

Early signatures of and developmental change in brain regions for theory of mind

Hilary Richardson and Rebecca Saxe

Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, United States

Chapter outline

21.1. Early sensitivity to mental states: prior neural and behavioral evidence	468	21.4. Future directions: open questions and challenges	474
21.2. Early sensitivity to mental states: neuroimaging studies of young children and infants	468	21.4.1. Neural correlates of structural changes in theory of mind	474
21.3. Neural correlates of ongoing theory of mind development in childhood	470	21.4.2. Discovering reliable neural markers of individual differences in theory of mind	475
21.3.1. Response selectivity: fine-tuning preferential responses	471	21.4.3. The role of developmental experience: language	475
21.3.2. Reliable spontaneous (uninstructed) responses to movies	472	21.4.4. The role of developmental experience: culture	477
21.3.3. Integration and separation of functional networks	473	21.4.5. The role of family on ToM: shared environment and shared genes	478
		21.5. Conclusion	479
		References	480

Humans are social creatures: we spend much of our time thinking about and building relationships with other people. One cognitive foundation of our social abilities is an intuitive “Theory of Mind” (ToM) by which we reason about the beliefs, desires, and emotions of others, and relate these internal mental states to observable behaviors (Gopnik and Wellman, 1992, 1994). We use our ToM to predict behavior, explain disagreements, and make moral evaluations.

In human adults, multiple brain regions are preferentially engaged by thinking about others’ thoughts, including bilateral temporoparietal junction (TPJ), precuneus (PC), and medial prefrontal cortex (MPFC; for reviews, see Carrington and Bailey, 2009; Dubois and Adolphs, 2016b; for metaanalyses, see Mar, 2011; Schurz et al., 2014). Blood oxygenation in these regions, measured with fMRI, is high when people read stories or see cartoons or watch movies that evoke a character’s thoughts and feelings, compared to stories or cartoons or movies that focus on physical events. The right TPJ in particular has a highly selective response for thinking about the mental states of others: it responds preferentially to others’ thoughts, relative to a person’s physical appearance, personality traits, or bodily sensations (e.g., hunger, pain; Bruneau et al., 2012; Mitchell et al., 2005; Saxe and Kanwisher, 2003; Saxe and Powell, 2006). In many contexts, though, all of these regions function as an integrated network, showing highly correlated responses to social stimuli (Hasson, 2004; Jääskeläinen et al., 2008; Lerner et al., 2011; Nummenmaa et al., 2012; Wilson et al., 2007) and even in the absence of any external stimuli¹ (Fox et al., 2005; Greicius et al., 2003).

1. Studies measuring highly correlated responses in ToM brain regions at rest often refer to this network as the “Default Mode Network.” “Default” refers to evidence that these brain regions are activated at rest (“by default,” e.g., while internalizing), rather than during traditional attention tasks (Buckner et al., 2008; Raichle et al., 2001). However, DMN regions and ToM brain regions are overlapping (Gusnard et al., 2001; Laird et al., 2013; Smith et al., 2009), and responses in these regions are evoked by social tasks.

A fundamental question for developmental cognitive neuroscience is therefore: how do human children develop this cortical network for thinking about others' minds? When do specific cortical regions begin to respond preferentially while thinking about others' minds, and operate together as a functional network? What are the computations performed by these regions, and how do they change to support children's increasingly sophisticated ToM? What are the environmental and genetic factors that promote or hinder this development?

In this chapter, we review recent research that begins to address these questions using neuroimaging of young human children and infants (research on adolescence is not reviewed in this chapter; for reviews, see [Foulkes and Blakemore, 2018](#); [Kilford et al., 2016](#)). First, we ask about the origins of the ToM network: at what age do the TPJ, PC, and MPFC first function as an integrated network, and respond preferentially to considering others' minds? Second, we ask about developmental change in the ToM network: do the cognitive computations performed by these regions change during childhood? Third, we consider the limits of existing research, and conclude with avenues for future research.

21.1 Early sensitivity to mental states: prior neural and behavioral evidence

In adults, the TPJ, PC, and MPFC are functionally correlated with one another and show strong preferential responses while thinking about others' minds. How early in children's development are these cortical signatures of ToM present?

Initial fMRI studies of children revealed that the ToM network is present as early as 5 years of age. Children listened to stories or viewed sequential pictures that evoked reasoning about mental states or about nonmentalistic control content while undergoing fMRI. These initial studies found that 5 to 12-year-old children, like adults, recruit bilateral TPJ, PC, and MPFC preferentially to reason about the minds of others ([Gweon et al., 2012](#); [Kobayashi et al., 2007](#); [Saxe et al., 2009](#)).

A key open question concerned the origins of the ToM network in children younger than 5 years. Before age 5, children appear to undergo dramatic changes in their understanding of others' internal mental lives. By far the most studied transition in children's ToM is the ability to understand that another person's beliefs about the world can be false ([Premack and Woodruff, 1978](#); [Wimmer and Perner, 1983](#)). The standard version of the false belief task involves a scenario in which a character (Sally) has a false belief about the location of an object (a ball), because she was not present when it was moved. Participants must form and access a representation of Sally's (false) belief in order to answer the key test question: Where will Sally look for her ball? Across hundreds of behavioral studies using variants of this task, children perform at chance on test questions that require representing others beliefs until approximately age 4 years ([Wellman et al., 2001](#)). This robust result led to the hypothesis that the transition from failure to success on false belief tasks reflects an important conceptual leap in ToM development ([Callaghan et al., 2005](#); [Perner et al., 1987](#); [Wimmer and Perner, 1983](#)): perhaps children gain insight into the representational nature of others minds around 4 years of age. Thus, one possibility was that the cortical ToM network might also undergo dramatic changes around age 4 years.

On the other hand, the claim that cognitive ToM changes dramatically around age 4 years has been questioned ([Bloom and German, 2000](#); [Carlson and Moses, 2001](#); [Carlson et al., 1998](#)). In particular, standard false belief tasks might underestimate younger children's cognitive sophistication, because of other task demands. Studies using nonlinguistic eye-tracking and interactive, naturalistic helping paradigms have suggested that toddlers as young as 2 years of age make action predictions based on others beliefs ([Buttelmann et al., 2009](#); [Southgate et al., 2007](#); [Surian et al., 2007](#); [Knudsen and Liszkowski, 2012a,b](#); [Priewasser et al., 2017](#)). If 18-month-olds are already sensitive to others' (false) beliefs in their anticipatory looking and helping behaviors, then it is plausible that cortical regions supporting ToM are already supporting these abstract cognitive functions in toddlers, or even earlier.

21.2 Early sensitivity to mental states: neuroimaging studies of young children and infants

Thus, to address important open questions about ToM brain development, it was necessary to conduct neuroimaging studies of ToM network activity in children younger than 5 years. This created a substantial technical hurdle. Traditional fMRI studies are challenging for young children and infants: fMRI measurements require participants to stay very still and attend to stimuli for an extended period of time.

A key challenge since these initial studies has been to design experimental paradigms and use measurement tools that are appropriate and feasible for measuring ToM responses in young children and infants. In the past 5 years, methodological advances have enabled studies to directly measure responses in ToM brain regions in early childhood and infancy. These studies have begun to provide a much needed link between the wealths of knowledge about early behavioral ToM development and the mature neural profile of ToM reasoning. Additionally, these studies provide compelling evidence that ToM brain regions operate together as a functional network and show preferential responses for processing mental states early in development.

In order to acquire high-quality fMRI data from young children, fMRI studies of ToM have begun to employ naturalistic movie-viewing paradigms. Unlike traditional fMRI experiments, which involve measuring responses to multiple trials of stimuli designed to isolate and evoke specific cognitive processes, naturalistic movie-viewing paradigms measure neural responses to a single narrative over time. Movie-viewing paradigms are engaging for even young children (Cantlon and Li, 2013), and reduce participant motion relative to traditional experiments (Raschle et al., 2009; Vanderwal et al., 2015), thus expanding the range of children who can participate in fMRI experiments. For example, Richardson et al. (2018) recorded fMRI responses from 122 children while watching Disney Pixar's *Partly Cloudy* (Reher and Sohn, 2009), a 6-minute wordless animated short film. Sixty-five of these children were aged between 3 and 5 years (Richardson et al., 2018). Similarly, Moraczewski et al. (2018) measured responses from 40 four-year-olds, in addition to 6-year-olds and adults, while watching a clip from Disney Pixar's *Toy Story* (Guggenheim et al., 1995; Moraczewski et al., 2018). These two studies were the first to use fMRI to measure ToM-related activity in children this young.

On the other hand, movie-viewing paradigms create new challenges for fMRI data analysis and interpretation: How can we recover or infer the specific cognitive processes evoked over time in the movie? One strategy is to simply measure the response of ToM regions in adults. Thirty-three adults watched *Partly Cloudy* (Reher and Sohn, 2009), and their ToM network showed a highly reliable pattern of response over the 6-min movie. The plot of *Partly Cloudy* depicts the relationship between two main characters: Gus, a cloud who makes babies, and Peck, a stork who delivers those babies to their parents. Gus specializes in dangerous babies: porcupines, sharks, alligators, and so on. Peck is often injured while delivering these babies. A key misunderstanding occurs when Peck leaves Gus for another cloud: Gus initially feels abandoned and betrayed, but is relieved when Peck returns, revealing protective gear. Brain regions with highly correlated activity during movie-viewing are driven and deactivated by similar moments in the movie—meaning their functional response profile is similar. In adults, TPJ, PC, and MPFC regions show distinctive peaks of activity at moments in the movie that focus on the emotional feelings and misunderstandings of Gus and Peck, and troughs of activity during intervening scenes that depict physical events like Gus creating babies and the physical attacks on Peck (Richardson et al., 2018).

The observed response in adults' ToM network provides a template against which we can compare children, even without knowing the precise cognitive processes evoked at each moment in the movie. Children's ToM network showed a similar timecourse of response to the movie. In 5-, 4-, and even 3-year-old children, the ToM network timecourse was significantly correlated with the ToM network response in adults, and not with the response of other control networks. That is, regions of the ToM network are preferentially active during the same scenes in the movie as adults' ToM network (Richardson et al., 2018). These results suggest that at least somewhat similar cognitive processes are evoked by the movie, and supported by the ToM network of cortical regions, in adults and in children as young as 3 years of age.

What about children even younger than 3 years old? Task-driven fMRI measurements are exceptionally difficult to acquire in infants and toddlers, but a more appropriate neuroimaging method is functional near infrared spectroscopy (fNIRS). fNIRS, like fMRI, measures the relative amounts of oxygenated and deoxygenated blood in the brain, but uses a cap with optodes that emit and detect light, rather than an MRI machine. Because the cap and optodes are anchored to the participant's scalp, fNIRS is more tolerant to participant motion than MRI (though methods for identifying and correcting for motion artifacts remain critical (Cooper et al., 2012)). As such, fNIRS affords easier and more efficient measurement of neural responses in infants (Lloyd-Fox et al., 2010). However, fNIRS has a limited field of view compared to fMRI, allowing researchers to measure activity in a limited number of superficial cortical regions, and creates a challenge for identifying the precise cortical source of any observed activity.

