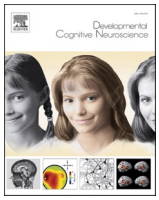
Developmental Cognitive Neuroscience 49 (2021) 100959

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn



Neural correlates of theory of mind in children and adults using CAToon: Introducing an open-source child-friendly neuroimaging task

Réka Borbás [a,b], Lynn V. Fehlbaum [a,b], Ursula Rudin [b], Christina Stadler [b], Nora M. Raschle [a,b,*]

[a] Jacobs Center for Productive Youth Development at the University of Zurich, Zurich, Switzerland
[b] Department of Child and Adolescent Psychiatry, University of Basel, University Psychiatric Clinics Basel, Basel, Switzerland

ARTICLE INFO

Keywords:
Theory of mind
Mentalizing
Development
Neuroimaging
Childhood
fMRI

ABSTRACT

Theory of Mind (ToM) or mentalizing is a basic social skill which is characterized by our ability of perspective-taking and the understanding of cognitive and emotional states of others. ToM development is essential to successfully navigate in various social contexts. The neural basis of mentalizing is well-studied in adults, however, less evidence exists in children. Potential reasons are methodological challenges, including a lack of age-appropriate fMRI paradigms. We introduce a novel child-friendly and open-source ToM fMRI task, for which accuracy and performance were evaluated behaviorally in 60 children ages three to nine (32♂). Furthermore, 27 healthy young adults (14♂; mean = 25.41 years) and 33 children ages seven to thirteen (17♂; mean = 9.06 years) completed the **C**ognitive and **A**ffective **T**heory of Mind Cart**oon** task (**CAToon;** www.jacobscenter.uzh.ch/en/research/developmental_neuroscience/downloads/catoon.html) during a fMRI session. Behavioral results indicate that children of all ages can solve the CAToon task above chance level, though reliable performance is reached around five years. Neurally, activation increases were observed for adults and children in brain regions previously associated with mentalizing, including bilateral temporoparietal junction, temporal gyri, precuneus and medial prefrontal/orbitofrontal cortices. We conclude that CAToon is suitable for developmental neuroimaging studies within an fMRI environment starting around preschool and up.

1. Introduction

Theory of Mind (ToM), or mentalizing, describes our ability to represent and understand the mental states (feelings, beliefs, desires and intentions) of others and ourselves (Gallagher and Frith, 2003). ToM is an essential social skill (Frith and Frith, 2008; Korkmaz, 2011; Saxe, 2006; Schnell et al., 2011) and a failure to develop adequate ToM skills is associated with different neurodevelopmental disorders, including autism spectrum disorder, developmental language disorder, attention deficit hyperactivity disorder and conduct disorder (Clegg et al., 2005; Frith, 2001; Sebastian et al., 2012b; Senju, 2012; Sharp, 2008; Uekermann et al., 2010). ToM may be differentiated into an affective (the understanding of emotions of others) and cognitive subcomponent (inferences a person makes about other people's beliefs and intentions; (Shamay-Tsoory et al., 2010)). Behaviorally, ToM has commonly been associated with empathy (O'Connell, 1995; Saxe, 2006; Völlm et al., 2006; Walter, 2012), and functional neuroimaging evidence indicates a partial overlap of brain regions commonly associated with empathy and mentalizing (Bzdok et al., 2012; Powell et al., 2017; Völlm et al., 2006).

The foundations of mentalizing are laid during the first few years of life, though they become more refined throughout childhood and adolescence. Early conceptualizations of ToM tasks have particularly focused on explicit measures (e.g., Sally and Anne Task (Baron-Cohen et al., 1985)), which are mastered around the age of 4 (Baron-Cohen et al., 1985; Wimmer and Perner, 1983). However, studies employing implicit ToM assessments during infancy, as for example through the investigation of an infant's anticipatory looks, have suggested that ToM may develop as early as 13–15 months (Onishi and Baillargeon, 2005; Southgate et al., 2007; Surian et al., 2007). Consequently, the type of task employed to measure ToM, or mentalizing, may have a significant influence on the interpretation of the reported skill levels.

Studies assessing the neural correlates of ToM in adults (Bzdok et al., 2012; Molenberghs et al., 2016; Van Overwalle and Baetens, 2009) have consistently linked mentalizing to brain areas within the frontal (e.g., anterior dorsal medial and ventromedial PFC, inferior frontal- and precentral gyri and the anterior cingulate cortex), temporal and parietal cortices (e.g., bilateral temporoparietal junction, middle temporal gyri, posterior superior temporal sulci and the precuneus (Molenberghs et al.,

* Corresponding author at: Jacobs Center for Productive Youth Development, University of Zurich, Andreasstrasse 15, Zürich, CH 8050, Switzerland.
E-mail address: nora.raschle@jacobscenter.uzh.ch (N.M. Raschle).

https://doi.org/10.1016/j.dcn.2021.100959
Received 3 April 2020; Received in revised form 27 April 2021; Accepted 1 May 2021
Available online 4 May 2021

2016)). In accordance with the conceptual separation of affective and cognitive ToM, distinct networks can be identified (Schlaffke et al., 2015; Sebastian et al., 2012a; Shamay-Tsoory and Aharon-Peretz, 2007). While affective and cognitive mentalizing are controlled by a shared network comprising bilateral temporal poles, superior temporal sulci and the dorsomedial prefrontal cortex, the specific role of orbitofrontal and ventromedial prefrontal cortices in affective mentalizing has been highlighted based on research in clinical and healthy populations (Hynes et al., 2006; Sebastian et al., 2012a; Shamay-Tsoory and Aharon-Peretz, 2007). Affective ToM has also particularly been associated with basal ganglia functioning (Schlaffke et al., 2015; Bodden et al., 2013). On the other hand, cognitive mentalizing processes are more specifically linked to activation in the dorsomedial prefrontal cortex, precuneus, cuneus, bilateral temporoparietal junction, and the middle of the superior temporal gyri (Molenberghs et al., 2016; Van Overwalle and Baetens, 2009; Schlaffke et al., 2015).

While various reports describe the neural correlates of ToM in adults, less is known for younger children, with or without neurodevelopmental disorders. Potential reasons may include practical and technical challenges as well as a lack of age-adequate scanner tasks (Raschle et al., 2012, 2009; Thieba et al., 2018). However, in recent years new studies have emerged investigating mentalizing in young populations through task-based functional magnetic resonance imaging (fMRI) or functional near infrared spectroscopy (Gweon et al., 2012; Hyde et al., 2018; Moraczewski et al., 2018; Richardson et al., 2018; Richardson and Saxe, 2020a). Such studies implement auditory paradigms, false belief tasks or incorporate more naturalistic settings such as passive movie viewing tasks. Alternatively, task-free functional (e.g., resting state fMRI (Xiao et al., 2019)) or structural measures, (e.g., white matter measures (Wiesmann et al., 2017)) can be further substantiated through the use of additional behavioral ToM measures.

