review. *Research in Autism Spectrum Disorders*, 7(12), 1595–1616. <https://doi.org/10.1016/j.rasd.2013.09.006>
- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, 62(1), 103–134. <https://doi.org/10.1146/annurev-psych-120709-145406>
- McCleery, J. P., Surtees, A. D. R., Graham, K. A., Richards, J. E., & Apperly, I. A. (2011). The neural and cognitive time course of theory of mind. *Journal of Neuroscience*, 31(36), 12849–12854. <https://doi.org/10.1523/JNEUROSCI.1392-11.2011>
- Meinhardt, J., Kühn-Popp, N., Sommer, M., & Sodian, B. (2012). Distinct neural correlates underlying pretense and false belief reasoning: Evidence from ERPs. *Neuroimage*, 63(2), 623–631. <https://doi.org/10.1016/j.neuroimage.2012.07.019>
- Meinhardt, J., Sodian, B., Thoermer, C., Döhnel, K., & Sommer, M. (2011). True- and false-belief reasoning in children and adults: an event-related potential study of theory of mind. *Developmental Cognitive Neuroscience*, 1(1), 67–76. <https://doi.org/10.1016/j.dcn.2010.08.001>
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *British Journal of Developmental Psychology*, 5(2), 125–137. <https://doi.org/10.1111/j.2044-835X.1987.tb01048.x>
- Poon, V. W. K., Shu, D. S. T., Chan, R. W. S., Leung, C. N. W., & Leung, P. W. L. (2020). Comparing the psychometric properties of the self- and parent-report versions of autism-spectrum quotient-adult in Hong Kong (AQ-adult-HK). *Journal of Autism and Developmental Disorders*, 50(2), 524–528. <https://doi.org/10.1007/s10803-019-04276-7>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>
- Quesque, F., & Brass, M. (2019). The role of the temporoparietal junction in self-other distinction. *Brain Topography*, 32(6), 943–955. <https://doi.org/10.1007/s10548-019-00737-5>
- Qureshi, A. W., Monk, R. L., Samson, D., & Apperly, I. A. (2020). Does interference between self and other perspectives in theory of mind tasks reflect a common underlying process? Evidence from individual differences in theory of mind and inhibitory control. *Psychonomic Bulletin & Review*, 27(1), 178–190. <https://doi.org/10.3758/s13423-019-01656-z>
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, 27(2), 224–260. <https://doi.org/10.1016/j.dr.2007.02.001>
- Ramsey, R. (2018). What are reaction time indices of automatic imitation measuring? *Consciousness and Cognition*, 65, 240–254. <https://doi.org/10.1016/j.concog.2018.08.006>

- Ramsey, R., Hansen, P., Apperly, I., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes. *Journal of Cognitive Neuroscience*, 25(5), 670–684. https://doi.org/10.1162/jocn_a_00345
- Santesteban, I., White, S., Cook, J., Gilbert, S., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to theory of mind. *Cognition*, 122(2), 228–235. <https://doi.org/10.1016/j.cognition.2011.11.004>
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, 42(5), 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>
- Steinbeis, N. (2016). The role of self-other distinction in understanding others' mental and emotional states: Neurocognitive mechanisms in children and adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686), 20150074–20150074. <https://doi.org/10.1098/rstb.2015.0074>
- Stone, J. V. (2002). Independent component analysis: An introduction. *Trends in Cognitive Sciences*, 6(2), 59–64. [https://doi.org/10.1016/S1364-6613\(00\)01813-1](https://doi.org/10.1016/S1364-6613(00)01813-1)
- Symeonidou, I., Dumontheil, I., Chow, W., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology*, 149, 81–97. <https://doi.org/10.1016/j.jecp.2015.11.007>
- Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, 142(1), 151–162. <https://doi.org/10.1037/a0028232>
- Tamir, D. I., Mitchell, J. P., & Taylor, S. E. (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences of the United States of America*, 107(24), 10827–10832. <https://doi.org/10.1073/pnas.1003242107>
- Wang, Y., Liu, Y., Gao, Y., Chen, J., Zhang, W., & Lin, C. (2008). False belief reasoning in the brain: An ERP study. *Science in China, Series C: Life Sciences*, 51(1), 72–79. <https://doi.org/10.1007/s11427-008-0014-z>
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684. <https://doi.org/10.1111/1467-8624.00304>
- White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, 80(4), 1097–1117. <https://doi.org/10.1111/j.1467-8624.2009.01319.x>
- White, S., Hill, E., Winston, J., & Frith, U. (2006). An islet of social ability in Asperger Syndrome: Judging social attributes from faces. *Brain and Cognition*, 61(1), 69–77. <https://doi.org/10.1016/j.bandc.2005.12.007>
- White, S. J., & White, S. J. (2013). The triple I hypothesis: Taking another('s) perspective on executive dysfunction in autism. *Journal of Autism and Developmental Disorders*, 43(1), 114–121. <https://doi.org/10.1007/s10803-012-1550-8>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.