Hyde et al. (2018) chose to test the specific hypothesis that RTPJ in particular is preferentially active while thinking about others' thoughts, even in young infants (Hyde et al., 2018). Participants watched simple sequences of actions: a target person placed an object in a container, and later came back to retrieve the object. In critical (false belief) trials, the object's location was changed while the target person was not looking. In a prior study of adults, Hyde and colleagues found that the RTPJ showed increased blood oxygenation specifically when the target with the false belief returned to retrieve the object (compared to other times during the event, and to other conditions that did not involve a false belief; Hyde et al., 2015).

Hyde et al. used three approaches to identify responses in RTPJ in infants, with fNIRS. First, they used an existing infant scalp fNIRS-MRI atlas (Lloyd-Fox et al., 2014) to guide probe placement, such that channels plausibly recorded from RTPJ. Second, they conducted photo migration simulations to estimate the cortical sensitivity profile of each channel, produce a sensitivity map based on these estimations, and estimate coordinates for the center of sensitivity for each channel. They subsequently compared channel locations to infant scalp surface maps (Lloyd-Fox et al., 2014), as well as to RTPJ coordinates reported by a prior fMRI study in adults (Deen et al., 2015). Finally, they implemented a "functional channel of interest" (fCOI) approach (Powell et al., 2017), which involves an iterative process of using a subset of trials to select channels that responded preferentially to the false belief condition relative to the true belief and direct perception

conditions (on average), and extracting the response from that channel in independent trials. This procedure selects channels based on their functional responses (false belief > true belief/direct perception), rather than relying on reliable channel placement on the scalp across participants.

Using this approach, in two samples of 7-month-olds (total $n = 50$) Hyde et al. observed preferential increases in oxygenated blood over the RTPJ, specifically during false belief sequences, and specifically when the target person returned to retrieve the object. That is, the RTPJ response in 7-month-olds was strikingly similar to the RTPJ response in adults!

Thus, aggregating the results of these very different studies, cortical regions of the ToM network appear to be preferentially engaged by similar stimuli in adults, 3-year-old children, and 7-month-old infants. These results point to a very early developmental origin of a cortical network for ToM.

Although no other studies have directly measured preferential functional responses during ToM in very young children or infants, other approaches hint at the same conclusion. Rather than measuring stimulus evoked responses in the ToM network, another strategy is to simply measure the correlation in responses, over time, between the different regions in the ToM network. Interregion correlation analyses can be used to characterize functional networks in the brain, because brain regions that operate within a functional network are more highly correlated with one another than with brain regions outside of that functional network (Blank et al., 2014; Hasson, 2004). During the *Partly Cloudy* movie, responses in ToM brain regions were significantly more correlated with one another than with other regions, even among 3-year-old children (Richardson et al., 2018).

Brain regions that are correlated during cognitive tasks are also correlated at rest (i.e., in absence of stimuli; Cole et al., 2014; Greicius et al., 2003; Miall and Robertson, 2006). Unlike preferential responses to specific stimuli, correlations in functional network properties can be measured with fMRI in toddlers and infants, because they can be measured during sleep. Studies that take this approach often use dimensional reduction techniques, like independent component analysis (ICA), to identify brain regions with similar (correlated) timecourses of activity. Using ICA, Gao et al. (2009) found similar response timecourses in MPFC and PC regions in 1-year-old infants, and MPFC, PC, and TPJ responses were correlated by age 2 years (Gao et al., 2009). An alternative method to ICA is to measure the correlation between responses in a particular “seed” region and spheres that surround every other voxel ($2 \times 2 \times 2$ mm cube) in cortex. For example, using the PC as a seed region, Gao et al. (2013) found correlated responses in MPFC and TPJ in 1- and 2-year-olds (Gao et al., 2013). Results from resting state connectivity studies of 1-month-old neonates and term infants have provided some evidence for correlated activity in MPFC and PC (Smyser et al., 2010; though see Gao et al., 2009, 2013); evidence for connectivity between these regions in preterm infants is sparse (Cao et al., 2016; Doria et al., 2010; for review, see van den Heuvel and Thomason, 2016). Overall, these results converge to support the hypothesis that the ToM network is functionally correlated and distinct from other cortical networks very early in development.

Evidence from studies of macaques provides an additional indirect source of evidence about the early origins of social brain regions. Sliwa and Freiwald (2017) used fMRI to measure neural responses in the brains of macaques while they watched short videos of two monkeys interacting (grooming, fighting, chasing, and mounting). Control condition videos showed a single monkey interacting with objects (food, toys), two objects interacting (bouncing against each other, getting entangled), or a single inactive monkey (lying down or sitting quietly) or inactive object (hanging). Functional regions in the MPFC and superior temporal sulcus (STS), a region adjacent to temporoparietal junction and involved in social perception in humans, were selectively recruited when macaques watched clips that involved social interactions between two monkeys. These regions responded to social interaction clips and deactivated to all other kinds of clips. Regions that responded preferentially to social interactions were functionally correlated, and more correlated with one another than with other regions (Sliwa and Freiwald, 2017). Given the similar location and functional response profile of these regions, the STS and MPFC in macaques may very well be functional homologs of the ToM network in humans. Future comparative work is necessary to better understand the similarities and differences of the functional roles of these homologous regions across species, to characterize the ontogenetic development of these regions in nonhuman primates (Rosati et al., 2014), and to link neural measures of social brain regions to the vast behavioral literature on social understanding and ToM in nonhuman primates (Drayton and Santos, 2016; Martin and Santos, 2016).

In all, by the most basic and general definition, the ToM network appears to emerge ontogenetically and possibly phylogenetically early. TPJ, MPFC, and PC are preferentially recruited by stimuli that evoke consideration of others’ thoughts and feelings. These regions are functionally correlated with one another, and thus constitute a distinct functional network.

21.3 Neural correlates of ongoing theory of mind development in childhood

The recent imaging studies reviewed above have begun to fill in the gap between behavioral studies on ToM development in early childhood and infancy, and fMRI studies of ToM brain regions in older children and adults. The results of these studies are generally consistent with the behavioral literature on ToM development, which increasingly places false belief

task performance in the context of gradual, continued ToM development throughout childhood. The ability to represent others' beliefs is preceded by the ability to represent others' goals and knowledge access (Wellman and Liu, 2004), and followed by the ability to reason about nonliteral speech (Filippova and Astington, 2008; Peterson and Wellman, 2018; Peterson et al., 2012; Wellman et al., 2011) and evaluate moral blameworthiness in accidents (Cushman et al., 2013). A fully developed Theory of Mind incorporates each of these abilities, and as such, Theory of Mind development encompasses more than the achievement of one of them. Analogously, ToM brain regions show early signatures of functional organization prior to explicit false belief task success, but should also continue to undergo developmental change throughout childhood. How do the functions and computations of ToM brain regions change to support ongoing ToM development? Below, we review three kinds of neural measures that capture developmental change in ToM brain regions in childhood and correlate with behavioral improvements in Theory of Mind, cross-sectionally: response selectivity, response reliability, and functional network segregation.

21.3.1 Response selectivity: fine-tuning preferential responses

Given the highly selective functional responses observed in ToM brain regions among adults, initial fMRI studies on ToM in children were designed specifically to test hypotheses about the development of selective responses. In other domains, the development of increasingly selective functional responses supports cognitive change in childhood. For example, the magnitude of selective responses in cortical regions specialized for faces (fusiform face area, FFA) and places (parahippocampal place area, PPA) is larger in adults, relative to children (despite similar whole-brain volume), and is correlated with behavioral recognition memory of faces and scenes (Golarai et al., 2007; Gomez et al., 2017; Grill-Spector et al., 2008). These cortical regions respond less to nonpreferred categories with age, and the reduced response to nonpreferred categories corresponds to improvements on category-relevant behavioral recognition tasks (Cantlon et al., 2010). Similarly, the volume of symbol selective cortex (visual word form area (VWFA); McCandliss et al., 2003) increases rapidly in children upon learning to read (Dehaene-Lambertz et al., 2018), and development of this region involves reduced responses to nonpreferred visual categories (Dehaene et al., 2010). Increasingly selective cortical responses could reflect or support the specialization of a region for the particular computational demands of a particular cognitive domain.

Thus, in addition to measuring neural responses to mental states and nonmentalistic control conditions, two of the early studies of ToM in children (5–12 years old) measured responses to nonmentalistic social stimuli, which included information about the characters' stable relationships ("Sarah and Lori played together on the school soccer team ..."; "Jenny and Samantha were twins ..."), physical appearances ("Old Mr. McFeegle is a gray wrinkled old farmer..."), and actions and abilities ("Once upon a time, a girl and her little brother went out to pick flowers ..."; "... She was so good at playing the flute that when she played everyone immediately started dancing"). By measuring responses to social, but nonmentalistic, stimuli, these two studies found evidence for developmental change in ToM responses: responses in ToM brain regions, and in particular in RTPJ, became more selective for mental states between 5 and 12 years of age (Gweon et al., 2012; Saxe et al., 2009). Responses to mental state stories remained high throughout childhood, but responses to non-mentalistic social control stories decreased with age. Additionally, selectivity for mental states in the RTPJ correlated with behavioral performance on a ToM task that involved reasoning about similar/diverse desires, true/false beliefs, moral blame, and mistaken referents (Gweon et al., 2012).

Richardson et al. (2018) provides converging evidence for developmental change in response selectivity throughout childhood. In addition to conducting interregion correlation analyses (reviewed above), Richardson et al. conducted reverse correlation analyses on ToM responses in adults ($n = 33$) and in 3-year-old children ($n = 17$). By identifying timepoints that reliably evoke positive responses in a brain region or functional network across participants, reverse correlation analyses offer a data-driven way to identify the kind of stimulus that drives responses in that region or network. In adults, reverse correlation analyses have offered satisfying replications of functional selectivity profiles initially discovered with well-controlled, experimenter-driven methods: the fusiform face area is driven by faces, and the parahippocampal place area is driven by scenes (Hasson, 2004). Similarly, Richardson et al. found that the pain network (a set of brain regions recruited to reason about physical/bodily pain) is driven by scenes involving salient physical sensations, and the ToM network is driven by scenes that evoke mental state reasoning among adults. Scenes that evoked responses in the pain network were nonoverlapping with those that evoked responses in the ToM network. A majority of the scenes identified as pain events (9/12) and all of the ToM scenes (7/7) were identified in a replication analysis of a second (independent) sample of adults (Richardson et al., 2018).