Here we present three experimental studies conducted to assess the development and implementation of the ***C**ognitive and **A**ffective **T**heory of Mind Cart**oon** task* (in short **CAToon;** available at: www.jacobscenter.uzh.ch/en/research/developmental_neuroscience/downloads/catoon.html), a novel, open-source, engaging and child-friendly fMRI mentalizing task. **Study 1** introduces development and behavioral assessment of CAToon in children. Specifically, we aimed to assess the age at which children were able to complete CAToon behaviorally. We hypothesized that CAToon may be completed starting around preschool/kindergarten (around 4 years of age).

Study 2 aimed to investigate whether CAToon will reliably elicit activation in brain regions previously associated with mentalizing in adults. Activation increases were expected in brain regions including dorsomedial PFC, bilateral temporoparietal junction, middle medial PFC, precuneus, right superior temporal sulcus, and ventromedial PFC (Molenberghs et al., 2016). Furthermore, affective and cognitive stories were hypothesized to elicit distinct (cognitive ToM: dorsomedial PFC, precuneus, superior temporal gyrus; affective ToM: OFC, ventromedial PFC, bilateral pars opercularis, basal ganglia,) as well as shared neural activation patterns (e.g., dorsomedial PFC, bilateral temporoparietal junction; (Bodden et al., 2013; Dufour et al., 2013; Hynes et al., 2006; Völlm et al., 2006)).

Study 3 aimed to assess the neural correlates during CAToon performance in a first group of typically developing children. We hypothesized, that CAToon may elicit similar, though still developing neuronal activation patterns in children (Richardson et al., 2018; Richardson and Saxe, 2020b).

2. Materials and methods

2.1. CAToon task

Task creation of the ***C**ognitive and **A**ffective **T**heory of Mind Cart**oon** task* **(CAToon)** included the following steps: (**1**) standardized literature review on ToM fMRI studies (as described in (Fehlbaum et al., 2021), (**2**) evaluation for child-appropriateness according to the following requirements: the task should be feasible for young children, including non-readers (no text), has to be engaging and fun, and should ideally be visually entertaining since this has previously been reported to reduce head motion (Huijbers et al., 2017). As a result, we decided on the use of cartoon stories, which are commonly used in the literature, successfully evoke distinct neuronal activation associated with ToM (e.g., (Schlaffke et al., 2015; Sebastian et al., 2012a; Völlm et al., 2006; Walter et al., 2009)) and which adhere to the aforementioned requirements. In **Study 1** (behavioral) and **3** (fMRI) these requirements were re-evaluated.

CAToon consists of a total of 30 hand-drawn stories, including two experimental conditions targeting affective ToM (AT) and cognitive ToM (CT) and a control condition (physical causality (PC); Fig. 1). Each condition comprises 10 stories of similar visual complexity. Ten backgrounds were prepared for the task and each background occurs only once in each condition. The two experimental conditions were designed to differentially motivate affective versus cognitive aspects of ToM reasoning. That means participants should have to infer how a character would react to a fellow character's expressed or expected emotions during AT trials, whereas during CT trials participants should assume how characters would act based on another character's intentions or beliefs. PC trials serve as a control condition, requiring a basic understanding of cause and effect and basic physical laws.

All trials start with three consecutively presented images, followed by a single image displaying three possible endings. CT trial endings consist of one possible, one improbable and one highly improbable/impossible solution. AT trial endings consist of two possible solutions (negative expectancy/positive expectancy) and one impossible solution. In positive expectancy endings a character's emotional needs are met with care or reassurance, whereas in negative expectancy outcomes the character is scolded, ridiculed or ignored. This manipulation allows the investigation of differences in positive or negative outcome expectancy. PC trial endings consist of one possible and two impossible solutions. As physical causality and cognitive ToM conditions have only one correct answer the chance of getting a correct answer is 33 % (1/3). For affective ToM (AT) we present two correct (negative and positive expectancy) answers resulting in a higher chance of 66 % (2/3). Therefore, for the overall task the chance level is 44 %. In other words, in the PC and CT conditions participants have to get at least 3.3 tasks correctly to reach the level of chance (that adds up to 6.6) and in AT condition they have to get 6.6 answers correctly to reach level of chance. Across the 30 trials that adds up to 13.2, which is 44 % of 30 tasks.

2.1.1. CAToon task evaluation (Study 1: behavioral)

CAToon was evaluated behaviorally in a group setting (first and second grade school classes) or in an individualized manner (preschoolers/kindergarteners). Each participant was asked to look at three images presented in a row and then indicate their choice of the most likely story ending out of three options by either pointing to it (preschoolers/kindergarteners) or by crossing off their choice in a booklet (school-aged children/group setting; details in **Supplement 1**).

2.1.2. CAToon task evaluation (Studies 2 & 3: fMRI)

A total of 30 cartoon stories were rear-projected onto a screen behind the scanner, viewed by participants via a prism attached to the head coil and displayed using Presentation® software (V16.5, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). Adults completed the fMRI task in one single run (trial order in **Supplement 2**), while children completed the task over the course of two runs (15 trials each) as suggested for fMRI in developmental population (Raschle et al., 2009). Both runs included 5 AT, 5 CT, and 5 PC stories. The run order for children completing the fMRI experiment was alternated (starting with either run1 or run2).

The task had a rapid event-related design with fixed inter-trial intervals of 3 s. Before the start of each run, a fixation cross was present for 2 s. Each trial started with the consecutive presentation of three images

