



Two systems for thinking about others' thoughts in the developing brain

Charlotte Grosse Wiesmann^{a,b,1} , Angela D. Friederici^c, Tania Singer^{d,2}, and Nikolaus Steinbeis^{e,2}

^aResearch Group Milestones of Early Cognitive Development, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; ^bCentre for Early Childhood Cognition, Institute for Psychology, Copenhagen University, 1353 Copenhagen, Denmark; ^cDepartment of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; ^dSocial Neuroscience Lab, Max Planck Society, 10099 Berlin, Germany; and ^eDivision of Psychology and Language Sciences, Faculty of Brain Sciences, University College London, WC1E 6BT London, United Kingdom

Edited by Uta Frith, University College London, London, United Kingdom, and approved February 3, 2020 (received for review September 26, 2019)

Human social interaction crucially relies on the ability to infer what other people think. Referred to as Theory of Mind (ToM), this ability has long been argued to emerge around 4 y of age when children start passing traditional verbal ToM tasks. This developmental dogma has recently been questioned by nonverbal ToM tasks passed by infants younger than 2 y of age. How do young children solve these tests, and what is their relation to the later-developing verbal ToM reasoning? Are there two different systems for nonverbal and verbal ToM, and when is the developmental onset of mature adult ToM? To address these questions, we related markers of cortical brain structure (i.e., cortical thickness and surface area) of 3- and 4-y-old children to their performance in novel nonverbal and traditional verbal ToM tasks. We showed that verbal ToM reasoning was supported by cortical surface area and thickness of the precuneus and temporoparietal junction, classically involved in ToM in adults. Nonverbal ToM reasoning, in contrast, was supported by the cortical structure of a distinct and independent neural network including the supramarginal gyrus also involved in emotional and visual perspective taking, action observation, and social attention or encoding biases. This neural dissociation suggests two systems for reasoning about others' minds—mature verbal ToM that emerges around 4 y of age, whereas nonverbal ToM tasks rely on different earlier-developing possibly social-cognitive processes.

Theory of Mind | false belief | cortical thickness | gray matter | brain development

Our daily interaction with other humans crucially relies on our ability to understand what they think or believe. Unlike their behavior, other people's mental states are not visible, and we, therefore, need to *infer* what is on their mind. This ability has been referred to as ToM. Reasoning about others' unobservable mental states allows for far more sophisticated predictions about how other individuals will act than merely observing their behavior. ToM, thus, constitutes a cornerstone of the complex social interaction and communication that characterizes humans.

For decades, it was believed that ToM was uniquely human (1) and emerged around 4 y of age (2–4). This traditional view was based on the standard ToM task that requires participants to reason about others' false beliefs (FBs), considered as the critical test of ToM. In recent years, however, this view has been questioned by new types of ToM tasks, which show that, if not explicitly prompted, infants younger than 2 y already display correct expectations of how others will act based on their FBs (5, 6). Because, in these novel tasks, the object of study is not made explicit, the tasks have been referred to as *implicit* ToM tasks. Recently, even apes have been shown to pass these implicit ToM tasks (7). These findings have led to one of the biggest puzzles of current developmental psychology. How do infants and apes solve these implicit ToM tasks? Do they represent other individuals' mental states? And if so, why do they then consistently fail the traditional explicit ToM tasks until several years later? At the core of this debate is the question of the relation between the processes underlying implicit

and explicit ToM tasks. While some claim that both task types measure the same early-developing core ToM capacity and only extraneous linguistic and executive task demands impede young preschoolers success on the explicit verbal tasks (6), others have argued, for two different systems, a mature explicit ToM process that develops at 4 y of age and an earlier-developing process measured by the implicit tasks (8, 9).

To address these questions, we studied the brain regions in which cortical structure supports success in implicit and in explicit ToM tasks in 3- to 4-y.o. children. We reasoned that, if both task types measure the same cognitive process, the maturation of similar brain regions should be relevant to their development. In contrast, if the maturation of different brain regions was important for succeeding on these tasks, this would support different underlying cognitive processes. Finally, comparing the relevant brain regions for children to those involved in ToM in adults would clarify which of these processes reflect mature adultlike ToM reasoning.

Mature verbal ToM reasoning has been shown to recruit a consistent network of brain regions including the temporoparietal junction (TPJ), middle temporal gyrus (MTG), precuneus (PC), and medial prefrontal cortex in adults (10) and in children (11–14). Increased functional and structural connectivities of this brain network have been associated with passing the explicit ToM tasks at around 4 y of age (13, 14). This leads to the prediction that this traditional milestone of passing explicit ToM tasks should be supported by cortical maturation in these regions. Much less is

Significance

The ability to reason about other people's thoughts and beliefs characterizes the complex social interaction among humans. This ability, called Theory of Mind (ToM), has long been argued to develop around 4 y when children start explicitly reasoning about others' beliefs. However, when tested nonverbally, infants already show action expectations congruent with others' beliefs before the age of 2 y. Do these behaviors reflect different systems for understanding others' minds—an early and a later developing one—or when does ToM develop? We show that these abilities are supported by the maturation of independent brain networks, suggesting different systems for explicit verbal ToM and early nonverbal action expectations.

Author contributions: C.G.W., A.D.F., T.S., and N.S. designed research; C.G.W. performed research; C.G.W. analyzed data; C.G.W. wrote the paper; and A.D.F., T.S., and N.S. edited the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: wiesmann@cbs.mpg.de.

²T.S. and N.S. contributed equally to this work.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1916725117/-DCSupplemental>.

First published March 9, 2020.

known about the neural structures involved in implicit ToM tasks. Neural activation for implicit and explicit ToM tasks have only been compared in adults thus far (15–19). Adults might engage in explicit verbal ToM while watching nonverbal ToM videos, however, possibly confounding brain regions for implicit and explicit ToM processes. Most of these studies have emphasized overlapping brain regions but have studied activation for implicit ToM only within the regions of interest defined by explicit ToM tasks (15–17), making it impossible to detect regions that differ from those recruited for explicit ToM. However, a recent multistudy analysis comparing whole-brain activation for implicit and explicit ToM tasks found a number of distinct brain regions for implicit ToM and an overlap with explicit ToM only in the right (R) TPJ (19). A recent study using functional near-infrared spectroscopy (fNIRS) in infants aged 7 mo gives further indications of activation over temporoparietal regions (20). However, the spatial resolution of fNIRS does not allow for more specific localization, and a direct comparison with explicit ToM tasks, which are designed for considerably older children, is not possible in infants. The critical question, thus, remains unanswered: Which brain regions and ultimately cognitive processes underlie young children's success in the implicit ToM tasks, and are these the same or different from those involved in the explicit ToM tasks? To address this question, we tested children with brain-structural magnetic resonance imaging (MRI) and a specifically tailored behavioral task battery of implicit and explicit ToM tasks at the critical age of 3–4 y when the ability to explicitly reason about others' beliefs develops. Only at this age are children able to perform both types of tasks, possibly succeeding in the implicit task without yet being able to engage in explicit verbal ToM. Thus, this approach uniquely allowed us to study and compare the brain regions associated with early success in implicit ToM tasks and the later emergence of explicit ToM reasoning. Furthermore, with our approach, we studied the structural cortical maturation associated with the development of ToM and, thus, shed light on the underlying neurodevelopmental mechanisms.