As a data-driven approach, reverse correlation analyses aren't constrained by experimenter-constructed hypotheses, and therefore could discover preferential responses for kinds of stimuli that are unexpected, especially in participants with different or developing cognitive abilities. Do the same ToM events evoke reliably high responses in ToM brain regions

among 3-year-old children? As in adults, scenes that evoked responses in the pain network were nonoverlapping with those that evoked responses in the ToM network in 3-year-old children. However, only 8 timepoints (of 168) reliably evoked positive responses in the ToM network across 3-year-old participants. Of these 8 timepoints, half corresponded to an adult ToM event; the other half corresponded to an adult pain event. In the entire cross-sectional sample of children ($n = 122$), ToM network responses to ToM events increased and ToM network responses to pain events decreased between 3 and 12 years of age. Thus, while ToM brain regions are functionally distinct from other networks and show some evidence for early preferential responses (i.e., by being significantly correlated with adult responses), responses become more selective for mental state content throughout childhood.

Together, these studies provide evidence for slow, gradual development of increasingly selective brain regions for ToM. However, there is a small discrepancy in these results. In the initial fMRI studies (Gweon et al., 2012; Saxe et al., 2009), developmental change in response selectivity manifested as stable, high responses to mental state stimuli paired with decreasing responses to nonmentalistic social stimuli. In Richardson et al. (2018) responses to nonmentalistic pain events decreased with age, and responses to mental state events increased. We suggest that this discrepancy may be explained by another aspect of the response in ToM brain regions that changes with age in childhood—spontaneous responses in ToM brain regions to naturalistic stimuli become increasingly reliable and stereotyped. In Richardson et al. (2018) developmental increases in ToM responses to mental state events could reflect reduced variability in the timing and duration of the responses to these events. Consistent with this idea, functional responses in the ToM network become more similar to the average adult response with age. We review additional evidence for developmental change in the reliability of spontaneous ToM responses below.

21.3.2 Reliable spontaneous (uninstructed) responses to movies

Moments or situations that require or would benefit from ToM reasoning aren't always marked. Individuals may have different thresholds for spontaneously engaging in mental state reasoning, and this could contribute to individual differences in ToM. For example, children who spontaneously produce mentalistic explanations for behaviors subsequently achieve consistent successful performance on false belief tasks (Amsterlaw and Wellman, 2006). These individual differences could reflect an individual's propensity to shift their focus to mental states (reliably and quickly), anticipate or predict future mental states, and/or interpret current events in the context of past mental states. Univariate measures of response magnitude or selectivity are limited in their sensitivity to this kind of individual difference, because they average responses across time and trials to calculate a single binary contrast. By contrast, timecourse analyses of data collected during naturalistic movie paradigms can capture individual differences in responses to a continuous, dynamic narrative that unfolds over time (Dubois and Adolphs, 2016b).

Similarity of response timecourses across individual participants can be quantified via intersubject correlation (ISC) analyses (Hasson, 2004). Among adults, ISC analyses find highly similar responses in ToM brain regions during movie-viewing or narrative-listening (Hasson, 2004; Jääskeläinen et al., 2008; Lerner et al., 2011; Nummenmaa et al., 2012; Wilson et al., 2007), as well as during free recall of a movie (Chen et al., 2016). Identifying response profiles that are generally common across individuals potentially enables discovering meaningful differences in individuals whose response timecourses vary. Indeed, adults with more correlated responses in ToM brain regions during movie-viewing have similar interpretations of the movie (Nguyen et al., 2017), and similar memories about the movie (Chen et al., 2016). Thus, ISC analyses are a promising approach for capturing meaningful neural individual differences in ToM responses to naturalistic viewing or listening paradigms (Dubois and Adolphs, 2016b).

Can naturalistic movie-viewing paradigms be used to capture developmental change in ToM? One way to measure individual differences in ToM responses to naturalistic paradigms across children is to compare each child's timecourse to an adult template time course. This measure is sometimes referred to as "functional maturity," as it quantifies the similarity of the functional profiles between children and mature adults. By calculating this measure over responses to naturalistic stimuli, this measure likely reflects developmental change in both response selectivity to functionally preferred content and developmental change in spontaneous, unprompted cognitive computations. In other cognitive domains, functional maturity is correlated with cognitive abilities. For example, Cantlon and Li (2013) calculated functional maturity in 4- to 6-year-old children by comparing their response timecourses to an episode of *Sesame Street* to an average adult timecourse. Among children, functional maturity of the intraparietal sulcus, a cortical region that responds preferentially to process numbers, correlated with behavioral performance on a math assessment, and functional maturity of Broca's area, a cortical region involved in language processing, correlated with children's verbal abilities (Cantlon and Li, 2013). While some aspects of the developmental increase in functional maturity could be very general, i.e., reflective of more consistent and adultlike eye movements during naturalistic viewing tasks (Franchak et al., 2016; Frank et al., 2011), correlations between

functional maturity in functionally selective cortical regions and performance on relevant behavioral tasks suggests that this measure captures domain-specific development as well.

Two studies have measured functional maturity of ToM brain regions in children. In [Richardson et al. \(2018\)](#), reviewed above, functional maturity of the ToM network increased with age: responses in these brain regions to *Partly Cloudy* became more adultlike between ages 3 and 12 years ([Richardson et al., 2018](#)). While functional maturity (of the entire ToM network) correlated with performance on a behavioral ToM task, this correlation did not remain significant when additionally controlled for age. In the second study, [Moraczewski et al. \(2018\)](#) measured response timecourses in 4- and 6-year-old children, as well as adults, while they watched a 6-min clip from the movie *Toy Story* ([Guggenheim et al., 1995](#)). Responses in ToM brain regions to *Toy Story* were more similar across individual adults, than across individual children (that is, the adult response profile was more stereotyped or reliable, across subjects). Additionally, responses in TPJ to *Toy Story* became more similar to that of adults between 4 and 6 years of age ([Moraczewski et al., 2018](#)). While behavioral ToM was not measured in this sample, [Moraczewski et al. \(2018\)](#) confirmed their results in subsequent analyses that used response timecourses from the frontal eye field as a nuisance regressor, in order to account for individual differences in attention. Nonetheless, additional work is necessary to better understand the relationship between functional maturity and individual behavioral differences in ToM.

Like traditional fMRI experiments, naturalistic movie-viewing paradigms allow for measurements of neural responses to particular events. However, because events are embedded within the movie, naturalistic movie-viewing paradigms are additionally sensitive to individual differences in anticipation of future mental states, and (re-)interpretation of current events in the context of past mental states. Spontaneous activity in response to specific events whose meanings depend on integration over long periods of time may be particularly relevant to individual differences in ToM. In [Richardson et al. \(2018\)](#), multiple measures of the response in ToM brain regions correlated with behavioral ToM score (proportion correct on questions that involved reasoning about same/different desires, true/false beliefs, moral blame, sarcasm, lies, second-order false beliefs, and mistaken referents). However, only the magnitude of response to a particular event during *Partly Cloudy* ([Reher and Sohn, 2009](#)) correlated with ToM score when additionally controlling for age. As reviewed above, similar response timecourses in adults reflect similar narrative interpretations of movies or stories. Thus, by looking at the content of the event that evoked ToM responses differentially by ToM behavioral score among children, we can hypothesize about the kinds of differences in narrative interpretations of that event that might be relevant to ToM development.

The event in question involved one character (Peck, the stork) offering an alternative explanation for his previous behavior (flying away) to another character (Gus, the cloud), who had not only made an incorrect inference about the cause of the behavior (Peck was abandoning him in favor of working with a different cloud), but who had also been very upset about this inferred cause. Based on the content of this event, the correlation between responses during this event and behavioral ToM may reflect participant's (re-)consideration of mental states in absence of explicit cues to do so. This kind of brain-behavior correlation has since been replicated in confirmatory analyses of an independent sample of 5- to 12-year-old children ($n = 186$; [Richardson, 2019](#)). Thus, more reliable response timecourses in ToM brain regions with age may reflect developmental improvements in spontaneous, flexible engagement of mental state reasoning. Spontaneous, flexible mental state reasoning may be directly tied to the development of increasingly refined ToM concepts, which could enable children to more easily recognize the relevance of particular ToM concepts across different contexts and without cues ([Amsterlaw and Wellman, 2006](#); also see [Moreira et al., 2018](#) for a relevant study in adolescents).

21.3.3 Integration and separation of functional networks

A third kind of neural change that may support ongoing ToM development is ongoing separation of functional networks. Functional network separation can be measured via interregion correlation analyses (reviewed above), which compare response timecourses within individuals across different brain regions ([Blank et al., 2014](#); [Hasson, 2004](#)). Brain regions with highly correlated activity during movie-viewing are driven and deactivated by similar moments in the movie, i.e., their functional response profile is similar, and they support a specific set of cognitive functions. Naturalistic movie-viewing and story-listening experiments have been used to characterize functional networks for processing faces and scenes ([Hasson, 2004](#)), and for dissociating brain networks for processing language versus domain-general cognitively difficult tasks (the “multiple demand” network; [Blank et al., 2014](#)).

These kinds of experiments have begun to characterize the functional network properties of ToM brain regions. Among adults, ToM brain regions are more correlated with one another than with brain regions involved in processing language ([Paunov et al., 2017](#)) and brain regions that process physical (bodily) pain ([Richardson et al., 2018](#)). Interestingly, ToM and pain brain regions are not just uncorrelated, but they are actually *anticorrelated*, among adults. As reviewed above, ToM

brain regions are functionally dissociated from the pain network by age 3 years. However, the extent of this dissociation increases with age throughout childhood—ToM brain regions become more correlated with one another (see also [Xiao et al., 2019](#)), and increasingly *negatively* correlated with brain regions in the pain network, with age ([Richardson et al., 2018](#)). These developmental trends were replicated in an independent sample of children and adolescents ([Richardson, 2019](#)).

Is the developmental separation of functional networks a plausible neural correlate of improvements in ToM behavior? Increases in within-network correlations could reflect faster, more efficient, and less noisy communication between ToM brain regions. Similarly, decreases in across-network correlations could reflect less interference, or more functionally precise responses, across networks. [Richardson et al. \(2018\)](#) collected behavioral ToM data and tested for neural measures that (1) differed between young children ($n = 65$, 3–5 years old) who passed and failed explicit false belief tasks, and (2) increased with overall ToM reasoning (including questions about concepts that develop prior to and after passing false belief tasks) in the full sample of children ($n = 122$, 3–12 years old). Interestingly, ToM brain regions were significantly more correlated with one another in children who passed explicit false belief tasks, relative to those who failed, controlling for age. However, this measure also increased with age in the full sample of children. In the full sample of children, within-ToM network correlations correlated with overall ToM performance, but this correlation did not survive controlling for age. Thus, one possibility is that within-ToM network correlations are particularly important for or reflective of ToM behavior in young children, but more correlated with age across development.