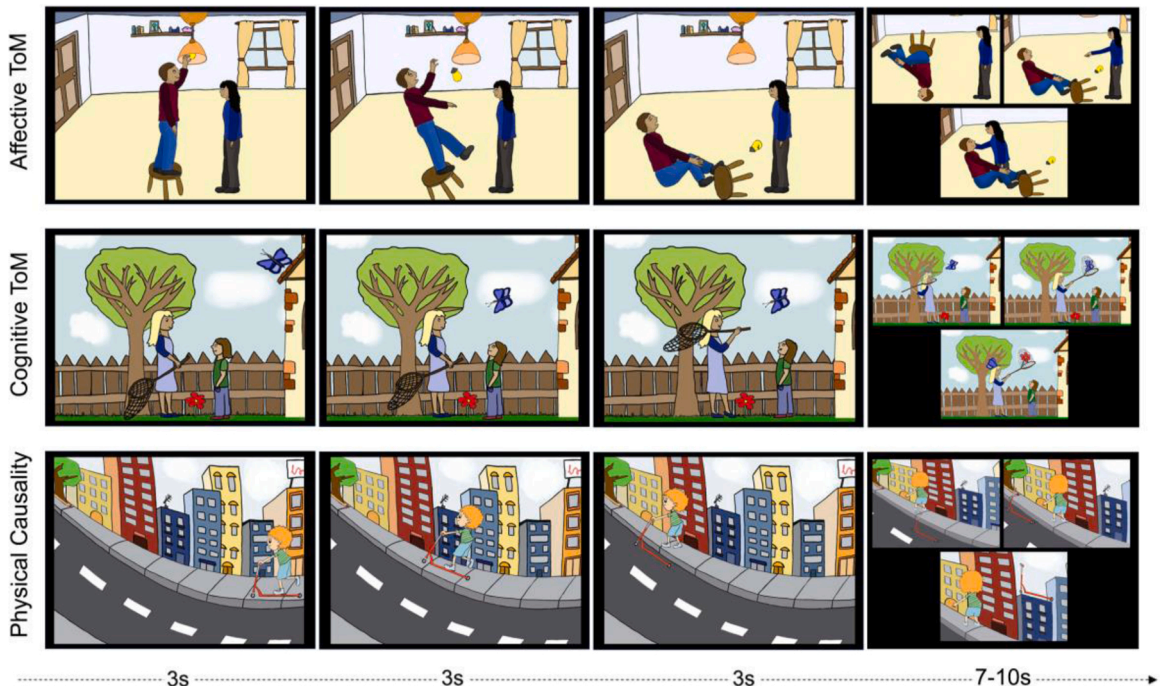


Fig. 1. Three CAToon example trials demonstrating one example of every condition included: affective ToM (experimental condition; top row), cognitive ToM (experimental condition; middle row), and physical causality (control condition; bottom row). The timeline shows the presentation duration for each image presented during fMRI with an answer time window of 7 s and 10 s for adult and child participants, respectively.

(3 s each), followed by a decision phase of 7 s for adults. Based on adult feedback from **Study 2** (i.e., challenged by the relatively short answer time), the decision time was extended to 10 s for children to assure age appropriateness. Participants had to choose one out of three possible endings through use of a button box (task design in Fig. 1). Task duration was 8 min 36 s for adults, and 11 min 4 s for children (two runs of 5 min 32 s). Before solving the CAToon neuroimaging task inside the MRI scanner, participants completed three practice trials on a laptop and by use of a cardboard model button box. After these practice trials it was verbally assured that participants understood the task and key points were repeated prior to the start of neuroimaging (further info in **Supplement 2**).

2.2. Participants and analyses

2.2.1. Participants Study 1: CAToon task evaluation in children (behavioral)

For **Study 1** 60 children ages three to nine years (mean age: 5.77 years; 32 boys; group characteristics in **Supplementary Table S1**) completed the CAToon task behaviorally. All children were recruited through daycare, kindergarten or schools. Answer choices, age and gender, but no identifiable personal data, was collected. In line with approval by the local ethics board (Ethikkomission Nordwest- und Zentralschweiz), informed assent to participate was provided by the daycare teachers, school principals or parents. Families were informed about the participation and had the option to withhold contribution.

2.2.2. Analyses Study 1

The mean percentage of correct answers for overall task performance was calculated for each participant. A one-way analysis of variance (ANOVA) was employed to inspect overall performance difference across age groups (by year). The group of nine-year old children was excluded from this analysis due to small sample size (n = 2). Prior to conducting the ANOVA, assumption of normality (kurtosis and skewness) and homogeneity of variance (Levene's test) was tested. The results from the ANOVA were followed up with Bonferroni-corrected post-hoc group comparisons. In addition to the categorical investigation of change in performance based on age in years, changes based on age in months were assessed (dimensionally) using partial F-tests to select the best-fitting regression model. Projected changes in performance based on age were calculated using the CurveExpert Professional Software (https://www.curveexpert.net/) by displaying the instantaneous rate of change (the slope of the tangent line at a given point on the curve). All behavioral analyses were performed in SPSS (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.) or R version 4.0.3 (R Core Team, 2020; https://www.R-project.org/).

2.2.3. Participants Studies 2 & 3: CAToon task evaluation in adults and children (fMRI)

28 healthy young adults and 37 typically developing children took part in the fMRI experiments assessing the neural correlates of mentalizing using CAToon. Participants included in the fMRI studies had normal or corrected-to-normal vision, sufficient German skills, no previous neuropsychological disorder, and average to above average IQ based on their level of education (for adults) or an IQ $\geq$ 70 (for children; verbal and non-verbal subtests of the German version of the Wechsler Intelligence Scale for Children (WISC-IV; (Daseking et al., 2007)). One adult was excluded from the study due to strong visual prescription glasses that could not be used or adjusted for within the scanner. Four children were excluded due artefacts caused by braces (n = 2), and claustrophobia (n = 2). The final sample therefore included 27 adults (20–39 years), and 33 children (7–13 years; group characteristics are listed in Table 1).

Adult participants further completed standardized questionnaires assessing callous-unemotional traits (callous-unemotional dimension of the Youth Psychopathic traits Inventory (YPI; (Andershed et al., 2002)) and empathy (Interpersonal Reactivity Index (IRI; (Davis, 1980)). This allowed the investigation of the association between behavioral scores of callous-unemotional traits or empathy and neural activation during mentalizing using post-hoc assessments. All participants and in case of the children their legal caretakers provided verbal and written consent for taking part in the study.

Table 1
Group characteristics and behavioral scores of adult and child participants of **Studies 2** and **3** (fMRI).

		Adults (n = 27;14♂)		Children (n = 33;18♂)	
		Mean	SD	Mean	SD
Age		25.41	4.16	9.06	2.11
ISCED		4.22	0.97	n.a.	n.a.
IQ	WISC-IV (verbal)	n.a.	n.a.	113.18	17.18
	WISC-IV (matrices)	n.a.	n.a.	110.45	13.83
IRI	Empathic Concern	18.56	4.04	n.a.	n.a.
	Fantasy	16.37	5.06	n.a.	n.a.
	Personal Distress	10.96	4.03	n.a.	n.a.
	Perspective Taking	18.93	3.10	n.a.	n.a.
	Total	64.81	8.91	n.a.	n.a.
YPI	Callous-Unemotional	9.22	1.76	n.a.	n.a.
	Grandiose-Manipulative	8.67	1.95	n.a.	n.a.
	Impulsive-Irresponsible	9.99	2.14	n.a.	n.a.
	Total	9.29	1.48	n.a.	n.a.

Notes. ISCED = International Standard Classification of Education, IRI = Interpersonal Reactivity Index (sum scores), YPI = Youth Psychopathic Traits Inventory (mean scores), WISC-IV = Wechsler Intelligence Scale for Children, Fourth Edition, SD = standard deviation, n.a. = not applicable.