During childhood, the human cerebral cortex undergoes important maturational changes as a result of the interaction of complementary microphysiological processes including synaptic pruning and myelination (21–23). These changes are associated with the development of cognitive abilities, such as intelligence and executive function, throughout middle childhood and adolescence (24–27). Cortical surface area increases during childhood and has been found to be positively related with cognitive development (22, 25). By contrast, cortical thickness appears to decrease from early childhood, partly as a result of myelination (22, 23), although its developmental trajectory depends on complementary microphysiological processes and is subject to considerable local and interindividual variation (22–24). Consequently, cortical thickness has been found to be either positively or negatively related to cognitive function depending on the age, cognitive domain, and brain region (24, 26). In preschool age, the relation of cortical brain structure and cognitive function has been rarely studied to date, and, in particular, research on the relation to early social-cognitive development is lacking.

We correlated these measures of cortical brain structure (i.e., cortical surface area and thickness) in 38 children aged 3 and 4 y with their performance on traditional explicit ToM tasks (2, 28) and a nonverbal implicit ToM task (28). In addition, we assessed children's executive function, linguistic and general cognitive abilities, known to be related with ToM development (28), and controlled for developments in these domains. This approach served to: 1) determine which markers of cortical brain structure underlie the developmental breakthrough in the explicit ToM tasks; 2) investigate the regions associated with early success in implicit ToM tasks; and 3) compare these brain regions to one another and to the ToM network recruited in adults.

We reasoned that, if children's developmental breakthrough in the traditional explicit ToM tasks, indeed, reflects their emerging ability to reason about mental states, their explicit ToM performance should correlate with cortical brain structure in the regions recruited for ToM reasoning in adults. If implicit ToM tasks, in turn, tap into the same processes, the same or overlapping brain regions should be found for both task types. Distinct and clearly dissociated brain regions, in contrast, would support two different systems for implicit and explicit ToM.

Our results show that the developmental breakthrough in explicit ToM tasks around the age of 4 y is, indeed, related to structural changes in the brain regions of the adult ToM network, including the TPJ and PC. Earlier-developing success in the implicit ToM tasks, in contrast, is associated with different regions in the developing brain, including the supramarginal gyrus (SMG). The observed neural effects are independent of one another, independent of age and of developments in other cognitive domains. These findings clearly support a different-systems account of implicit and explicit ToM performance where mature adultlike ToM emerges with passing the explicit ToM tasks around the age of 4 y, supported by cortical brain structure in what is known as the ToM network in adults. The brain regions involved in implicit ToM tasks, in contrast, suggest that earlier-developing, possibly social-cognitive processes underlie younger children's success in these tasks.

Results

Behavioral Results. A behavioral assessment of explicit and implicit ToM tasks and codeveloping cognitive functions was conducted in 60 children aged 3 to 4 y (28) from which 38 children successfully performed an anatomical MRI (for details, see *Methods*). Explicit ToM performance was assessed with two explicit ToM tasks—a false location and a false content FB task—from which an aggregate explicit ToM score was computed (see *Methods*). In the subsample of 38 children with usable MRI data, there was a significant difference between 3- and 4-y.o.s (Mann-Whitney u test $P < 0.001$), in that 3-y.o. children performed below chance (median = 0, Wilcoxon signed rank test $P < 0.001$) and 4-y.o.s marginally above chance (median = 0.5, Wilcoxon signed rank test $P = 0.082$), confirming the results of the full behavioral sample (28).

In the implicit ToM task, children watched nonverbal animated hiding/chasing scenarios between two animals while their gaze direction was recorded with an eye tracker (for details, see *Methods*). The chasing animal had a FB about where the chased animal was hiding. Rather than asking children verbally as in the explicit tasks, in the implicit FB task, children's spontaneous looking behavior was observed to track where they anticipated the chasing animal to search for the hiding one. In the full behavioral sample, children looked significantly more to the correct than the incorrect location (mean = 54.0%, SD = 10.8%, one-sample t test $t[56] = 2.809$, $P = 0.007$) (28) and marginally significantly in the subsample of children with usable MRI data (mean = 53.4%, SD = 10.9%, one-sample t test $t[37] = 1.907$, $P = 0.064$). There was no significant age difference between 3 and 4 y (3-y.o.s: 54.3%, 4-y.o.s: 52.6%, independent sample t test $t[36] = 0.494$, $P = 0.778$) as in the full behavioral sample (28). Furthermore, as in the full sample, no correlation between implicit and explicit ToM performance was found (Spearman's rank correlation coefficient $\rho[38] = -0.124$, $P = 0.459$).

Explicit ToM Tasks and Brain Structure. To test whether ToM performance was related to circumscribed cortical brain structure, we reconstructed cortical surface area and thickness from high-resolution anatomical MRI using surface-based analysis in FreeSurfer (29) in the same 3- and 4-y.o. children that participated in the behavioral task battery. We then computed the linear relation of children's explicit ToM scores with these indices of gray matter structure on the entire cortical surface. This yielded a significant positive correlation of children's explicit

Table 1. MNI coordinates, effect size, exact significance, and cluster size of significant brain-behavior relations

	Anatomical region	Peak voxel coordinate in MNI 305 space (X, Y, Z)			Correlation in peak voxel (Spearman's ρ)	Clusterwise P value	Cluster-size (in mm ²)	
Explicit ToM	Surface area	R PC*	10.0	-28.5	25.9	0.547	0.00380 [†]	937.45
		R pMTG/ITS	47.5	-58.3	5.4	0.399	0.04136 [†]	660.93
		R TPJ*	56.1	-48	33.4	0.472	0.00020 [‡]	210.21
Cortical thickness	R pMTG*	59.2	-48.8	6.7	0.470	0.00020 [‡]	189.85	
	L PC	-11.6	-57.6	35.6	0.336	0.00479 [‡]	89.64	
Implicit ToM	Surface area	R SMG*	59.8	-17.4	31.8	0.331	0.00400 [†]	945.90
		Cortical thickness	L PC	-9.0	-71.9	55.5	0.396	0.00320 [‡]

The effects are controlled for age, gender, and total surface area/mean thickness. The explicit ToM effects were independent of implicit ToM and vice versa.