These results are consistent with two studies that used two different neural measures of within-network integration. First, [Grosse Wiesmann et al. \(2017\)](#) used diffusion tensor imaging (DTI) to measure white matter connections—rather than functional responses—in the brain, to study ToM development in young children ([Grosse Wiesmann et al., 2017](#)). [Grosse Wiesmann et al.](#) found increased connectivity in tracts surrounding ToM brain regions (including RTPJ, PC, and VMPFC) in 3- and 4-year-old children who passed false belief tasks, relative to those who failed. Second, [Xiao et al. \(2019\)](#) found that resting state connectivity between RTPJ and other ToM ROIs (LTPJ, PC) correlated with scores on a parent-questionnaire assessing ToM abilities in 4- to 8-year-old children. While both studies found neural differences by ToM behavior that remained significant while controlling for age, future work with a wider and older age range is necessary to determine if these neural measures are especially important for early ToM development, or if ongoing developmental change in these measures supports subsequent ToM advances. In any case, it is encouraging to see consistency in results across three types of neural measures (interregion correlations measured during movie-viewing and at rest, and physical connectivity measures).

In sum, imaging studies of children provide evidence for ongoing development in ToM brain regions throughout childhood. While ToM brain regions show early preferential responses and early segregation from other functional networks, the responses in these brain regions become increasingly selective and fine-tuned for processing mental states. Responses in these regions also appear to converge with typical adult responses, and diverge from responses in other functional networks, with age. Each of these neural measures appears to relate to performance on ToM tasks in childhood. Of course, these measures are not independent from one another—they each capture a slightly different aspect of the same functional response. Future work is necessary to understand the relationships between each measure. For example, what is the causal relationship between the development of selective responses, and the development of segregated functional networks? Does intrinsic or functional connectivity restrict, predict, or guide physical connectivity, or vice versa? Understanding how these measures relate to one another in development may clarify their relative contributions to behavioral ToM development.

21.4 Future directions: open questions and challenges

21.4.1 Neural correlates of structural changes in theory of mind

As children get older, they not only acquire and refine ToM concepts, but they also achieve a more sophisticated understanding of the relationships between different mental states (beliefs, desires, emotions), and the relationship between mental states and their causes and consequences. While children show relatively early understanding of probable causes of different emotions (surprise, sadness, awe) ([Lagattuta and Wellman, 2001](#); [Skerry and Spelke, 2014](#); [Wu et al., 2017](#)), there is also evidence for ongoing change in the conceptual organization of emotions in childhood ([Gao and Maurer, 2010](#); [Weisman et al., 2017](#); [Widen and Russell, 2003, 2008](#)). One challenge for future neuroimaging studies of ToM development will be to discover neural correlates of this kind of conceptual change. Traditional measures of selectivity—which measure the relative response of a given brain region to different conceptual categories (e.g., mental states vs. bodily sensations)—may not be sensitive to structural changes *within* the category of mental states. That is, because mental state stimuli evoke high responses in ToM brain regions, selectivity may not be a sensitive neural measure of the development conceptual distinctions *between* different mental states within the broader category. Indeed, among adults, different kinds of mental states (e.g., expected vs. unexpected) evoke the same magnitude of response in ToM brain regions ([Young et al., 2010b](#)).

Multivariate pattern analyses may provide a method for capturing developmental change in the structure of mental state representations. In adults, multivariate analyses have revealed information about the structure of representations within ToM brain regions by determining the features of mental state stimuli that drive response similarity. For example, [Tamir et al. \(2016\)](#) used multivariate analyses to suggest that almost half of the variation in responses in ToM brain regions to mental states can be accounted for collectively by three abstract dimensions: rationality, social impact, and valence ([Tamir et al., 2016](#); see also [Saxe, 2018](#)). To discover these dimensions, Tamir and colleagues first used principal component analysis (PCA) on behavioral ratings of 166 mental states (e.g., “embarrassment,” “ecstasy,” “planning,” “dominance,” “friendliness,” “imagination,” “self-pity,” “satisfaction,” “affection,” “disgust,” “disarray”) on 16 plausible dimensions derived from the literature, collected in a large sample of adults ($n = 1205$). They found that behavioral rating similarity across mental states was driven by four orthogonal dimensions: rationality, social impact, human mind, and valence. Next, they used fMRI to measure responses in an independent sample of adults ($n = 20$) who considered and rated the extent to which a given scenario (e.g., “finding \$5 on the sidewalk”, $n = 16$ unique scenarios total) evoked a particular mental state (see examples above; $n = 60$ unique mental states total). They used a “searchlight analysis” to examine the relative response surrounding every voxel of cortex to each mental state. The extent to which mental states were similar in rationality, social impact, and valence ratings predicted the extent to which mental states evoked similar neural responses in social brain regions. Thus, the same ToM brain regions implicated by univariate analyses contain multidimensional representations of mental states, and rationality, social impact, and valence may be particularly important dimensions by which responses in these regions are organized.

Multivariate analyses can additionally provide evidence about the different functional roles of distinct social brain regions. [Koster-Hale et al. \(2017\)](#) used fMRI to measure neural responses in adults who listened to short stories that involved mental state reasoning. Across stories, characters formed beliefs that differed in justification (i.e., based on strong vs. weak evidence), source modality (i.e., based on hearing vs. seeing evidence), and valence (positive vs. negative). Koster-Hale and colleagues found that the epistemic features— that is, features that described the source and evidence for the belief— organized responses in temporoparietal junction, whereas the valence of the resulting emotion (caused by the belief) organized responses in medial prefrontal cortex. Thus, multivariate methods can be used to discover the divisions of labor among ToM brain regions, in addition to the features that organize neural responses, and the underlying representations, within each region. Developmental change in the structure of ToM representations could be reflected in developmental change in the features that drive response similarity in ToM brain regions. Additionally, developmental change in the structure of ToM representations could result in divisions of labor among ToM brain regions, by which certain brain regions are increasingly fine-tuned to particular aspects of mental states. Studies that use multivariate analyses to characterize ToM responses in child populations are an important next step in the developmental cognitive neuroscience of ToM.

21.4.2 Discovering reliable neural markers of individual differences in theory of mind

Perhaps the greatest challenge faced by developmental cognitive neuroscientists studying ToM is to discover robust neural markers that reliably measure individual differences in ToM ([Dubois and Adolphs, 2016a](#)). To date, there is no such evidence for a neural marker that reliably predicts individual differences in ToM longitudinally, despite multiple longitudinal behavioral studies suggesting that there are reliable individual differences in ToM behavior in development ([Peterson and Wellman, 2018](#); [Wellman et al., 2011](#); [Wellman et al., 2008](#); [Yamaguchi et al., 2009](#)). Reliable neural individual difference measures would be useful for testing hypotheses about environmental and genetic factors that promote or hinder development of ToM brain regions, and are critical for designing and testing the effectiveness of clinical interventions that aim to improve social cognitive abilities. While many studies find group differences between older children and adolescents who have been diagnosed with autism or schizophrenia and neurotypical individuals, none of these neural differences are sensitive or early markers that could be used in a clinical or treatment setting (studies of clinical populations are not reviewed in this chapter; see [Happé and Conway, 2016](#); [Happé and Frith, 2014](#)). Longitudinal evidence is necessary for testing candidate neural markers: such a marker should show reliable neural individual differences over time that relate to ToM behavior, and early measurement of the marker should predict subsequent ToM development and ability.

Neuroimaging evidence from individuals with different developmental experiences is also useful for learning about drivers of individual differences in ToM, and about the impact of experience on social and brain development more broadly.

21.4.3 The role of developmental experience: language

There is significant need for more research on the developmental factors that drive the specialization of brain regions for ToM. One open question concerns the role of experience on the maintenance and development of selective responses. By one extreme, evolutionary pressures and genetic makeup could drive ToM brain regions to develop increasingly

functionally selective responses throughout childhood, regardless of developmental experience. On the other extreme, the development of functionally selective responses could be quite fragile or flexible: lack of a necessary input or experience during a critical or sensitive time could preclude the development of brain regions selective for ToM processes. By this account, individual differences in the development of specialized responses would be a direct result of developmental experience.

What kinds of developmental experiences might be important for ToM development? Once again, behavioral studies of ToM development in children provide useful hypotheses. For example, language abilities are highly correlated with ToM performance in childhood (Astington, 2006; Astington and Jenkins, 1999; Milligan et al., 2007). One intriguing hypothesis is that this correlation reflects a facilitative role of linguistic experience for ToM development. That is, language might not only enable children to express their understanding of ToM concepts—linguistic experience might directly drive the development of ToM concepts per se. If linguistic experience is important for ToM-specific development (and not just for task performance), then ToM development should be delayed as a function of delayed access to language. Behavioral studies have provided evidence in favor of this hypothesis by measuring ToM performance in children who are d/Deaf.² Deaf children born to hearing parents (DoH) often experience delayed access to sign language, whereas d/Deaf children born to signing parents receive access to sign language at birth. DoH children show delayed ToM development corresponding to the length of delay prior to exposure to language, even on nonlinguistic ToM tasks (Meristo et al., 2012; Peterson and Siegal, 1999; Peterson and Wellman, 2018; Schick and Hoffmeister, 2001; Schick et al., 2007; Woolfe et al., 2002). Does delayed access to language result in delayed development of brain regions specialized for ToM?

A recent study by Richardson et al. (2019) used fMRI in conjunction with behavioral measures to ask this question. Native ($n = 21$) and delayed ($n = 12$) signing children (4–12 years old) completed linguistic and nonlinguistic behavioral tasks and fMRI experiments. Delayed signers experienced delays ranging from 0.25 to 7 years in length, but all children were proficient in American Sign Language (ASL) at the time of the study. Consistent with previous behavioral studies, children with delayed access to language showed corresponding delays on behavioral ToM tasks. These delays were most apparent in linguistic questions involving “advanced” ToM reasoning, like considering intended meaning in nonliteral speech (sarcasm), lies, and second-order false beliefs, and assignment of moral blame in accidents, as measured by the linguistic task. However, delayed signing children showed typical ToM performance on moral blame questions in the nonlinguistic behavioral task, and all children showed higher ToM performance on questions about false beliefs when tested linguistically. Behavioral differences across linguistic and nonlinguistic task formats suggest that language plays a role in the *expression* of ToM understanding. Does language additionally play a role in development of ToM per se?