2.2.4. fMRI data acquisition (Studies 2 & 3)

For the fMRI task whole-brain T2-weighted echo-planar images were collected on a Siemens 3 T Prisma MR scanner using a 20-channel head coil (transverse slice orientation, interleaved acquisition) and the following specifics: field of view = 220 mm, TR = 2000 ms, TE = 30 ms, 42 slices, slice thickness = 2 mm, voxel size = 2.0 × 2.0 × 2.0 mm, 333 volumes. One additionally structural image was acquired for co-registration during image preprocessing, using the following specifics: voxel size: 1.0 × 1.0 × 1.0 mm; TR = 1900 ms; TE = 3.42 ms; TA = 4.26; flip angle = 9 degrees; field of view = 256 × 256 mm, 192 slices with a slice thickness of 1.00 mm. For fMRI acquisition, the first twelve seconds prior to the start of the first stimulus included simultaneous multislice acquisition and dummy scans (discarded), which allowed accounting for T1 equilibration effects. The ToM task lasted 8 min and 38 s for adults, and 11 min and 4 s (5 min 32 s per run) for children. The structural image acquisition lasted 4 min and 26 s.

2.2.5. Analyses of in-scanner data (Studies 2 & 3)

In-scanner performance (correct versus incorrect answers) and differences between conditions were investigated employing one-way ANOVAs in adults and children. Bonferroni-corrected post-hoc tests were employed to further investigate significant differences between trials.

2.2.6. Whole brain fMRI analyses (Study 2 & 3)

fMRI data was analyzed using the Statistical Parametric Mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm/) in MATLAB 2016a (Mathworks, Natick, MA). Regressors of interest were created using a boxcar-function for experimental and control condition and contrasts of interest were calculated for affective ToM (AT > PC), cognitive ToM (CT > PC) and mentalizing (CT|AT > PT). For adults three additional contrasts were calculated. To detect shared activation of affective and cognitive ToM we conducted a conjunction analysis, testing areas activated in both, AT and CT conditions ((CT > PC) & (AT > PC)). Finally, contrasts of distinct activation representing affective ToM (AT > CT) and cognitive ToM (CT > AT) were calculated. For all contrasts the statistical parametric maps were cluster-level FWE-corrected at $p < 0.05$, with an initial cluster-building threshold of $p < 0.001$, uncorrected. Regressors of interest were implemented for the full trial duration of 16s (adults) or 19s (children), including story presentation and decision time. To assess whether regressor length significantly impacted neural activation, post-hoc analyses were also conducted implementing a reduced regressor (i.e., only considering the story phase of the trials in adults, excluding the decision phase).

Standard fMRI preprocessing included realignment and unwarping, co-registration to each participant's structural image and segmentation prior to normalization into standard space (ICBM152 template). All images were smoothed using an 8-mm full width at half maximum isotropic kernel. Using the art imaging toolbox (https://www.nitrc.org/projects/artifact_detect/) seven additional regressors accounting for motion and variations in mean signal intensity as well as a high-pass filter of 0.01 Hz (128 s) were added to the first-level model of each participant.

3. Results

3.1. Study 1: CAToon evaluation in children (behavioral)

3.1.1. Performance across conditions

Summarizing the ratio of correct answers in all children for each condition showed that AT trials were solved correctly in 90.7 %, CT trials in 60 % and PC trials in 73.3 % of all trials. Overall task performance results in a ratio of correct answers ranging from 40 to 97%. All children reached an accuracy above chance; Table 2. Considering only the incorrect solutions within the cognitive ToM conditions, children selected improbable scenarios in 54.9 % and highly improbable/impossible solution in 45.1 % of the cases. One-way analysis of variance showed a significant effect of condition for correct answers ($F(2,177) = 43.214$, $p < 0.001$). Bonferroni post-hoc tests for pairwise comparisons indicated that the ratio of correct answers differed significantly between all three conditions ($p < 0.001$). It has to be noted, that a direct comparison may not be warranted and has to be interpreted with great caution, since AT conditions consisted of two possibly correct endings unlike CT and PC trials (one possible ending).

3.1.2. Performance across age

A one-way ANOVA investigating the effect of age on accuracy rate was conducted after Levene's test indicated equal variances ($F(5, 52) = 0.420$, $p = 0.832$) and normal distribution within the age groups (**Supplementary Table S2**; (Field, 2013; Ghasemi and Zahediasl, 2012)). The analysis of variances showed a significant effect of age on performance, $F(552) = 39.215$, $p < 0.001$. Bonferroni post-hoc test for pairwise comparisons revealed that 3 and 4-year-olds scored

Table 2
Ratio of correct answers (in %) in the different conditions across age groups (in years).

		Age (in years)						
	Overall	3(n = 11)	4(n = 9)	5(n = 8)	6 (n = 5)	7(n = 11)	8(n = 14)	9(n = 2)
AT	89.4	76.4	81.1	93.8	88.0	100	98.6	100
CT	57.3	35.5	36.7	65.0	70.0	72.7	73.6	90
PC	71.9	57.3	52.2	78.8	76.0	79.1	87.1	100
Overall	74.6	56.4	56.7	79.1	77.6	83.9	86.4	96.7

Note: AT = affective ToM, CT = cognitive ToM, PC = Physical causality.

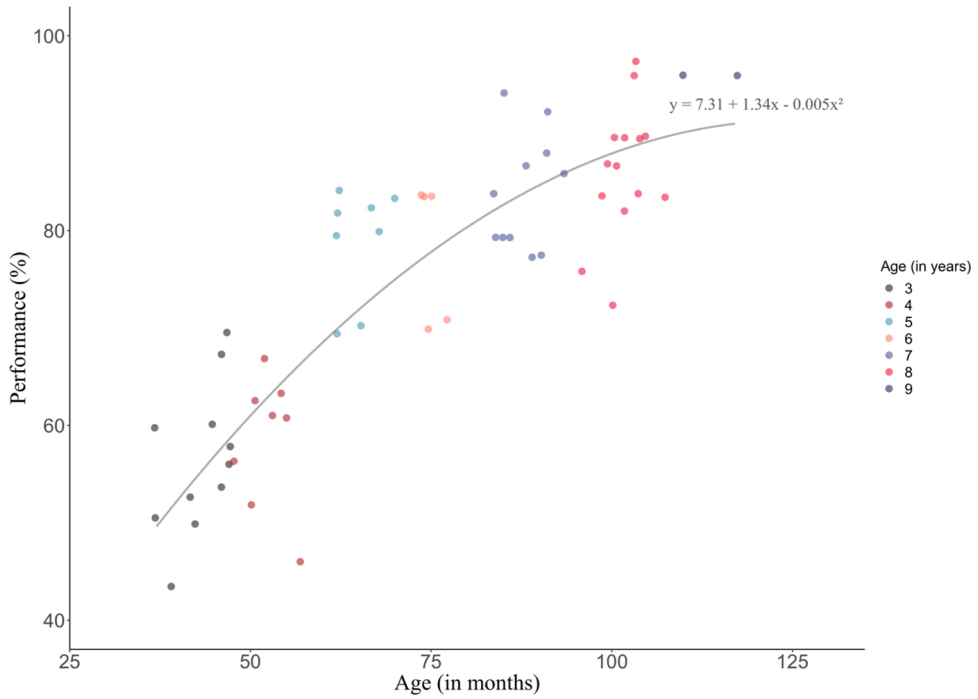


Fig. 2. Scatterplot and quadratic model best representing a performance-age relationship across all children.

significantly less correct compared to all other age groups, with no further difference between the two youngest age groups ($p = 1.000$). Also, while variations in performance remained, no significant differences between 5, 6, 7 and 8-year-olds were observed (**Supplementary Table S3**).