*These effects were independent of codeveloping abilities (i.e., language, executive function, and general intelligence).

[†]Corrected at whole brain level.

[‡]Small volume correction within the regions of a ToM meta-analysis (10).

ToM performance with their cortical surface area in the PC (Fig. 1A and Table 1), which remained significant when controlling for total surface area, chronological age, and gender. Given our clear a priori hypotheses on the brain regions relevant for explicit ToM reasoning, in a second step, we computed a small volume correction within the regions of a ToM meta-analysis on FB reasoning (10). In addition to the effect in the R PC, this showed a significant positive correlation of children's explicit ToM performance with their cortical surface area in the R TPJ (Fig. 1B and Table 1). Moreover, explicit ToM performance was significantly positively correlated with cortical thickness in the R pMTG and L PC. These effects again remained significant when controlling for age and gender as well as total surface area and mean thickness, respectively.

Implicit ToM Tasks and Brain Structure. To identify cortical brain structure associated with success in the implicit ToM task, we

computed the linear relation of children's implicit ToM performance with cortical surface area and thickness on the whole cortical surface. This yielded a significant positive correlation with surface area in the R SMG (Fig. 2A and Table 1) without any overlap with the regions found for explicit ToM, which was significant when controlling for total surface area, age, and gender.

Dissociation of Implicit and Explicit ToM. To test whether the implicit ToM task also recruited typical adult ToM regions, we computed a small volume correction within the regions of the FB reasoning meta-analysis (10). In contrast to the explicit task, there was no significant linear relation with implicit ToM performance within the adult ToM network except for a positive relation with cortical thickness in a small cluster in the L dorsal PC, which was more dorsal than the cluster reported for explicit ToM (Figs. 2B and 3). To test for the independence of the effects

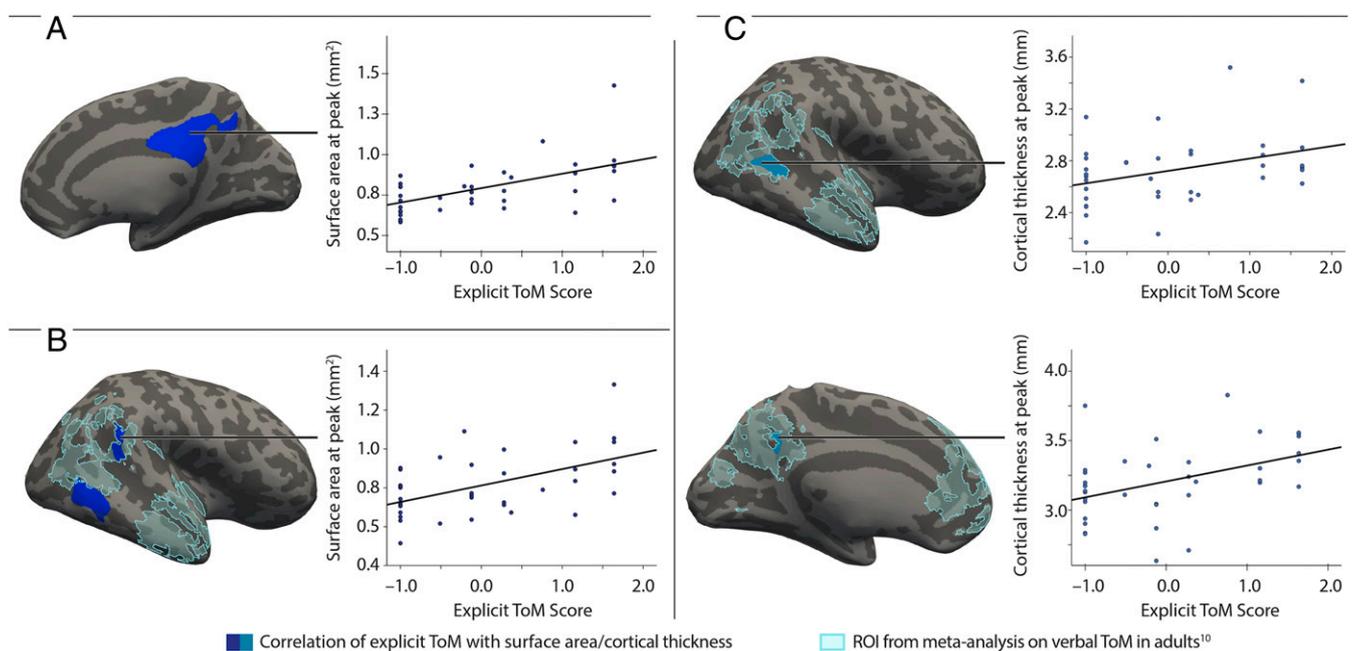


Fig. 1. Linear relation of explicit ToM performance with cortical surface area in the R PC (A, whole brain analysis) and R TPJ [B, within the regions of a ToM meta-analysis on FB reasoning in adults (10)]. (C) Linear relation of explicit ToM performance with cortical thickness in the R posterior MTG (pMTG, *Top*) and left (L) PC (*Bottom*) [within the regions of the adult ToM meta-analysis (10)]. These relations were independent of age, gender, the implicit ToM task, and codeveloping cognitive abilities. All effects are cluster-size corrected with a significance threshold of $P < 0.05$ and are shown on the inflated surface of the common group template.