The linguistic fMRI experiment provided a way to test this question. This experiment was modeled after the paradigm used in initial studies of ToM brain regions in 5- to 12-year-old children, and therefore was designed to measure the selectivity of ToM responses for mental states relative to nonmentalistic social information (Gweon et al., 2012; Saxe et al., 2009). As reviewed above, developmental specialization of ToM brain regions typically occurs via suppression of responses to nonmentalistic social content. In delayed signers, the selectivity of the response in RTPJ was delayed as a function of the length of delay prior to language exposure: the response to the nonmentalistic social condition was higher in children who experienced longer language delays. This effect was present despite proficiency in sign language at the time of the study, and no differences in the responses of brain regions involved in processing language. Thus, developmental experience, and in particular, early and extensive exposure to rich linguistic input, is important for refining the functional response of RTPJ in childhood. Importantly, typical response profiles were observed in delayed signing adults ($n = 16$), which suggests that delayed access to sign language delays, but does not permanently disrupt, the development of selective responses in RTPJ.

Interestingly, children with delayed access to language did not show differences in ToM responses during the nonlinguistic fMRI experiment, which involved watching a silent version of Disney Pixar’s *Partly Cloudy* (Reher and Sohn, 2009). As described above, this movie stimulus has previously been used to measure responses in ToM brain regions in children ages 3–12 years old (Richardson et al., 2018). As in the original study, ToM brain regions responded preferentially to mental state content and deactivated during moments depicting physical pain during this experimental context, regardless of age of exposure to language. Given that the functional dissociation between ToM brain regions and the pain network is apparent by age 3 years (Richardson et al., 2018), one possibility is that preferential responses in RTPJ for minds relative to bodies develop early and are less dependent on linguistic experience.

This study suggests that language is a key aspect of developmental experience that impacts ToM development. Is early linguistic input sufficient for typical development of brain regions specialized for ToM reasoning? Studies of children who

2. Note that in the literature, the capitalized form of the word “Deaf” has been used to refer to the cultural and linguistic minority group, and the lower-case “deaf” to refer to the audiological status.

are congenitally blind could help to address this question. Blind children have typical linguistic input, but reduced access to information about minds that is conveyed through vision. Vision provides a way to perceive consequences of mental states (e.g., if a person reaches for a teddy bear, she prefers it to the ball; if a person expresses sadness upon seeing a puppy, she's remembering when the puppy stole her snack), and facilitates early interactions and social bonding (e.g., through eye contact, joint attention, and attention to facial expressions). While previous neuroimaging research with adults suggests that by adulthood, blindness has no effect on the functional responses in theory of mind brain regions (Bedny et al., 2009; Koster-Hale et al., 2014), behavioral studies find some evidence for delayed ToM development in children who are blind (Brambring and Asbrock, 2010; Brown et al., 1997; Minter et al., 1998; Peterson et al., 2000). Future work investigating the development of functionally selective responses in children who are congenitally blind could clarify whether linguistic input is particularly important for refining functional responses, and if visual input during development plays a similar role in refining the functional responses in RTPJ.

21.4.4 The role of developmental experience: culture

A second way to study the impact of developmental experience on ToM development is to study children growing up in different cultures. Cultures provide social norms and instill individuals with a particular set of values. In adults, cultural differences appear to influence ToM reasoning: large-scale industrialized societies (e.g., Los Angeles, USA) place greater weight on intentions when assigning blame or punishment than small-scale traditional societies (e.g., Yasawa Island, Fiji; Barrett et al., 2016). Does culture also impact the development of ToM in children?

Initial studies of the influence of culture on ToM development compared ToM performance in children from WEIRD (Western, Educated, Industrialized, Rich, Democratic; Henrich, 2011) and non-WEIRD societies. Studies of non-WEIRD societies are extremely difficult and expensive to conduct, and can be risky in that the presence of the research itself could have unintended consequences for these societies. As such, there have been relatively few behavioral studies and, to our knowledge, no fMRI studies of ToM in individuals living in smallscale societies. Possibly because of the sparse number of studies, those that do exist focus primarily on false belief task performance: a reasonable initial step for understanding ToM reasoning in a different culture. Interestingly, these studies suggest that children from non-WEIRD societies, e.g., children of the Baka (pygmies living in southeast Cameroon rainforests; Avis and Harris, 1991), children living in an impoverished mountain village of Peru (Callaghan et al., 2005), and children from a Polynesian settlement in Samoa pass false belief tasks by age 5 years (Callaghan et al., 2005), similar to children in WEIRD societies (Wellman et al., 2001). Of course, false belief tasks may not be the most sensitive test of the impact of culture on ToM development. Additionally, the differences between the societies examined may not be relevant for false belief task performance.

Subsequent studies of the role of culture on ToM development suggest that by emphasizing or deemphasizing particular ToM concepts, cultures can influence the *trajectory* of ToM development; that is, the order in which children master particular concepts in ToM. For example, individualistic communities, like the United States of America or Australia, tend to value personal expression and independence, whereas collectivist or interdependent communities, like China, Iran, or Turkey, place more emphasis on shared knowledge. These different social norms and values highlight different ToM concepts for young children. Accordingly, American and Australian children understand that individuals can hold different beliefs earlier than Chinese, Iranian, and Turkish children, whereas Chinese, Iranian, and Turkish children understand the relationship between seeing and knowing earlier than Americans and Australians (Selcuk et al., 2018; Shahaecian et al., 2011; Wellman et al., 2006). The social norms of a culture can also shape how children seek out social information. For example, by 7 months of age, infants from Western (UK) versus Eastern (Japan) cultures fixate on different aspects of the face (mouth vs. eyes) to discriminate emotional expression (happy vs. fearful). These perceptual strategies map onto cultural differences in information value of mouths and eyes for communicating emotion (Geangu et al., 2016). Thus, culture shapes social input as well as the trajectory of ToM development in childhood. Does culture also influence the development of ToM brain regions, and if so, how?

The role of culture for the development of ToM brain regions may differ from the role of language. Multiple studies suggest that a person's native language does not affect the development of ToM brain regions. Individuals who are native English (Saxe and Kanwisher, 2003), Swedish (Happé et al., 1996), French (Brunet et al., 2000, 2003), German (Sommer et al., 2007; Vogeley et al., 2001), Japanese (Kobayashi et al., 2007), and American Sign Language (Richardson et al., 2019) speakers/signers recruit the same ToM brain regions during social reasoning tasks. Additionally, the semantic content, rather than the linguistic format, of stories drives response similarity in these brain regions, across adults (Honey et al., 2012). However, as reviewed above, children with delayed access to language show corresponding delays in the development of selective responses in the RTPJ (Richardson et al., 2019). Thus, while the differences in phonetic distinctions, grammatical structures, and modality that exist across languages do not appear to alter the location or

functional profiles of brain regions recruited for ToM, early exposure to language is important for the development of ToM brain regions. The provision of some set of norms and values that guide and help to understand social behaviors may similarly be important for the development of ToM brain regions. However, given the behavioral literature reviewed above, the particular norms and values that a culture highlights may also alter the developmental trajectory of ToM brain regions, or the particular features of mental states that drive responses in ToM brain regions.

What kinds of neural measures might be sensitive to cultural differences in ToM? Traditional univariate measures could reflect cultural differences if an experiment evokes a particular aspect of mental state reasoning that differs across the cultures studied. Additionally, naturalistic movie-viewing or narrative-listening paradigms could plausibly capture differences in narrative interpretation or perspective driven by cultural background. Among adults, similar response timecourses in ToM brain regions between speakers and listeners predict successful communication (that is, agreement about what was communicated; [Stephens et al., 2010](#)). Thus, differences in neural responses in ToM brain regions could indicate differences in culturally shaped perspectives. Multivoxel pattern analyses might also be useful for measuring differences in ToM reasoning driven by culture, because they could reveal differences in the features or dimensions of mental states that drive responses in ToM brain regions (e.g., degree of intent). Given evidence that culture shapes the *trajectory* of ToM development in children, studying children may be necessary to understand cultural differences in the trajectory of development of ToM brain regions ([Selcuk et al., 2018](#); [Shahaeian et al., 2011](#); [Wellman et al., 2006](#)).

There is strikingly little evidence about the development of ToM brain regions across cultures. [Lloyd-Fox et al. \(2017\)](#) used fNIRS to measure cortical responses to social stimuli (e.g., movies of adults playing Peek-A-Boo), compared to nonsocial control stimuli (e.g., images of cars and helicopters) in infants (0–24 months old) growing up in rural Gambia. Preferential responses in the right posterior temporal lobe (plausibly STS/TPJ) to social stimuli in these infants did not differ from those of children growing up in the UK ([Lloyd-Fox et al., 2017](#)). In addition to showing some preservation of preferential responses to social stimuli in the temporal lobe, this experiment demonstrates the feasibility of using fNIRS, which is less expensive and easier to transport than fMRI, to measure neural responses in children living in many diverse areas of the world. Future work is necessary to understand the degree of specialization of ToM brain regions in infants across varying cultures, and to characterize the impact of culture on the developmental trajectory of ToM brain regions in childhood.

Ultimately, studying the impact of culture on the development of ToM brain regions could refine theories about functional origins of these brain regions. For example, WEIRD adults recruit ToM brain regions when making moral judgments—e.g., when deciding how much blame to assign or how severely to punish an individual who has caused harm. In particular, the consideration of an individual’s intentions—whether she *meant* to cause harm—drives responses in RTPJ ([Koster-Hale et al., 2013](#); [Young and Saxe, 2009](#); [Young et al., 2010a,c](#)). But, if WEIRD people tend to focus on individual beliefs much more than people growing up with other cultural norms, then it seems likely that the root cognitive function of RTPJ (if there is one) is more general than thinking about others’ thoughts: i.e., the RTPJ could plausibly be specialized for considering any latent variables that people use to explain and evaluate others’ actions. These two hypotheses about the root function of the RTPJ can’t be teased apart in WEIRD adults, because individual beliefs *are* the latent variables WEIRD adults use to explain and evaluate others’ actions. However, studies of individuals who prioritize other information to understand actions (e.g., outcomes, rather than intent) could inform the extent to which ToM brain regions are specialized for beliefs per se, versus the function they serve for understanding others’ actions.

21.4.5 The role of family on ToM: shared environment and shared genes

Parents and siblings play a large role in shaping the smaller-scale culture of a child’s home life. Biological parents and siblings share their genes, in addition to shaping a child’s environment. As a result, it is sometimes difficult to tease apart the relative contributions of shared environment and genetic heritability when characterizing the role of family on social cognitive development ([Avinun and Knafo, 2014](#)). Indeed, some genes, like the oxytocin receptor gene, may impact ToM development in early childhood both through heritability and through their direct impact on social interactions ([Wade et al., 2015](#); [Wu and Su, 2015](#)). One way to isolate the impact of heritable genes, relative to shared or nonshared environmental factors, on ToM development is to study monozygotic and dizygotic twin pairs. While both kinds of twin pairs have similar cultural, home, and school environments across twins, monozygotic twins are genetically identical, whereas dizygotic twins share on average 50% of their genes. A handful of studies have now taken this approach in order to determine the relative influences of genes and environment on ToM. However, the methods and results of these studies are quite variable. In a study of 119 three-year-old twin pairs, [Hughes and Cutting \(1999\)](#) found that genetic influences accounted for 60% of the variance in performance on false belief and deception ToM tasks ([Hughes and Cutting, 1999](#)). In a second study of

1116 five-year-old twin pairs, heritable genes accounted for only 7% of the variance in performance on first and second-order false belief tasks, relative to 45% of the variance accounted for by nonshared environmental factors (e.g., siblings' different relationships with parents, with each other, and with peers, as well as child-specific life events like illness), and 28% by shared environmental factors (e.g., other siblings and SES-related factors). Similarity in ToM performance did not vary by zygosity of twin pairs in this sample (Hughes et al., 2005).