Assessing a dimensional age-performance model revealed a significant improvement of fit comparing the linear and quadratic models ($F(1) = 4.918$, $p = 0.031$). The age-performance relationship was best described by a quadratic model (Fig. 2), no further improvement was observed when using a cubic model. The regression model indicates that the instantaneous rate of change (f'(x)) is higher in younger ages and becomes lower in older children. This implicates bigger steps of improvement taking place in younger children. More specifically, in the youngest participants the performance is predicted to improve by 0.97 % with each passing month (f'(37) = 0.97; equaling one additional correctly solved trial every 3 months). In contrast, the rate of change drops to an improvement of 0.57 % at 77 months (f'(77) = 0.57), and decreases even further with progressing age (f'(117) = 0.17). This means, that the oldest participant within our sample is predicted to improve by one additional correct answer when aging 17 months, reflecting a deceleration of improvement in the older participants.

3.2. Studies 2 and 3: CAToon task evaluation in adults and children (fMRI)

3.2.1. In-scanner task performance

Overall, adult participants scored above chance with 87.90 % correct answers across the 30 trials (AT: 97.80 %, CT:71.90, PC: 94.10), while children scored 81.92 % correct across all trials (AT: 95.15 %, CT: 72.73 %, and PC: 80.91 %). The range of correctly answered trials was between 73–100 % for adults. Children scored between 63–93 % corrects overall. The analysis of variance showed a significant effect of condition on the ratio of correct answers in both, adults ($F(2,78) = 45.373$, $p < 0.001$) and children ($F(2,96)=24.35$, $p < 0.001$). In adults the ratio of correct answers within CT was significantly lower as compared to AT and PC (both $p < 0.001$), however there was no significant difference between AT and PC trials ($p = 0.636$) according to the Bonferroni post-hoc test for pairwise comparisons. In children, the post-hoc Bonferroni pairwise comparison yielded significant differences between AT and CT ($p < 0.001$), AT and PC ($p < 0.001$), and also PC and CT ($p = 0.041$). It has to be noted though, that a direct comparison between AT, CT and PC conditions may not be warranted and has to be interpreted with great caution, since AT conditions consisted of two possibly correct endings unlike CT and PC trials (one possible ending). When looking at the incorrect solutions within the cognitive ToM trials, adults selected improbable scenarios in 87.1 % and highly improbable/impossible solutions in 12.9 % of the cases. Children selected improbable scenarios in 62.7 % and highly improbable/impossible solution in 37.3 % of the cases.

3.2.2. Whole brain fMRI analyses

Mentalizing ($(CT \mid AT) > PT$) yielded a significant increase in activation in adults and children in frontal brain regions, including medial prefrontal, and orbitofrontal cortices, and inferior frontal gyrus. Activation increase was further observed in temporal regions, such as bilateral temporoparietal junctions, temporal poles and superior temporal sulcus. Parietal regions with heightened activity during mentalizing included inferior parietal lobule, precuneus and supramarginal and angular gyri. Further areas with an increased activation included limbic regions (e.g., cingulate cortex, insula, hippocampus), and basal nuclei (e.g., right thalamus). Affective (AT > PC) and cognitive (CT > PC) ToM-related activation increases were within expected areas, such as medial PFC, temporoparietal junction and precuneus; **see** Fig. 3, Tables 3 & 4 . The repeated analyses of the three main contrasts in the adult group with a shortened regressor including only the story-phase yielded similar activation pattern when compared to analyses implementing the full regressor (data provided through NeuroVault: https://identifiers.org/neurovault.collection:9698).

In addition to parametric correction methods, we also conducted post-hoc non-parametric correction methods in order to test the stability of our findings using different approaches. More specifically, permutation-based multiple comparison correction using SnPM (SnPM13.1.06; http://www.nisox.org/Software/SnPM13/) was computed for the main contrasts of interest (AT > PC, CT > PC, (AT|CT) > PC), employing a cluster-level inference of $p < 0.05$ FWE correction after an initial cluster-forming threshold of $p < 0.0001$ (as recommended

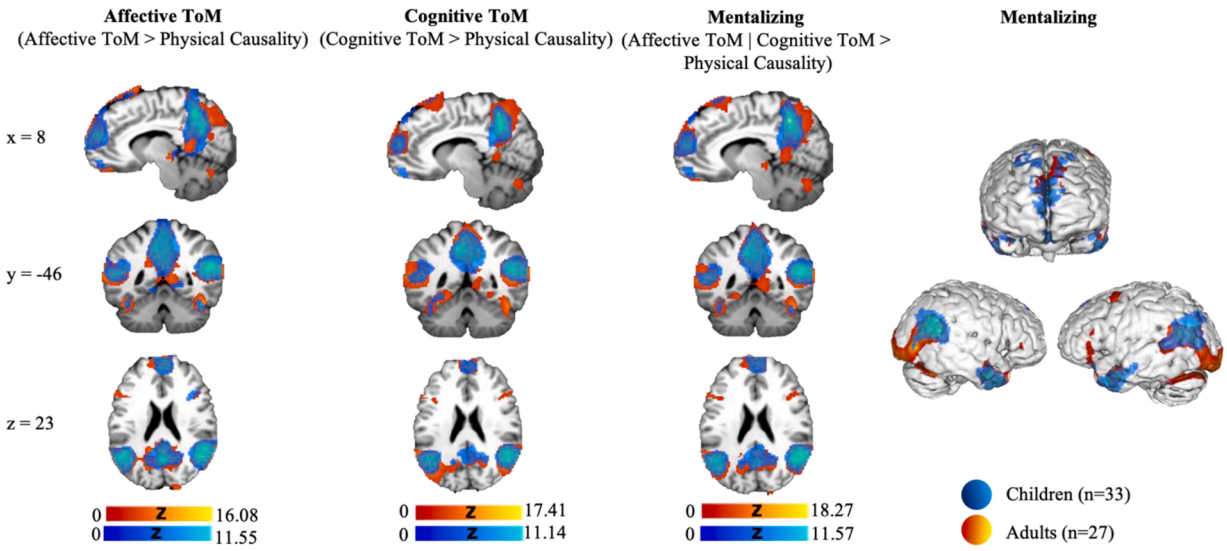


Fig. 3. Statistical parametric maps displaying neural activation during affective ToM, cognitive ToM, and mentalizing in the adult (red) and child (blue) groups (cluster-level FWE-corrected $p < 0.05$, using a cluster-building threshold of $p < 0.001$, uncorrected) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 3
Peak activation reports for affective ToM, cognitive ToM and mentalizing in the adult group.