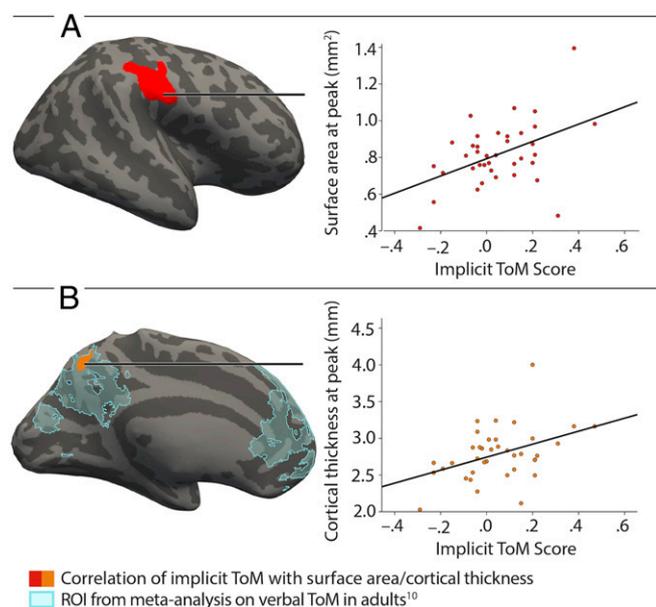


Fig. 2. (A) Linear relation of implicit ToM performance with cortical surface area in the R SMG (whole brain analysis). (B) Correlation of implicit ToM performance with cortical thickness in the L PC [within the regions of the adult ToM meta-analysis (10)]. These relations were independent of age, gender, the explicit ToM task, and the effect in A of codeveloping cognitive abilities. All effects are cluster-size corrected with a significance threshold of $P < 0.05$ and are shown on the inflated surface of the common group template.

observed for the implicit and explicit ToM tasks, we controlled for performance in the other task type, respectively. This showed that all regions of significant relation with explicit ToM performance were independent of implicit ToM performance and vice versa. While the regions found for explicit ToM performance fully overlapped with the brain network activated for explicit verbal FB reasoning in adults (10), success in the implicit ToM task was associated with different brain regions. These regions, in turn, lie within the area of activation found by a multistudy analysis on an implicit ToM paradigm (30) in adults (Fig. 3) with some differences in the peak coordinates (*SI Appendix, Table S1*).

Independence of Codeveloping Cognitive Abilities. To test for the specificity of the reported effects for ToM, we controlled for children's linguistic abilities, executive function, and general cognitive abilities. These were assessed with a comprehensive task battery of standardized tests of language development (31) and general cognitive function (32) as well as specifically tailored tasks known to be related to ToM in development (28) (for details, see *Methods* and *SI Appendix*). Including these scores as covariates in the linear model showed that all reported effects, except for the relatively small clusters in the L PC, were independent of children's linguistic, executive, and general cognitive abilities. The effects were, thus, specific to children's implicit or explicit ToM performance and not explained by more general cognitive development in other domains.

Discussion

Developmental psychology has been facing a puzzle: Our ability to reason about others' minds has long been believed to develop around the age of 4 y when children start passing the traditional explicit ToM tests. Presently, however, novel implicit ToM tasks have shown behavior that looks like ToM in infants even before the age of 2 y (6). When does ToM develop, and what is the relation between the processes subserving implicit and explicit ToM tasks? We addressed this question by studying structural brain

markers (cortical surface area and thickness) in relation to implicit and explicit ToM tasks at the critical age of 3 to 4 y—before and after explicit ToM reasoning emerges. The results show that the development of explicit ToM reasoning in young children was associated with increases in cortical thickness and surface area in brain regions of the classical ToM network (i.e., temporoparietal regions and the PC). Implicit ToM performance, in contrast, was associated with increases in cortical thickness and surface area in independent brain regions (i.e., the SMG and a more dorsal portion of the PC). These associations were independent of one another and independent of codeveloping abilities in other cognitive domains (i.e., language, executive function, and general cognitive abilities).

The finding that passing explicit ToM tasks was associated with brain structure in the core regions of the ToM network clearly supports the theory that the breakthrough observed on these tasks at 4 y reflects a development toward mature adultlike ToM reasoning. This speaks against the view that a full-blown ToM is innate or develops early in infancy. In particular, it is inconsistent with the view that young children only fail explicit ToM tasks due to extrinsic linguistic or executive demands of these tasks (5, 6) because, in that case, brain regions involved in language or executive function should be related to success in the tasks. On the contrary, passing the explicit ToM tasks seems to rely on the same brain regions and processes recruited when adults reason about others' mental states. This argument receives additional support from the fact that the observed effects were independent of children's linguistic abilities and executive function. These results complement previous findings that the structural and functional connectivities between these brain regions are important for the emergence of ToM reasoning in the explicit ToM tasks (13, 14).

Earlier-developing success in implicit ToM tasks, in contrast, was associated with brain structure in brain regions distinct and independent from those involved in explicit ToM tasks, including the SMG rather than the TPJ (i.e., angular gyrus). Indeed, we observed a double dissociation in that all of the reported effects correlated only with one task type but not the other with no overlap between the two networks. Furthermore, the effects found for the implicit ToM task were independent of performance on the explicit ToM tasks and vice versa. These findings clearly support independent processes involved in implicit and explicit ToM tasks—mature ToM is measured by the explicit tasks and only develops late in preschool age, whereas looking behavior in the implicit ToM tasks is driven by independent earlier-developing processes. The dissociation was independent of developments in other cognitive domains (i.e., language,

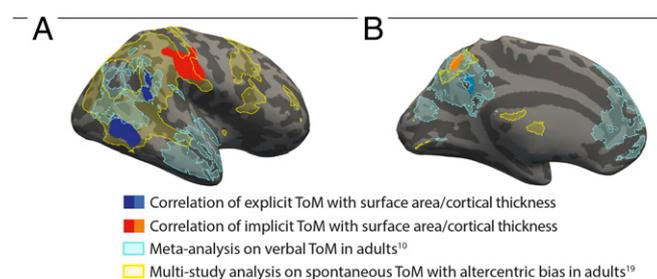


Fig. 3. Distinct and independent brain regions were associated with success on explicit (blue) and implicit ToM tasks (red, orange) as shown for the surface area in A and cortical thickness in B. While the effects found for the explicit ToM tasks were in regions activated by verbal ToM stories in adults (10), the implicit ToM effects lay within regions activated by a spontaneous ToM task in which adults experienced biases depending on another person's belief (19). The effects are shown on the inflated surface of the common group template.

executive function, working memory, and general intelligence) and was supported by the lack of a behavioral correlation between implicit and explicit ToM performances (28).

Our implicit ToM findings in the SMG and dorsal PC overlapped with the activation maps of a recent multistudy fMRI investigation on an implicit ToM paradigm in adults (19) with some differences in the peak coordinates (Fig. 3 and *SI Appendix, Table S1*). In the adult implicit ToM paradigm, participants experienced social biases when asked to detect a ball, depending on whether the other person believed the ball to be there or not (18, 19, 30). In contrast to our implicit ToM task in young children, this social bias task in adults additionally recruited the R TPJ—a core ToM region (18, 19). This suggests that, during such tasks, adults recruit those processes underlying implicit ToM success in early childhood, and, in addition, engage in explicit ToM.