What could explain the discrepancy in these results? As the authors point out, one possibility is that the larger sample of the second study afforded the sensitivity to measure the effects of shared environment. A second possibility is that the relative impact of genes and environment changes over development, with genes having a larger impact on ToM in early childhood (Study 1, age 3 years), and environment playing a larger role in ToM development later on (Study 2, age 5 years). While the standard pattern in most aspects of development (McGue et al., 1993; Plomin and Kosslyn, 2001), including prosocial behavior (Knafo and Plomin, 2006), is that the role of genes *increases* with age, an initial drop followed by a rise in the role of genes on prosocial sharing and comforting behaviors has been observed in young children (Knafo et al., 2008). However, results from a longitudinal study of >1000 twin pairs studied at 2, 3, and 4 years of age argue against these two possibilities. This study, which was similar in sample size as the initial study, found a moderate role of genes, accounting for 25%–57% of the variance on a questionnaire-based assessment of ToM. Additionally, the role of shared environment on ToM tasks decreased with age (Ronald et al. (2005); though see Ronald et al. (2006) and Warrier and Baron-Cohen (2018), for evidence of little to no role of genes on ToM in 9- and 13-year-olds, respectively). A third possibility is that the inclusion of children from very low socioeconomic status (SES) families in the second study increased its sensitivity to the effects of shared environment. In other aspects of cognition, like IQ, shared environment is a stronger predictor than genes in impoverished families, whereas the opposite is true in affluent families (Turkheimer et al., 2003). Thus, the effects of shared environment on ToM may be stronger in samples that include more children living in disadvantageous environments. Future studies that take into account developmental changes with age, differences by ToM task, and gene × environment interactions (Knafo and Israel, 2009; Knafo-Noam et al., 2018) are necessary to clarify the relative role of genes on ToM development.

Could neuroimaging measures help to clarify the heritability of ToM in children? Among adults, there is some evidence that cortical responses in functionally selective brain regions are heritable. For example, the pattern of response to faces and scenes in the ventral visual cortex is more similar in monozygotic twins than in dizygotic twins (Polk et al., 2007). Thus, it is plausible that responses in ToM brain regions could similarly reflect the role of genes in ToM. Additionally, more similar responses in monozygotic twins compared to dizygotic twins in the response of the RTPJ would provide strong evidence that the genetic impact is on ToM per se, rather than domain-general abilities that are heritable and correlated with ToM (like language and executive functions), given the functional specialization of this brain region. By comparing RTPJ responses in twin pairs of different ages, this kind of research could also inform developmental changes in the relative roles of genes and environment on ToM. Finally, studies that measure correlations between particular genes and functionally selective responses could provide specific insight into how genes shape ToM development. In sum, neuroimaging measures may be a promising way to make progress in this area.

21.5 Conclusion

Throughout this chapter, we have summarized exciting progress in developmental cognitive neuroscience, as well as several avenues for future research. Recent methodological advances, like the utilization of naturalistic movie-viewing paradigms and fNIRS, have enabled us to begin to describe the early neural signatures of ToM brain regions. Simultaneously, studies of older children have begun to test different hypotheses about the kinds of neural changes that might support ongoing behavioral ToM development after age 5. Future studies on neural markers that reflect structural changes in ToM, or the impact of developmental or cultural experience, will provide exciting insight into ToM development: What does it mean to become better at ToM in mid- to late childhood, and what is the basis for individual differences in ToM in adults? This work will also lead to a better understanding of the development of functionally specialized brain regions for ToM: What experiences are necessary or sufficient for the maintenance and fine-tuning of preferential responses for ToM? Finally, a key challenge for future studies is to discover reliable neural markers of ToM via longitudinal studies. These markers will provide specific hypotheses about neural differences in clinical populations that struggle disproportionately with ToM reasoning. We are excited to see progress made toward these aims by developmental cognitive neuroscientists in the years to come.

References

- Amsterlaw, J., Wellman, H.M., 2006. Theories of mind in transition: a microgenetic study of the development of false belief understanding. *J. Cogn. Dev.* 7 (2), 139–172. https://doi.org/10.1207/s15327647jcd0702_1.
- Astington, J.W., 2006. The developmental interdependence of theory of mind and language. In: *The Roots of Human Sociality: Culture, Cognition, and Human Interaction*, pp. 179–206.
- Astington, J.W., Jenkins, J.M., 1999. A longitudinal study of the relation between language and theory-of-mind development. *Dev. Psychol.* 35 (5), 1311.
- Avinun, R., Knafo, A., 2014. Parenting as a reaction evoked by children’s genotype: a meta-analysis of children-as-twins studies. *Personal. Soc. Psychol. Rev.* 18 (1), 87–102.
- Avis, J., Harris, P.L., 1991. Belief-desire reasoning among Baka children: evidence for a universal conception of mind. *Child Dev.* 62 (3), 460–467.
- Barrett, H.C., Bolyanatz, A., Crittenden, A.N., Fessler, D.M., Fitzpatrick, S., Gurven, M., et al., 2016. Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proc. Natl. Acad. Sci. U.S.A.* 113 (17), 4688–4693.
- Bedny, M., Pascual-Leone, A., Saxe, R.R., 2009. Growing up blind does not change the neural bases of Theory of Mind. *Proc. Natl. Acad. Sci. U.S.A.* 106 (27), 11312–11317.
- Blank, I., Kanwisher, N., Fedorenko, E., 2014. A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *J. Neurophysiol.* 112 (5), 1105–1118.
- Bloom, P., German, T.P., 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77 (1), B25–B31.
- Brambling, M., Asbrock, D., 2010. Validity of false belief tasks in blind children. *J. Autism Dev. Disord.* 40 (12), 1471–1484. <https://doi.org/10.1007/s10803-010-1002-2>.
- Brown, R., Hobson, R.P., Lee, A., Stevenson, J., 1997. Are there “Autistic-like” features in congenitally blind children? *J. Child Psychol. Psychiatry* 38 (6), 693–703.
- Bruneau, E.G., Pluta, A., Saxe, R., 2012. Distinct roles of the “shared pain” and “theory of mind” networks in processing others’ emotional suffering. *Neuropsychologia* 50 (2), 219–231.
- Brunet, E., Sarfati, Y., Hardy-Baylé, M.-C., Decety, J., 2000. A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* 11 (2), 157–166. <https://doi.org/10.1006/nimg.1999.0525>.
- Brunet, E., Sarfati, Y., Hardy-Baylé, M.-C., Decety, J., 2003. Abnormalities of brain function during a nonverbal theory of mind task in schizophrenia. *Neuropsychologia* 41 (12), 1574–1582.
- Buckner, R.L., Andrews Hanna, J.R., Schacter, D.L., 2008. The brain’s default network. *Ann. N. Y. Acad. Sci.* 1124 (1), 1–38.
- Buttelmann, D., Carpenter, M., Tomasello, M., 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112 (2), 337–342. <https://doi.org/10.1016/j.cognition.2009.05.006>.
- Callaghan, T., Rochat, P., Lillard, A., Claux, M.L., Odden, H., Itakura, S., et al., 2005. Synchrony in the onset of mental-state reasoning: evidence from five cultures. *Psychol. Sci.* 16 (5), 378–384.
- Canlon, J.F., Li, R., 2013. Neural activity during natural viewing of Sesame Street statistically predicts test scores in early childhood. *PLoS Biol.* 11 (1), e1001462.
- Canlon, J.F., Pinel, P., Dehaene, S., Pelphrey, K.A., 2010. Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebr. Cortex* 21 (1), 191–199. <https://doi.org/10.1093/cercor/bhq078>.
- Cao, M., He, Y., Dai, Z., Liao, X., Jeon, T., Ouyang, M., et al., 2016. Early development of functional network segregation revealed by connectomic analysis of the preterm human brain. *Cerebr. Cortex* 27 (3), 1949–1963.
- Carlson, S.M., Moses, L.J., 2001. Individual differences in inhibitory control and children’s theory of mind. *Child Dev.* 72 (4), 1032–1053.
- Carlson, S.M., Moses, L.J., Hix, H.R., 1998. The role of inhibitory processes in young children’s difficulties with deception and false belief. *Child Dev.* 69 (3), 672–691.
- Carrington, S.J., Bailey, A.J., 2009. Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum. Brain Mapp.* 30 (8), 2313–2335. <https://doi.org/10.1002/hbm.20671>.
- Chen, J., Leong, Y.C., Norman, K.A., Hasson, U., 2016. Shared Experience, Shared Memory: A Common Structure For Brain Activity During Naturalistic Recall. *bioRxiv*, p. 035931.
- Cole, M.W., Bassett, D.S., Power, J.D., Braver, T.S., Petersen, S.E., 2014. Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83 (1), 238–251.
- Cooper, R., Selb, J., Gagnon, L., Phillip, D., Schytz, H.W., Iversen, H.K., et al., 2012. A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Front. Neurosci.* 6, 147.
- Cushman, F., Sheketoff, R., Wharton, S., Carey, S., 2013. The development of intent-based moral judgment. *Cognition* 127 (1), 6–21.
- Deen, B., Koldewyn, K., Kanwisher, N., Saxe, R., 2015. Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebr. Cortex* 25 (11), 4596–4609.
- Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes Filho, G., Jobert, A., et al., 2010. How learning to read changes the cortical networks for vision and language. *Science* 330 (6009), 1359–1364.
- Dehaene-Lambertz, G., Monzalvo, K., Dehaene, S., 2018. The emergence of the visual word form: longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biol.* 16 (3), e2004103.
- Doria, V., Beckmann, C.F., Arichi, T., Merchant, N., Groppo, M., Turkheimer, F.E., et al., 2010. Emergence of resting state networks in the preterm human brain. *Proc. Natl. Acad. Sci. U.S.A.* 107 (46), 20015–20020.