		T	p_{FWE}	k	MNI		
Brain region	Hem.				x	y	z
Affective ToM (AT > PC)							
IFG, OFC, temporal pole, inf., mid. and sup. temporal gyri, sup. temporal sulcus, precuneus, fusiform, supramarginal and angular gyri, hippocampus, parahippocampal gyrus amygdala, insula	R/L	16.08	<0.001	25519	48	−44	18
ACC, medial mid. and sup. frontal gyri, supplementary motor area	R/L	8.39	<0.001	2756	4	56	18
medial OFC	R/L	6.81	0.001	422	4	50	−18
mid. frontal gyrus, precentral gyrus	L	6.68	0.012	345	−40	6	60
IFG (partes opercularis and triangularis)	R	4.75	0.006	249	56	30	4
Cognitive ToM (CT > PC)							
temporal pole, mid., inf.& sup. temporal gyrus, middle frontal gyrus, precuneus, angular gyrus, fusiform gyrus, insula, mid & post. cingulate cortex	R	17.41	<0.001	13605	50	−48	18
temporal pole, inf., med. & sup. temporal gyri,sup. temporal sulcus, supramarginal gyrus, parahippocampal gyrus, IFG, precuneus, fusiform gyrus, insula, amygdala	L	12.42	<0.001	11535	−48	−52	22
mid. & sup. frontal gyri, supplementary motor area, anterior & mid. cingulate cortex, precentral gyri	L	7.01	<0.001	4155	−38	2	48
mid. frontal gyrus, precentral gyrus	R	4.51	0.004	315	34	4	46
Mentalizing: ((CT | AT) > PT)							
temporal pole, inf., mid. and sup. temporal gyri, sup. temporal sulcus, fusiform gyrus, supramarginal & angular gyri, OFC, mid. frontal gyrus	R	18.28	<0.001	9459	50	−46	18
temporal pole, inf. & mid. temporal gyri, inf. parietal lobule, middle & posterior cingulate cortex, precuneus, cuneus, lingual gyrus, fusiform gyrus, OFC, insula	R/L	12.8	<0.001	17954	4	−56	34
medial sup. frontal gyrus, mid. & sup.frontal gyrus, supplementary motor area, anterior cingulate cortex	R/L	8.35	<0.001	4330	6	56	18
hippocampus, lingual gyrus, parahippocampal gyrus, fusiform gyrus, thalamus	R	5.79	0.002	367	20	−36	−12
medial OFC, gyrus rectus	R/L	5.49	0.011	261	4	50	−18
mid.frontal gyrus, precentral gyrus	R	4.96	0.028	212	36	4	46

Note. Hem = hemisphere, ACC = anterior cingulate cortex, IFG = inferior frontal gyrus, OFC = orbitofrontal cortex, inf. = inferior, mid. = middle, sup. = superior, L/R = left/right, T-scores, k = cluster size and xyz co-ordinates of peak voxel according to Montreal Neurological Institute (MNI).

by (http://www.nisox.org/Software/SnPM13/exnew)). Importantly, using non-parametric tests the relevant clusters remained similar (data provided through NeuroVault: https://identifiers.org/neurovault.collection:9699).

3.2.2.1. Shared and distinct activation for AT and CT. The conjunction of affective and cognitive trials revealed areas of shared activation in bilateral temporal poles and temporoparietal junctions, right superior temporal sulcus, anterior cingulum, precuneus, bilateral inferior frontal gyri and dorsomedial prefrontal cortex (Fig. 4, Table 5). An increase in activation was observed for affective versus cognitive ToM (AT > CT) trials in anterior precuneus, middle and posterior cingulate cortex bilaterally, inferior temporal gyrus, ventromedial prefrontal and orbitofrontal cortices. Significantly greater activation in cognitive versus affective ToM (CT > AT) trials was observed in left middle and superior frontal gyri, right insula, left inferior and middle temporal, bilateral angular and right supramarginal gyri, hippocampus, and posterior precuneus (Fig. 4, Table 5).

In order to assess whether the analysis timeframe had any effect on the neural activation obtained during cognitive and affective ToM, we re-analyzed contrasts of interest (e.g., 'AT > CT'; 'CT > AT') using a shorter regressor, which resulted in an overall similar activation pattern. However, for the cognitive trials a relative increase in activation in right insular and inferior frontal gyrus was no longer observed employing the

Table 4
Peak activation reports for affective ToM, cognitive ToM and mentalizing in children.

Brain region	Hem.	*T*	p_{FWE}	k	MNI		
					x	y	z
Affective ToM (AT > PC)							
temporal pole, paracentral lobule, precuneus, mid. & sup. temporal g., mid. & post. cingulate cortex, hippocampus, supramarginal g., angular g., lingual g., thalamus, amygdala, insula	L/R	11.55	<0.001	12972	2	−60	24
temporal pole, inf., mid., & sup. temporal g., fusiform g., angular g., lingual g, supramarginal g., parahippocampal g., hippocampus, amygdala, insula	R	10.61	<0.001	7339	46	−56	20
medial OFC, anterior cingulate c., medial sup. frontal gyrus	R/L	8.29	<0.001	2467	−2	50	18
inf. temporal gyrus, fusiform gyrus, parahippocampal gyrus	R	6.93	0.008	340	42	−44	−22
inf., mid. occipital gyrus, lingual gyrus	L	6.13	0.078	184	−20	−100	−8
mid., & sup. frontal gyrus supplementary motor area	R	5.62	0.014	296	12	36	60
precentral gyrus, inf., & mid., frontal gyrus	R	5.39	0.017	284	30	12	26
Cognitive ToM (CT > PC)							
precuneus, mid., & post. cingulate cortex, precuneus, cuneus	R/L	11.14	<0.001	4004	4	−54	40
mid., & sup. temporal g., angular gyrus, supramarginal gyrus	R	9.7	<0.001	2447	50	−56	24
mid. & sup. temporal gyri, parietal inf. lobulus, angular gyrus, supramarginal gyrus	L	8.65	<0.001	2746	−42	−64	28
temporal pole, inf., mid., & sup. temporal gyrus, inferior OFC, insula, fusiform gyrus	L	7.81	<0.001	2107	−62	−8	−20
temporal pole, inf., mid., & sup. temporal gyrus, inferior OFC, insula, fusiform gyrus	R	7.07	<0.001	1380	52	10	−34
Inf. temporal gyrus, amygdala, parahippocampal gyrus, hippocampus, fusiform gyrus, lingual gyurs	L	7.06	<0.001	892	−26	−36	−14
supplementary motor area, superior medial frontal gyrus	R/L	6.79	0.005	443	16	40	56
superior medial frontal gyrus, anterior cingulate cortex	R/L	6.31	<0.001	936	4	58	18
medial OFC	L/R	6.21	0.012	265	4	56	−18
mid., sup. frontal gyrus	L	5.31	0.006	330	−16	36	58
Mentalizing ((CT | AT) > PC)							
precuneus, paracentral lobule, middle & posterior cingulate cortex, cuneus, lingual g.	R/L	11.57	<0.001	5046	6	−52	42
temporal pole, inf., mid., & sup. temporal gyri, angular g., parahippocampal g., hippocampus, amygdala, insula, fusiform g., supramarginal g.	R	10.89	<0.001	5731	46	−56	22
temporal pole, inf., mid., sup. temporal gyrus, parahippocampal gyrus, hippocampus, amygdala, insula, supramarginal g., angular g., lingual g.	L	8.99	<0.001	6475	−40	−68	28
medial sup., & orbitofrontal gyri, anterior cingulate cortex	R/L	7.33	<0.001	2143	8	56	18
supplementary motor area, superior frontal gyrus	R/L	5.86	<0.001	494	16	40	56
inf. temporal gyrus, fusiform gyrus, parahippocampal gyrus	R	5.27	0.006	220	42	−44	−24