What are the cognitive processes that allow young children to pass implicit ToM tasks? While explicit ToM reasoning relied on the brain regions of the classical ToM network, success in implicit ToM tasks was associated with different brain regions, that is, the SMG and a more dorsal portion of the PC. These regions are part of two distinct functional connectivity networks as shown with the automated fMRI synthesis software Neurosynth (<https://neurosynth.org/>) (33). The peak coordinates of explicit ToM (i.e., TPJ, ventral PC, and pMTG) are part of the default mode network commonly recruited for ToM, episodic and semantic memories (<https://neurosynth.org/>). The peak coordinates of the implicit ToM task (i.e., SMG and dorsal PC), in contrast, are coactivated in a superior parietal network associated with sensory, motor and body representations, and observation (<https://neurosynth.org/>). Similar dissociated networks have been observed in a number of studies reporting differences between ToM and other forms of perspective taking in adults (19, 34–36) and in children (14, 37). A series of studies comparing ToM and empathy found the SMG to be involved in overcoming egocentric biases when judging others' emotional states in both adults (34, 35) and children (37). The TPJ, in turn, was involved in explicit ToM reasoning. This dissociation of TPJ and SMG was confirmed in preschool-aged children when comparing ToM with observing others' bodily sensations or pain (*SI Appendix, Table S1*) (14). Interestingly, a similar dissociation of SMG versus TPJ and dorsal versus ventral PC was found when comparing visual perspective taking with ToM in tasks where participants had to overcome social biases between their own and another's incongruent visual perspective (38), similar to the implicit social bias ToM paradigms investigated in adults (18, 19, 30).

In cognitive development, similar to implicit ToM tasks, these social-cognitive processes [i.e., empathy, visual perspective taking, social biases between incongruent perspectives, as well as action understanding, which also recruits the SMG (39, 40)] have been shown to develop earlier than explicit ToM reasoning (30, 41–44). Based on the parallels in dissociated developmental trajectories and brain networks (SMG and dorsal PC versus TPJ and ventral PC), we speculate that success in the implicit ToM task might be related to a stronger focus on other agents, possibly resulting in a social modulation of encoded events. In line with recent theoretical suggestions, this could lead to social biases toward the others' (rather than one's own) perspective, which may generate success in implicit ToM tasks without a full-fledged ToM (45–48).

In addition to being involved in biases between one's own and others' emotional and visual perspectives and action observations, the SMG has also been discussed to be involved in attention control (49). However, the observed relation was independent of children's performance in executive function and general cognitive tasks, which involved visual attention. Moreover, attention tasks typically recruit a more posterior portion of the SMG or the angular gyrus (49–52) (see also <https://neurosynth.org/>). A closer region to our effect in the SMG was only observed in a study

where social context modulated attention (52), in line with our suggestion that socially modulated attention or encoding might drive success in implicit ToM tasks. Support for the involvement of general cognitive, in addition to social-cognitive, processes might come from the relation of the implicit ToM task with the dorsal PC, which was not independent of general cognitive abilities and executive function, and has previously been associated with selection, working memory, and processing location (<https://neurosynth.org/>). While, at the current stage, we can only speculate about the cognitive processes underlying early success in the implicit ToM tasks, our data clearly support that these processes differ from those underlying mature verbal ToM reasoning.

Open Questions. It remains an open question which processes drive children's correct looking behavior in the implicit ToM tasks, and the exact function of the SMG and dorsal PC in these tasks requires further investigation. Future research will need to directly contrast brain activation during these tasks with activation during other social-cognitive processes that have been shown to recruit portions of the SMG, such as handling conflict and biases in visual and emotional perspectives (14, 36), action observation (39, 40), and social as well as nonsocial modulations of attention and encoding (19, 34, 37, 52) in children younger than 4 y of age. This will be a challenge as it requires task-based functional MRI with very young children (14) because the relatively low spatial resolution of infant neuroimaging methods (such as NIRS) does not allow reliably measuring and dissociating the relevant brain regions (i.e., TPJ versus SMG and different portions of the PC) (20). In addition to understanding the neural processes underlying implicit ToM success and their exact function, further behavioral experimental manipulations and correlates would help to clarify what drives children's correct looking behavior in the implicit ToM tasks and how to interpret this behavior. It is not only the underlying processes, but also the exact developmental trajectory and potential limitations of implicit ToM processes that remain an open question. Recent studies have had difficulty replicating findings from different implicit ToM tasks (51–53), have reported considerably later developmental onsets of success in these tasks (53), and have reported specific performance limitations (53, 54). Behavioral success in the current MRI sample was also weaker than in the original behavioral study of our implicit ToM paradigm (27), which likely resulted from reduced power in the smaller sample of children with usable MRI data. We believe that the present study contributes to the debate about limitations of implicit ToM performance, first, by indicating that a different process underlies success in these tasks, which would explain some of the observed performance limitations (47). Second, the robust correlation of the implicit ToM task with brain structure supports that the variance in performance on this task is meaningful and reflects robust interindividual differences.

Conclusions. Our finding of distinct and independent brain networks for implicit and explicit ToM tasks shows a dissociation of the processes underlying these achievements at different time points in early childhood. The emergence of verbal ToM reasoning in the traditional explicit ToM tasks was associated with cortical brain structure in the classical ToM network. This supports the traditional view that the behavioral breakthrough on these tasks around 4 y of age represents a major advance toward a mature adultlike ToM. Success in the implicit ToM tasks, in contrast, was associated with an independent brain network (including the SMG), which has previously been reported to be involved in processing conflict in emotional and visual perspective taking, action observation, and social attention or encoding biases. In sum, these results support the involvement of different systems in what has been thought to reflect the development of ToM—with mature adultlike ToM emerging around the age of

4 y and earlier-developing possibly social-cognitive processes that support success in implicit ToM tasks. In addition to providing a better understanding of the cognitive architecture and development of human social cognition, our findings might have implications for understanding neurodevelopmental disorders with social-cognitive impairments, such as autism spectrum disorder, which has been found to differentially impair implicit and explicit ToM performances (55).

Methods

Participants. MRI data and behavioral data of 38 typically developing 3- and 4-y.o. children were analyzed for the present study (median = 4.07 y, range: 3.07–4.58 y; 21 female). The behavioral assessment was conducted in a total sample of $n = 60$ children aged 3 and 4 y (28) from which $n = 9$ children did not participate in or aborted the MRI, one child was excluded due to an MRI acquisition error, one child was excluded because of an incidental neurological finding, and $n = 11$ children were excluded because of motion artifacts in the MRI data detected by visual inspection only including data of participants that had good to moderate MRI scans (56). A power analysis with G*Power (57) showed that the computed linear regression with $n = 38$, an α error of 5%, and an effect size of $f(2) = 0.25$ had a power of $1 - \beta = 85\%$. Parental informed consent was obtained for all children, and the study was approved by the Ethics Committee at the Faculty of Medicine of the University of Leipzig.