- Drayton, L.A., Santos, L.R., 2016. A decade of theory of mind research on Cayo Santiago: insights into rhesus macaque social cognition. *Am. J. Primatol.* 78 (1), 106–116.
- Dubois, J., Adolphs, R., 2016. Building a science of individual differences from fMRI. *Trends Cogn. Sci.* 20 (6), 425–443. <https://doi.org/10.1016/j.tics.2016.03.014>.
- Dubois, J., Adolphs, R., 2016. How the brain represents other minds. *Proc. Natl. Acad. Sci. U.S.A.* 113 (1), 19–21. <https://doi.org/10.1073/pnas.1522316113>.
- Filippova, E., Astington, J.W., 2008. Further development in social reasoning revealed in discourse irony understanding. *Child Dev.* 79 (1), 126–138. <https://doi.org/10.1111/j.1467-8624.2007.01115.x>.
- Foulkes, L., Blakemore, S.J., 2018. *Studying Individual Differences in Human Adolescent Brain Development*, vol. 1. Nature Publishing Group.
- Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102 (27), 9673–9678.
- Franchak, J.M., Heeger, D.J., Hasson, U., Adolph, K.E., 2016. Free viewing gaze behavior in infants and adults. *Infancy* 21 (3), 262–287.
- Frank, M.C., Vul, E., Saxe, R., 2011. Measuring the development of social attention using free-viewing. *Infancy* 17 (4), 355–375. <https://doi.org/10.1111/j.1532-7078.2011.00086.x>.
- Gao, X., Maurer, D., 2010. A happy story: developmental changes in children’s sensitivity to facial expressions of varying intensities. *J. Exp. Child Psychol.* 107 (2), 67–86.
- Gao, W., Zhu, H., Giovanello, K.S., Smith, J.K., Shen, D., Gilmore, J.H., Lin, W., 2009. Evidence on the emergence of the brain’s default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc. Natl. Acad. Sci. U.S.A.* 106 (16), 6790–6795.
- Gao, W., Gilmore, J.H., Shen, D., Smith, J.K., Zhu, H., Lin, W., 2013. The synchronization within and interaction between the default and dorsal attention networks in early infancy. *Cerebr. Cortex* 23 (3), 594–603.
- Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M.K., Caldara, R., Turati, C., 2016. Culture shapes 7-month-olds’ perceptual strategies in discriminating facial expressions of emotion. *Curr. Biol.* 26 (14), R663–R664.
- Golarai, G., Ghahremani, D.G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J.L., Gabrieli, J.D., Grill-Spector, K., 2007. Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* 10 (4), 512.
- Gomez, J., Barnett, M.A., Natu, V., Mezer, A., Palomero-Gallagher, N., Weiner, K.S., et al., 2017. Microstructural proliferation in human cortex is coupled with the development of face processing. *Science* 355 (6320), 68–71.
- Gopnik, A., Wellman, H.M., 1992. Why the child’s theory of mind really is a theory. *Mind Lang.* 7 (1-2), 145–171.
- Gopnik, A., Wellman, H.M., 1994. The theory theory. In: *Mapping the Mind: Domain Specificity in Cognition and Culture*, p. 257.
- Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100 (1), 253–258.
- Grill-Spector, K., Golarai, G., Gabrieli, J., 2008. Developmental neuroimaging of the human ventral visual cortex. *Trends Cogn. Sci.* 12 (4), 152–162.
- Grosse Wiesmann, C., Schreiber, J., Singer, T., Steinbeis, N., Friederici, A.D., 2017. White matter maturation is associated with the emergence of Theory of Mind in early childhood. *Nat. Commun.* 8, 14692.
- Guggenheim, R. (Producer), Arnold, B. (Producer), Lasseter, J. (Director), 1995. *Toy Story* [Motion Picture] USA: Pixar Animation Studios and Walt Disney Pictures.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98 (7), 4259–4264.
- Gweon, H., Dodell-Feder, D., Bedny, M., Saxe, R., 2012. Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Dev.* 83 (6), 1853–1868. <https://doi.org/10.1111/j.1467-8624.2012.01829.x>.
- Happé, F., Conway, J.R., 2016. Recent progress in understanding skills and impairments in social cognition. *Curr. Opin. Pediatr.* 28 (6), 736–742.
- Happé, F., Frith, U., 2014. Annual research review: towards a developmental neuroscience of atypical social cognition. *J. Child Psychol. Psychiatry* 55 (6), 553–577.
- Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., et al., 1996. “Theory of mind” in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport* 8 (1), 197–201.
- Hasson, U., 2004. Intersubject synchronization of cortical activity during natural vision. *Science* 303 (5664), 1634–1640. <https://doi.org/10.1126/science.1089506>.
- Henrich, J., 2011. Thematic Series: The Emerging Science of Culture the Weirdest People in the World.
- Honey, C.J., Thompson, C.R., Lerner, Y., Hasson, U., 2012. Not lost in translation: neural responses shared across languages. *J. Neurosci.* 32 (44), 15277–15283.
- Hughes, C., Cutting, A.L., 1999. Nature, nurture, and individual differences in early understanding of mind. *Psychol. Sci.* 10 (5), 429–432.
- Hughes, C., Jaffee, S.R., Happé, F., Taylor, A., Caspi, A., Moffitt, T.E., 2005. Origins of individual differences in theory of mind: from nature to nurture? *Child Dev.* 76 (2), 356–370.
- Hyde, D.C., Aparicio Betancourt, M., Simon, C.E., 2015. Human temporal-parietal junction spontaneously tracks others’ beliefs: a functional near-infrared spectroscopy study. *Hum. Brain Mapp.* 36 (12), 4831–4846.
- Hyde, D.C., Simon, C.E., Ting, F., Nikolaeva, J., 2018. Functional organization of the temporal-parietal junction for theory of mind in preverbal infants: a near-infrared spectroscopy study. *J. Neurosci.* 38 (18), 4264–4274.
- Jääskeläinen, I.P., Koskentalo, K., Balk, M.H., Autti, T., Kauramäki, J., Pomren, C., Sams, M., 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimaging J.* 2, 14.

- Kilford, E.J., Garrett, E., Blakemore, S.-J., 2016. The development of social cognition in adolescence: an integrated perspective. *Neurosci. Biobehav. Rev.* 70, 106–120.
- Knafo, A., Israel, S., 2009. Genetic and environmental influences on prosocial behavior. In: *Prosocial Motives, Emotions, and Behavior: the Better Angels of Our Nature*, pp. 149–167.
- Knafo, A., Plomin, R., 2006. Prosocial behavior from early to middle childhood: genetic and environmental influences on stability and change. *Dev. Psychol.* 42 (5), 771.
- Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J.L., Rhee, S.H., 2008. The developmental origins of a disposition toward empathy: genetic and environmental contributions. *Emotion* 8 (6), 737.
- Knafo-Noam, A., Vertsberger, D., Israel, S., 2018. Genetic and environmental contributions to children's prosocial behavior: brief review and new evidence from a reanalysis of experimental twin data. *Curr. Opin. Psychol.* 20, 60–65.
- Knudsen, B., Liskowski, U., 2012. Eighteen-and 24-month-old infants correct others in anticipation of action mistakes. *Dev. Sci.* 15 (1), 113–122.
- Knudsen, B., Liskowski, U., 2012. 18-Month-Olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy* 17 (6), 672–691.
- Kobayashi, C., Glover, G.H., Temple, E., 2007. Children's and adults' neural bases of verbal and nonverbal 'theory of mind'. *Neuropsychologia* 45 (7), 1522–1532.
- Koster-Hale, J., Saxe, R., Dungan, J., Young, L.L., 2013. Decoding moral judgments from neural representations of intentions. *Proc. Natl. Acad. Sci. U.S.A.* 110 (14), 5648–5653. <https://doi.org/10.1073/pnas.1207992110>.
- Koster-Hale, J., Bedny, M., Saxe, R., 2014. Thinking about seeing: perceptual sources of knowledge are encoded in the theory of mind brain regions of sighted and blind adults. *Cognition* 133 (1), 65–78. <https://doi.org/10.1016/j.cognition.2014.04.006>.
- Koster-Hale, J., Richardson, H., Velez, N., Asaba, M., Young, L., Saxe, R., 2017. Mentalizing regions represent distributed, continuous, and abstract dimensions of others' beliefs. *Neuroimage* 161, 9–18. <https://doi.org/10.1016/j.neuroimage.2017.08.026>.
- Lagattuta, K.H., Wellman, H.M., 2001. Thinking about the past: early knowledge about links between prior experience, thinking, and emotion. *Child Dev.* 72 (1), 82–102.
- Laird, A.R., Eickhoff, S.B., Rottschy, C., Bzdok, D., Ray, K.L., Fox, P.T., 2013. Networks of task co-activations. *Neuroimage* 80, 505–514.
- Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U., 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31 (8), 2906–2915. <https://doi.org/10.1523/JNEUROSCI.3684-10.2011>.
- Lloyd-Fox, S., Blasi, A., Elwell, C.E., 2010. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34 (3), 269–284.
- Lloyd-Fox, S., Richards, J.E., Blasi, A., Murphy, D.G., Elwell, C.E., Johnson, M.H., 2014. Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* 1 (2), 025006.
- Lloyd-Fox, S., Begus, K., Halliday, D., Pirazzoli, L., Blasi, A., Papademetriou, M., et al., 2017. Cortical specialisation to social stimuli from the first days to the second year of life: a rural Gambian cohort. *Accid. Anal. Prev.* 25, 92–104.
- Mar, R.A., 2011. The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* 62, 103–134.
- Martin, A., Santos, L.R., 2016. What cognitive representations support primate theory of mind? *Trends Cogn. Sci.* 20 (5), 375–382.
- McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7 (7), 293–299. [https://doi.org/10.1016/S1364-6613\(03\)00134-7](https://doi.org/10.1016/S1364-6613(03)00134-7).
- McGue, M., Bouchard Jr., T.J., Iacono, W.G., Lykken, D.T., 1993. Behavioral Genetics of Cognitive Ability: A Life-span Perspective.
- Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., Siegal, M., 2012. Belief attribution in deaf and hearing infants. *Dev. Sci.* 15 (5), 633–640.
- Miall, R.C., Robertson, E.M., 2006. Functional imaging: is the resting brain resting? *Curr. Biol.* 16 (23), R998–R1000.
- Milligan, K., Astington, J.W., Dack, L.A., 2007. Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78 (2), 622–646.
- Minter, M., Hobson, R.P., Bishop, M., 1998. Congenital visual impairment and "theory of mind." *Br. J. Dev. Psychol.* 16, 183.
- Mitchell, J.P., Banaji, M.R., Macrae, C.N., 2005. General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage* 28 (4), 757–762.
- Moraczewski, D., Chen, G., Redcay, E., 2018. Inter-subject synchrony as an index of functional specialization in early childhood. *Sci. Rep.* 8 (1), 2252.
- Moreira, J.F.G., McLaughlin, K.A., Silvers, J.A., 2018. Spatial and Temporal Cortical Variability Track With Age and Affective Experience During Emotion Regulation in Youth. *bioRxiv*, p. 291245.
- Nguyen, M., Vanderwal, T., Hasson, U., 2017. Shared Understanding is Correlated With Shared Neural Responses in The Default Mode Network. *bioRxiv*, p. 231019.
- Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I.P., Hari, R., Sams, M., 2012. Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci. U.S.A.* 109 (24), 9599–9604.
- Paunov, A., Blank, I., Fedorenko, E., 2017. Functionally Distinct Language and Theory of Mind Networks Work Together, Especially During Language Comprehension, 27. *PsyArXiv*.
- Perner, J., Leekam, S.R., Wimmer, H., 1987. Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5 (2), 125–137.
- Peterson, C.C., Siegal, M., 1999. Representing inner worlds: theory of mind in autistic, deaf, and normal hearing children. *Psychol. Sci.* 10 (2), 126–129.