Note. Hem = hemisphere, ACC = anterior cingulate cortex, OFC = orbitofrontal cortex, g. = gyrus, inf. = inferior, mid. = middle, sup. = superior, L/R = left/right, T-scores, k = cluster size and xyz co-ordinates of peak voxel according to Montreal Neurological Institute (MNI).

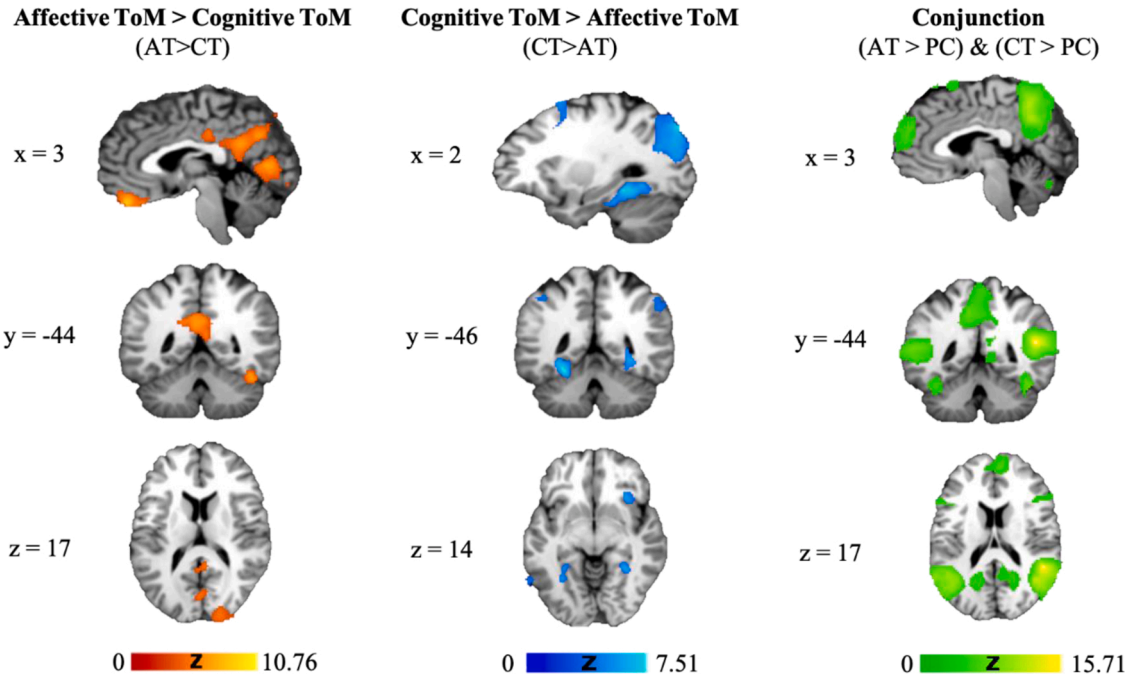


Fig. 4. Statistical parametric maps displaying shared (green: (Affective ToM > Physical Causality) & (Cognitive ToM > Physical Causality),) and distinct (red: Affective ToM > Cognitive ToM; blue: Cognitive ToM > Affective ToM) activation in adults. (cluster-level FWE-corrected $p < 0.05$, using a cluster-building threshold of $p < 0.001$, uncorrected) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 5
Peak activation reports for distinct (AT > CT), (CT > AT) and shared activation (conjunction analysis in adults ((AT > PC) & (CT > PC)).

Area	Hem.	*T*	p_{FWE}	k	MNI		
					x	y	z
AT > CT							
precuneus, mid. & post. cingulum, cuneus	L/R	10.76	<0.001	1734	−4	−58	32
gyrus rectus, medial OFC	R/L	8.6	0.001	304	0	42	−22
fusiform gyrus, cuneus, inf, mid. & sup. occipital gyrus, lingual gyrus, inf. & mid. temporal gyrus	R	8.42	<0.001	3595	48	−74	4
inferior temporal gyrus, fusiform gyrus	R	6.62	0.015	178	42	−46	−20
CT > AT							
lingual, fusiform & parahippocampal gyri	L	7.51	<0.001	496	−26	−46	−12
sup. & inf. parietal lobule, angular gyrus, middle occipital gyrus	L	7.48	<0.001	2106	−34	−82	−32
precuneus, superior parietal lobule	R	6.12	0.001	332	18	−70	58
precuneus, lingual gyrus, hippocampus, fusiform & parahippocampal gyri	R	6.01	0.004	263	30	−46	−6
mid. & sup. frontal gyrus	L	5.93	0.001	369	−18	18	42
Insula, inferior frontal gyrus	R	5.6	0.093	125	32	20	−4
angular gyrus, inf. parietal lobule	L	5.07	0.084	129	−40	−48	50
sup. parietal lobule, mid. & sup. occipital gyrus, angular gyrus	R	5.07	0.050	150	36	−80	38
mid. & sup. frontal gyrus, precentral gyrus	R	4.69	0.002	300	24	6	60
inf. & mid. temporal gyrus	L	4.47	0.11	118	−54	−56	−8
supramarginal & angular gyrus	R	4.22	0.002	185	52	−44	36
(AT > PC) & (CT > PC)							
temporal pole, inferior frontal gyrus, fusiform gyrus, inferior, middle & superior temporal gyri, sup. temporal sulcus supramarginal, angular & lingual gyri, insula, parahippocampal gyrus	R	15.72	<0.001	7188	48	−46	18
temporal pole, parancentral lobule, precuneus, supramarginal gyrus, fusiform gyrus, insula, inferior & middle temporal gyrus, mid. & post. cingulum, angular gyrus, lingual gyrus, hippocampus	R/L	11.31	<0.001	13641	4	−52	46
superior medial frontal gyrus, anterior cingulum, middle & superior frontal gyrus	R/L	6.31	<0.001	2565	6	56	18
middle frontal gyrus, precentral gyrus	L	5.52	0.0004	357	−42	2	60
inferior frontal gyrus	R	4.58	0.0003	374	56	30	4

Note. Hem. = hemisphere, OFC = orbitofrontal cortex, inf. = inferior, mid. = middle, sup. = superior, post. = posterior, L/R = left/right, T-scores, k = cluster size and xyz co-ordinates of peak voxel according to Montreal Neurological Institute (MNI).

shorter regressor (data provided through NeuroVault: https://identifiers.org/neurovault.collection:9698).