Cognitive Assessment. The children took the completed task battery including traditional explicit ToM tasks and an implicit ToM task. In addition, to ensure that the observed effects were specific to ToM and not explained by developments in other cognitive domains, we administered standardized test batteries of language development (31) and general intelligence (32) as well as a specifically tailored set of executive function tasks known to correlate with ToM (28) (for details, see *SI Appendix*). This comprehensive cognitive assessment took place on three different days within a median period of 13 d (interquartile range: 10–16 d). The tasks were conducted in counter-balanced order across participants with the exception of the explicit ToM tasks that always took place last to avoid any influence of the explicit on the implicit ToM task.

Explicit ToM Tasks. The children performed two traditional explicit FB tasks—a false location and a false content task—both presented in an interactive setting [procedure as in Grosse Wiesmann et al. (13, 28)]. In the false location task, children saw a mouse puppet finding a candy in a little bag and an empty box. The mouse then left the room, and the experimenter conspiratorially moved the candy from the bag to the box. When the mouse came back, children were asked three test questions about where the mouse would look for the candy, whether she knew where it was and where she believed it was, followed by a control question to make sure the child remembered the actual location of the candy. In the false content task, children were shown a Kinder chocolate bar box and were asked what they believed was inside the box. After suspecting chocolates inside the box, they were shown that it, in fact, contained pencils. A mouse puppet then entered the scene, and children were asked three test questions: whether the mouse knew what was in the box, what she believed was in it, and what the child itself had originally believed was inside the box, followed by a control question about the actual content of the box. In each of the tasks, children could obtain a total of three points, one for each of the test questions. The performance on the two tasks was highly intercorrelated [Spearman's $\rho(38) = 0.889$; $P = 9 \times 10^{-14}$], and we, therefore, combined them into a total explicit ToM score with equal weight for each of the six test questions.

Implicit ToM Task. In an anticipatory looking paradigm (28), children were presented with short film clips on a Tobii T120 eye-tracker monitor. These showed different animal agents observing a mouse running through a y-shaped tunnel to one of two boxes each positioned at one of the two exits of the tunnel (*SI Appendix, Fig. S1*). The children first saw two familiarization (FAM) videos in which the animal agent then followed the mouse through the tunnel and took the exit where the mouse was hiding. After that, children were shown 12 FB videos in which the animal agent had a FB about the location of the mouse, which had actually left the scene in the agent's absence. Once the agent entered the tunnel, children's anticipatory looking was evaluated as a measure of their expectation as to where the agent would come out of the tunnel and search for the mouse. The video ended before the agent came out of the tunnel to prevent children from learning throughout the trials. There were two different FB conditions, respectively,

controlling for different non-belief-related strategies (58). To keep up children's anticipatory looking, the FB videos were intermixed with further FAM videos (a total of 10 trials) and six true belief (TB) videos analogous to the FB videos, except that the mouse returned to the box that the agent had witnessed it going to. In the FAM and TB videos, the agent always exited the tunnel on the side where the mouse was hiding. As in the original article (28), gaze data were analyzed for a period of interest from the moment when the agent had disappeared in the tunnel until its reappearance in the FAM and TB conditions or until the end of the trial in the FB conditions. Two areas of interest (AOI) were defined, each covering one of the tunnel exits and the corresponding box. During the period of interest, a differential looking score of relative looking times to the correct compared to the incorrect AOI was computed in accordance with previous literature (55) and served as the implicit ToM score.

MRI Data Acquisition. High-resolution 3D T_1 -weighted MRI images were acquired on a Siemens 3T TIM Trio scanner with a 32-channel head coil using the MP2RAGE sequence (59) at $1.2 \times 1 \times 1$ mm resolution (inversion time $TI_1 = 700$ ms; flip angle $\alpha_1 = 4^\circ$; $TI_2 = 2,500$ ms; $\alpha_2 = 5^\circ$; repetition time $TR = 5,000$ ms; echo time $TE = 3.24$ ms; $FoV = 192 \times 192$ mm; 176 sagittal slices; GRAPPA 3; partial Fourier phase factor 6/8; bandwidth 240 Hz/Px; acquisition time 5:22 min). Children were acquainted with the scanning procedure by performing a mock scan in a playful setting a few days before the actual scan and watched a movie of their choice on MR-compatible goggles and headphones during the scan.

MRI Data Analysis. Individual brain images were processed in FreeSurfer 5.3.0 (<http://surfer.nmr.mgh.harvard.edu>) to reconstruct cortical surfaces and generate local estimates of cortical thickness and surface area following the standard surface-based stream in FreeSurfer (60). This processing stream includes intensity normalization, followed by skull stripping and tessellation of the gray/white matter cortical boundary, automated topology correction (61), surface deformations following intensity gradients to optimize the inner (gray/white matter), and the outer (gray matter/cerebrospinal fluid [CSF]) cortical border to the location of greatest shift in intensity (60). Surface area of the gray/white matter boundary was calculated as well as cortical thickness defined as the shortest distance from the gray/white matter boundary to the gray matter/CSF boundary at each vertex on the tessellated surface. Skull stripping, white matter segmentation, and cortical and pial surfaces were inspected visually for errors and corrected manually when necessary as recommended in the FreeSurfer pipeline (<http://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData>). The automated FreeSurfer pipeline was rerun for the surfaces that contained errors and then reinspected. The individual cortical surfaces were then registered to a common group template with the spherical registration implemented in FreeSurfer to allow for an accurate matching of local cortical thickness and surface area measures across participants. Cortical thickness and surface area maps were smoothed on the tessellated surfaces using a 10-mm FWHM Gaussian kernel to reduce measurement noise while preserving anatomical localizability.

Brain Template Creation. The common group template had been created from the individual T_1 -weighted images of all of the children included in the analysis using the advanced normalization tools (ANTs) script `buildtemplate_parallel.sh` (<http://stnava.github.io/ANTs/>) after affine alignment to Montreal Imaging Institute (MNI) space, intensity correction, normalization, and skull stripping in FreeSurfer. The group template was then processed in FreeSurfer to reconstruct the cortical surfaces as described for the individual brain images. Segmentation and surfaces were again inspected visually and corrected manually, and the automated FreeSurfer pipeline was rerun and reinspected.