- Peterson, C.C., Wellman, H.M., 2018. Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM Delay. *Child Dev.* 90 (6), 1917–1934.
- Peterson, C.C., Peterson, J.L., Webb, J., 2000. Factors influencing the development of a theory of mind in blind children. *Br. J. Dev. Psychol.* 18 (3), 431–447.
- Peterson, C.C., Wellman, H.M., Slaughter, V., 2012. The mind behind the message: advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Dev.* 83 (2) <https://doi.org/10.1111/j.1467-8624.2011.01728.x>.
- Plomin, R., Kosslyn, S.M., 2001. Genes, brain and cognition. *Nat. Neurosci.* 4 (12), 1153–1154.
- Polk, T.A., Park, J., Smith, M.R., Park, D.C., 2007. Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. *J. Neurosci.* 27 (51), 13921–13925.
- Powell, L.J., Deen, B., Saxe, R., 2017. Using individual functional channels of interest to study cortical development with fNIRS. *Dev. Sci.* 21 (4), e12595.
- Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1 (04), 515–526.
- Priewasser, B., Rafetseder, E., Gargitter, C., Perner, J., 2017. Helping as an early indicator of a theory of mind: mentalism or teleology? *Cogn. Dev.* 46, 69–78.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98 (2), 676–682.
- Raschle, N.M., Lee, M., Buechler, R., Christodoulou, J.A., Chang, M., Vakil, M., et al., 2009. Making MR imaging child’s play-pediatric neuroimaging protocol, guidelines and procedure. *J. Vis. Exp.* (29).
- Reher, K. (Producer), Sohn, P. (Director), 2009. *Partly Cloudy* [Motion Picture] USA: Pixar Animation Studios and Walt Disney Pictures.
- Richardson, H., 2019. Development of brain networks for social functions: confirmatory analyses in a large open source dataset. *Dev. Cognit. Neurosci.* 37, 100598.
- Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., Saxe, R., 2018. Development of the social brain from age three to twelve years. *Nat. Commun.* 9 (1), 1027.
- Richardson, H., Koster-Hale, J., Caselli, N., Magid, R.W., Benedict, R., Olson, H., Saxe, R., 2019. Reduced Neural Selectivity For Mental States in Deaf Children With Delayed Exposure to Sign Language. *PsyArxiv*. <https://doi.org/10.31234/osf.io/j5ghw> (In revision).
- Ronald, A., Happé, F., Hughes, C., Plomin, R., 2005. Nice and nasty theory of mind in preschool children: nature and nurture. *Soc. Dev.* 14 (4), 664–684.
- Ronald, A., Viding, E., Happé, F., Plomin, R., 2006. Individual differences in theory of mind ability in middle childhood and links with verbal ability and autistic traits: a twin study. *Soc. Neurosci.* 1 (3–4), 412–425.
- Rosati, A.G., Wobber, V., Hughes, K., Santos, L.R., 2014. Comparative developmental psychology: how is human cognitive development unique? *Evol. Psychol.* 12 (2), 147470491401200211.
- Saxe, R., 2018. Seeing other minds in 3D. *Trends Cogn. Sci.* 22 (3), 193–195.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind.” *Neuroimage* 19 (4), 1835–1842.
- Saxe, R., Powell, L.J., 2006. It’s the thought that counts specific brain regions for one component of theory of mind. *Psychol. Sci.* 17 (8), 692–699.
- Saxe, R.R., Whitfield-Gabrieli, S., Scholz, J., Pelphrey, K.A., 2009. Brain regions for perceiving and reasoning about other people in school-aged children. *Child Dev.* 80 (4), 1197–1209.
- Schick, B., Hoffmeister, R., 2001. ASL skills in deaf children of deaf parents and of hearing parents. In: Presented at the Society for Research in Child Development International Conference, Minneapolis, MN.
- Schick, B., De Villiers, P., De Villiers, J., Hoffmeister, R., 2007. Language and theory of mind: a study of deaf children. *Child Dev.* 78 (2), 376–396.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J., 2014. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* 42, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>.
- Selcuk, B., Brink, K.A., Ekerim, M., Wellman, H.M., 2018. Sequence of theory-of-mind acquisition in Turkish children from diverse social backgrounds. *Infant Child Dev.* e2098.
- Shahaeian, A., Peterson, C.C., Slaughter, V., Wellman, H.M., 2011. Culture and the sequence of steps in theory of mind development. *Dev. Psychol.* 47 (5), 1239.
- Skerry, A.E., Spelke, E.S., 2014. Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition* 130 (2), 204–216.
- Sliwa, J., Freiwald, W.A., 2017. A dedicated network for social interaction processing in the primate brain. *Science* 356 (6339), 745–749.
- Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., et al., 2009. Correspondence of the brain’s functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106 (31), 13040–13045.
- Smyser, C.D., Inder, T.E., Shimony, J.S., Hill, J.E., Degnan, A.J., Snyder, A.Z., Neil, J.J., 2010. Longitudinal analysis of neural network development in preterm infants. *Cerebr. Cortex* 20 (12), 2852–2862.
- Sommer, M., Döhnell, K., Sodian, B., Meinhardt, J., Thoermer, C., Hajak, G., 2007. Neural correlates of true and false belief reasoning. *Neuroimage* 35 (3), 1378–1384.
- Southgate, V., Senju, A., Csibra, G., 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18 (7), 587–592. <https://doi.org/10.1111/j.1467-9280.2007.01944.x>.
- Stephens, G.J., Silbert, L.J., Hasson, U., 2010. Speaker–listener neural coupling underlies successful communication. *Proc. Natl. Acad. Sci. U.S.A.* 107 (32), 14425–14430.
- Surian, L., Caldi, S., Sperber, D., 2007. Attribution of beliefs by 13-month-old infants. *Psychol. Sci.* 18 (7), 580–586.

- Tamir, D.I., Thornton, M.A., Contreras, J.M., Mitchell, J.P., 2016. Neural evidence that three dimensions organize mental state representation: rationality, social impact, and valence. *Proc. Natl. Acad. Sci. U.S.A.* 113 (1), 194–199. <https://doi.org/10.1073/pnas.1511905112>.
- Turkheimer, E., Haley, A., Waldron, M., d'Onofrio, B., Gottesman, I.I., 2003. Socioeconomic status modifies heritability of IQ in young children. *Psychol. Sci.* 14 (6), 623–628.
- van den Heuvel, M.I., Thomason, M.E., 2016. Functional connectivity of the human brain in utero. *Trends Cogn. Sci.* 20 (12), 931–939.
- Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L.C., Castellanos, F.X., 2015. Inscapes: a movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage* 122, 222–232.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., et al., 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* 14 (1), 170–181. <https://doi.org/10.1006/nimg.2001.0789>.
- Wade, M., Hoffmann, T.J., Jenkins, J.M., 2015. Gene–environment interaction between the oxytocin receptor (OXTR) gene and parenting behaviour on children's theory of mind. *Soc. Cogn. Affect. Neurosci.* 10 (12), 1749–1757.
- Warrier, V., Baron-Cohen, S., 2018. Genetic contribution to “theory of mind” in adolescence. *Sci. Rep.* 8 (1), 3465.
- Weisman, K., Dweck, C.S., Markman, E.M., 2017. Children's intuitions about the structure of mental life. In: Presented at the the Annual Meeting of the Cognitive Science Society.
- Wellman, H.M., Liu, D., 2004. Scaling of theory-of-mind tasks. *Child Dev.* 75 (2), 523–541.
- Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72 (3), 655–684.
- Wellman, H.M., Fang, F., Liu, D., Zhu, L., Liu, G., 2006. Scaling of theory-of-mind understandings in Chinese children. *Psychol. Sci.* 17 (12), 1075–1081.
- Wellman, H.M., Lopez-Duran, S., LaBounty, J., Hamilton, B., 2008. Infant attention to intentional action predicts preschool theory of mind. *Dev. Psychol.* 44 (2), 618–623. <https://doi.org/10.1037/0012-1649.44.2.618>.
- Wellman, H.M., Fang, F., Peterson, C.C., 2011. Sequential progressions in a theory-of-mind scale: longitudinal perspectives. *Child Dev.* 82 (3), 780–792.
- Widen, S.C., Russell, J.A., 2003. A closer look at preschoolers' freely produced labels for facial expressions. *Dev. Psychol.* 39 (1), 114.
- Widen, S.C., Russell, J.A., 2008. Children acquire emotion categories gradually. *Cogn. Dev.* 23 (2), 291–312.
- Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M., 2007. Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cerebr. Cortex* 18 (1), 230–242.
- 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).
- Woolfe, T., Want, S.C., Siegal, M., 2002. Signposts to development: theory of mind in deaf children. *Child Dev.* 73 (3), 768–778.
- Wu, N., Su, Y., 2015. Oxytocin receptor gene relates to theory of mind and prosocial behavior in children. *J. Cogn. Dev.* 16 (2), 302–313.
- Wu, Y., Muentener, P., Schulz, L.E., 2017. One-to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proc. Natl. Acad. Sci.* 114 (45), 11896–11901.
- Xiao, Y., Geng, F., Riggins, T., Chen, G., Redcay, E., 2019. Neural correlates of developing theory of mind competence in early childhood. *Neuroimage* 184, 707–716.
- Yamaguchi, M., Kuhlmeier, V.A., Wynn, K., vanMarle, K., 2009. Continuity in social cognition from infancy to childhood. *Dev. Sci.* 12 (5), 746–752. <https://doi.org/10.1111/j.1467-7687.2008.00813.x>.
- Young, L., Saxe, R., 2009. An fMRI investigation of spontaneous mental state inference for moral judgment. *J. Cogn. Neurosci.* 21 (7), 1396–1405.
- Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R., 2010. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci. U.S.A.* 107 (15), 6753–6758. <https://doi.org/10.1073/pnas.0914826107>.
- Young, L., Dodell-Feder, D., Saxe, R., 2010. What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia* 48 (9), 2658–2664.
- Young, L., Nichols, S., Saxe, R., 2010. Investigating the neural and cognitive basis of moral luck: it's not what you do but what you know. *Rev. Philos. Psychol.* 1 (3), 333–349. <https://doi.org/10.1007/s13164-010-0027-y>.