Statistical MRI Data Analysis. The statistical analyses were computed with FreeSurfer 6.0. The relation of cortical thickness and surface area, respectively, with our main variables (explicit and implicit ToM scores) were estimated in separate general linear models using the tool `mri_glmfit` implemented in FreeSurfer. In addition, we controlled for children's chronological age, gender, and mean thickness or total surface area, respectively, by including them as covariates in the linear model. We corrected for multiple comparisons with a clusterwise correction using the FreeSurfer tool `mri_glmfit-sim`, specifying a cluster-forming threshold of $P < 0.001$, clusterwise threshold of $P < 0.05$, positive relation with surface area, bidirectional relation with cortical thickness, and additional correction for the analyses on two hemispheres. For the clusterwise correction, a Monte Carlo simulation with 10,000 iterations was precomputed on our group template.

To test for the independence of the effects found for implicit and explicit ToM from one another, we additionally included the implicit ToM score as a covariate to the linear models for the explicit ToM score and vice versa. All effects remained significant.

Finally, to test for the specificity of the obtained effects for ToM, we controlled, for children's executive function, language, and general intelligence scores by including these as covariates into the linear models.

Regions of Interest. Given our clear a priori hypotheses for explicit ToM and to be able to compare the effects found in children to the mature adult ToM network, in addition to the whole brain analysis, we computed a small-volume correction within the regions of interest from a meta-analysis on reasoning about FBs in adults (10). For this, we registered the original meta-analysis maps from the MNI space to our group template with the ANTs script `WarpImageMultiTransform` (<http://stnava.github.io/ANTs/>) and then projected them on the surface using the FreeSurfer tool `mri_vol2surf`. The linear models for the relations of cortical thickness and surface area with explicit and implicit ToM, respectively (as well as the analyses that included the covariates described above) were computed within the obtained mask with `mri_glmfit` as before.

For purposes of qualitative comparison and visualization, the activation maps of the FB versus TB ("PxA") contrast of a recent multistudy analysis on implicit ToM in adults (19) was warped and projected to our group template surface as described above and displayed on the surface in Fig. 3. The effects were visualized on the inflated surface of the common group template with Freeview.

Data Availability Statement. All employed materials and data discussed in the paper are saved in a local repository at the Max Planck Institute for Human Cognitive and Brain Sciences, and data in fully anonymized format will be made available to the reader upon request (according to the data protection policy in the ethics agreement).

ACKNOWLEDGMENTS. We thank Riccardo Cafiero, Sofie Valk, and Clara Kühn for their methodological guidance on gray matter analyses in FreeSurfer; Julia Werner, Anne Grigutsch, Lisa Uhlich, and Alina Kowald for their help with correcting segmentation and surface reconstruction in FreeSurfer; Hung Nguyen Trong and Christiane Attig for their help with data acquisition and organization; and Emanuela Yeung for proofreading the paper. Parts of this project were funded by the Studienstiftung des deutschen Volkes and the Marie Skłodowska-Curie actions (MSCA) of the European Commission to C.G.W.

1. J. Call, M. Tomasello, Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci. (Regul. Ed.)* **12**, 187–192 (2008).
2. H. Wimmer, J. Perner, Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–128 (1983).
3. J. H. Flavell, F. L. Green, E. R. Flavell, Developmental changes in young children's knowledge about the mind. *Cogn. Dev.* **5**, 1–27 (1990).
4. H. M. Wellman, D. Cross, J. Watson, Meta-analysis of theory-of-mind development: The truth about false belief. *Child Dev.* **72**, 655–684 (2001).
5. R. M. Scott, R. Baillargeon, Early false-belief understanding. *Trends Cogn. Sci.* **21**, 237–249 (2017).
6. R. Baillargeon, R. M. Scott, Z. He, False-belief understanding in infants. *Trends Cogn. Sci.* **14**, 110–118 (2010).
7. C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, Great apes anticipate that other individuals will act according to false beliefs. *Science* **354**, 110–114 (2016).
8. I. A. Apperly, S. A. Butterfill, Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* **116**, 953–970 (2009).
9. C. D. Frith, U. Frith, Implicit and explicit processes in social cognition. *Neuron* **60**, 503–510 (2008).
10. M. Schurz, J. Radua, M. Aichhorn, F. Richlan, J. Perner, Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42**, 9–34 (2014).
11. D. Liu, M. A. Sabbagh, W. J. Gehring, H. M. Wellman, Neural correlates of children's theory of mind development. *Child Dev.* **80**, 318–326 (2009).
12. R. R. Saxe, S. Whitfield-Gabrieli, J. Scholz, K. A. Pelphrey, Brain regions for perceiving and reasoning about other people in school-aged children. *Child Dev.* **80**, 1197–1209 (2009).
13. C. Grosse Wiesmann, J. Schreiber, T. Singer, N. Steinbeis, A. D. Friederici, White matter maturation is associated with the emergence of theory of mind in early childhood. *Nat. Commun.* **8**, 14692 (2017).
14. H. Richardson, G. Lisandrelli, A. Riobueno-Naylor, R. Saxe, Development of the social brain from age three to twelve years. *Nat. Commun.* **9**, 1027 (2018).
15. Á. M. Kovács, S. Kühn, G. Gergely, G. Csibra, M. Brass, Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PLoS One* **9**, e106558 (2014).
16. C. K. Naughtin *et al.*, Do implicit and explicit belief processing share neural substrates? *Hum. Brain Mapp.* **38**, 4760–4772 (2017).
17. D. Schneider, V. P. Slaughter, S. I. Becker, P. E. Dux, Implicit false-belief processing in the human brain. *Neuroimage* **101**, 268–275 (2014).
18. L. Bardi, C. Desmet, A. Nijhof, J. R. Wiersema, M. Brass, Brain activation for spontaneous and explicit false belief tasks overlaps: New fMRI evidence on belief processing and violation of expectation. *Soc. Cogn. Affect. Neurosci.* **12**, 391–400 (2017).
19. S. Boccadoro *et al.*, Defining the neural correlates of spontaneous theory of mind (ToM): An fMRI multi-study investigation. *Neuroimage* **203**, 116193 (2019).
20. D. C. Hyde, C. E. Simon, F. Ting, J. I. Nikolova, Functional organization of the temporal-parietal junction for theory of mind in preverbal infants: A near-infrared spectroscopy study. *J. Neurosci.* **38**, 4264–4274 (2018).
21. K. B. Walhovd, A. M. Fjell, J. Giedd, A. M. Dale, T. T. Brown, Through thick and thin: A need to reconcile contradictory results on trajectories in human cortical development. *Cereb. Cortex* **27**, 1472–1481 (2017).
22. R. Cafiero, J. Brauer, A. Anwander, A. D. Friederici, The concurrence of cortical surface area expansion and white matter myelination in human brain development. *Cereb. Cortex* **29**, 827–837 (2019).
23. V. S. Natu *et al.*, Apparent thinning of human visual cortex during childhood is associated with myelination. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20750–20759 (2019).
24. P. Shaw *et al.*, Intellectual ability and cortical development in children and adolescents. *Nature* **440**, 676–679 (2006).
25. K. B. Walhovd *et al.*, Neurodevelopmental origins of lifespan changes in brain and cognition. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9357–9362 (2016).
26. H. G. Schnack *et al.*, Changes in thickness and surface area of the human cortex and their relationship with intelligence. *Cereb. Cortex* **25**, 1608–1617 (2015).
27. K. L. Mills, F. Lalonde, L. S. Clasen, J. N. Giedd, S. J. Blakemore, Developmental changes in the structure of the social brain in late childhood and adolescence. *Soc. Cogn. Affect. Neurosci.* **9**, 123–131 (2014).
28. C. Grosse Wiesmann, A. D. Friederici, T. Singer, N. Steinbeis, Implicit and explicit false belief development in preschool children. *Dev. Sci.* **20**, 1–15 (2017).
29. B. Fischl, FreeSurfer. *Neuroimage* **62**, 774–781 (2012).
30. Á. M. Kovács, E. Téglás, A. D. Endress, The social sense: Susceptibility to others' beliefs in human infants and adults. *Science* **330**, 1830–1834 (2010).
31. H. Grimm, *Sprachentwicklungstest für drei- bis fünfjährige Kinder: SETK 3-5: Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen* (Hogrefe, Verlag für Psychologie, 2001).
32. P. Melchers, U. Preuß, *K-ABC Kaufman Assessment Battery for Children, Deutsche Version Individualtest zur Messung von Intelligenz und Fertigkeit bei Kindern (6. teilweise ergänzte Aufl.)* (Hogrefe, Göttingen, 2003).
33. T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, T. D. Wager, Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
34. G. Silani, C. Lamm, C. C. Ruff, T. Singer, Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. *J. Neurosci.* **33**, 15466–15476 (2013).
35. P. Kanske, A. Böckler, F. M. Trautwein, F. H. Parianen Lesemann, T. Singer, Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Soc. Cogn. Affect. Neurosci.* **11**, 1383–1392 (2016).
36. M. Schurz *et al.*, Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity. *Neuroimage* **117**, 386–396 (2015).
37. N. Steinbeis, B. C. Bernhardt, T. Singer, Age-related differences in function and structure of rSMG and reduced functional connectivity with DLPFC explains heightened emotional egocentricity bias in childhood. *Soc. Cogn. Affect. Neurosci.* **10**, 302–310 (2015).
38. M. Schurz, M. Aichhorn, A. Martin, J. Perner, Common brain areas engaged in false belief reasoning and visual perspective taking: A meta-analysis of functional brain imaging studies. *Front. Hum. Neurosci.* **7**, 112 (2013).
39. D. Y. J. Yang, G. Rosenblau, C. Keifer, K. A. Pelphrey, An integrative neural model of social perception, action observation, and theory of mind. *Neurosci. Biobehav. Rev.* **51**, 263–275 (2015).
40. S. Morales, L. C. Bowman, K. R. Velnoskey, N. A. Fox, E. Redcay, An fMRI study of action observation and action execution in childhood. *Dev. Cogn. Neurosci.* **37**, 100655 (2019).
41. T. Singer, The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neurosci. Biobehav. Rev.* **30**, 855–863 (2006).
42. B. Sodian, C. Thoermer, U. Metz, Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Dev. Sci.* **10**, 199–204 (2007).
43. J. Decety, The neural pathways, development and functions of empathy. *Curr. Opin. Behav. Sci.* **3**, 1–6 (2015).
44. E. N. Cannon, A. L. Woodward, Infants generate goal-based action predictions. *Dev. Sci.* **15**, 292–298 (2012).
45. V. Southgate, Are infants altercentric? The other and the self in early social cognition. *Psychol. Rev.* **10.1037/rev0000182** (2019).
46. M. Tomasello, How children come to understand false beliefs: A shared intentionality account. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8491–8498 (2018).
47. C. Grosse Wiesmann, *The Emergence of Theory of Mind—Cognitive and Neural Basis of False Belief Understanding in Preschool Age* (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, 2017), vol. 193.
48. C. Grosse Wiesmann, V. Southgate, "Early Theory of Mind development—Are infants inherently altercentric?" in *The Neural Basis of Mentalizing—A Social-Cognitive and Affective Neuroscience Perspective*, M. Gilead, K. Ochsner, Eds. (Springer Nature, Basel, Switzerland, 2020), in press.
49. M. Corbetta, G. Patel, G. L. Shulman, The reorienting system of the human brain: From environment to theory of mind. *Neuron* **58**, 306–324 (2008).

50. S. C. Krall *et al.*, The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Struct. Funct.* **220**, 587–604 (2014).
51. R. M. Carter, S. A. Huettel, A nexus model of the temporal-parietal junction. *Trends Cogn. Sci.* **17**, 328–336 (2013).
52. T. Schuwerk, M. Schurz, F. Müller, R. Rupprecht, M. Sommer, The rTPJ's overarching cognitive function in networks for attention and theory of mind. *Soc. Cogn. Affect. Neurosci.* **12**, 157–168 (2017).
53. C. Grosse Wiesmann, A. D. Friederici, D. Disla, N. Steinbeis, T. Singer, Longitudinal evidence for 4-year-olds' but not 2- and 3-year-olds' false belief-related action anticipation. *Cogn. Dev.* **46**, 58–68 (2017).
54. D. Poulin-Dubois *et al.*, Do infants understand false beliefs? We don't know yet—A commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cogn. Dev.* **48**, 302–315 (2018).
55. A. Senju, V. Southgate, S. White, U. Frith, Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science* **325**, 883–885 (2009).
56. L. L. Backhausen *et al.*, Quality control of structural MRI images applied using FreeSurfer—a hands-on workflow to rate motion artifacts. *Front. Neurosci.* **10**, 558 (2016).
57. F. Faul, E. Erdfelder, A. G. Lang, A. Buchner, G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
58. V. Southgate, A. Senju, G. Csibra, Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* **18**, 587–592 (2007).
59. J. P. Marques *et al.*, MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage* **49**, 1271–1281 (2010).
60. B. Fischl, A. M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11050–11055 (2000).
61. B. Fischl, A. Liu, A. M. Dale, Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans. Med. Imaging* **20**, 70–80 (2001).