3.2.3. Post-hoc analyses in adults

3.2.3.1. Callous-unemotional traits/empathy and neuronal activation during mentalizing. Two post-hoc partial correlation analyses were conducted in the adult group to investigate the association of callous-unemotional traits and empathy with activation in right temporoparietal junction during mentalizing ((AT | CT)>PC). The analyses were controlled for age and gender and Bonferroni correction for multiple comparison testing was used to adjust for the number of tests conducted ($p<(0.05/2)$). Mean parameter estimates from the right temporoparietal junction as defined according Dufour et al. (2013; retrieved from: http://saxelab.mit.edu/use-our-theory-mind-group-maps) and extracted through the MarsBar toolbox [Brett et al., 2002; http://marsbar.sourceforge.net/]. The right temporoparietal junction was chosen as an area of interest because of its key implication across a wide range of mentalizing tasks and studies (Döhnel et al., 2012; Mahy et al., 2014; Powell et al., 2017; Saxe, 2010). Partial correlation analyses revealed a significant negative correlation between CU-traits and neural activation during mentalizing in right temporoparietal junction ($r(23) = -0.533$, $p = 0.006$, but no significant association between empathy levels and activation in right temporoparietal junction ($r(23) = 0.366$, $p = 0.072$). Further partial correlation tests revealed no significant relationship between callous-unemotional traits and (1) motion during scan (as measured by the average head motion during task or the number of outliers over 1.5mm), and (2) performance on the task, in order to test confounds (**Supplementary Table S5**).

4. Discussion

Here we evaluate feasibility and neural activation patterns evoked by CAToon, a newly developed child-friendly and open-source fMRI Theory of Mind cartoon task. Evaluation included one behavioral study (**Study 1**; behavioral assessment of 60 children; 3–9 years) and two neural evaluations (**Study 2**: fMRI in 27 adults and **Study 3:** fMRI conducted in 33 children). Behavioral results support task feasibility as early as three years of age. However, reliable performance skills are reached around 5 years, which we suggest as an ideal age for fMRI task implementation. fMRI evidence in children and adults confirmed that CAToon is associated with significant activation increases in brain regions associated with mentalizing (e.g., dorsomedial PFC, ventromedial PFC, bilateral temporoparietal junction, middle temporal gyrus, posterior superior temporal sulcus, precuneus, inferior frontal gyrus, precentral gyrus, anterior cingulate cortex and temporal pole). Affective and cognitive ToM trials led to brain activation increases of shared (e.g., bilateral temporal pole, temporoparietal junction, superior temporal sulcus, precuneus and parts of the dorsomedial prefrontal cortex) and distinct brain regions (e.g., AT-specific: orbitofrontal cortex, anterior parts of the precuneus, posterior cingulate cortex, CT-specific: right insula, parahippocampal and fusiform gyrus and posterior potions of the precuneus). Moreover, activation increases in the right temporoparietal junction were negatively correlated with levels of callous-unemotional traits, but not empathy, in adults.

4.1. Feasibility of the CAToon task for children

Behavioral data (**Study 1**) and fMRI data acquisition (**Study 3**) revealed that children of all ages tested were able to complete CAToon above chance level. More specifically, **Study 1** indicated that while all children were able to complete CAToon, children aged five years and up performed significantly better than three and four-year-olds. While children of five years and older still displayed variations in performance, no further significant change in task performance was observed, indicating reliable task performance. Behaviorally, children were most accurate in the affective ToM condition, followed by physical causality and cognitive trials. These findings have to be considered with caution, however, since outcome options were not identical for all conditions (e.g., two possible correct endings for AT compared to CT and PC

conditions).

An increasing performance accuracy of children ages five and up as reported here is in line with previous evidence of children performing reliably on explicit ToM tasks starting around four to six years of age (Frith and Frith, 2003; Wellman et al., 2001). Notably, implicit ToM tasks reveal false belief understanding in infants already (Southgate et al., 2007; Surian et al., 2007). However, demands posited by an explicit and/or fMRI task require complementary skills to basic false belief understanding (Lillard and Kavanaugh, 2014). Younger children have been reported to be more challenged by or fail mentalizing tasks that require inhibitory control and working memory (Carlson et al., 2002; Müller et al., 2012; Rakoczy, 2010; Scott and Baillargeon, 2017). The observed performance improvements may result from individual improvements in ToM skills and/or maturation of executive functions typically observed around this age (Roebers et al., 2011; Röthlisberger et al., 2010). Such skill improvements have been linked to the start of formal schooling (e.g., (Brod et al., 2017; Roebers et al., 2011)). For fMRI purposes we therefore recommend the use of CAToon starting around the age of five years and up, which considers increased challenges posed by an MRI environment (Raschle et al., 2012, 2009).

The use of implicit ToM fMRI tasks by passive movie has shown to be possible in children as young as three years of age (Richardson et al., 2018). Here we additionally evaluated neural activation associated with the story-phase of the CAToon trials only (as compared to the implementation of regressors that include the story and explicit answer phase) in adults, with comparable outcome. While this may be viewed as a first step towards testing CAToon's suitability as a potential passive viewing task, future investigations in younger children are warranted.

4.2. Neural correlates of mentalizing using CAToon in young adults

Study 2 revealed robust activation increases in brain areas commonly associated with mentalizing for adults, including the dorsomedial and ventromedial PFC, bilateral temporoparietal junction, middle temporal gyrus, posterior superior temporal sulcus, precuneus, inferior frontal gyrus, precentral gyrus, anterior cingulate cortex and temporal poles (Blakemore, 2012; Bzdok et al., 2012; Molenberghs et al., 2016; Van Overwalle and Baetens, 2009). More specifically, the role of the temporoparietal junction during mentalizing is supported by evidence associating this region to temporary mental state attribution of self and others (Mahy et al., 2014; Molenberghs et al., 2016; Van Overwalle and Baetens, 2009). Our findings are further in line with studies demonstrating an involvement of the bilateral temporal pole in context-specific mentalizing (C. D. Frith and Frith, 2006), ventromedial prefrontal cortex in social cognition and self-perception (Amodio and Frith, 2006) and precentral gyrus in the differentiation of self and other (Aichhorn et al., 2006; Ruby and Decety, 2001). Our data supports an involvement of regions specific for affective aspects of mentalizing (e.g., empathic judgment, emotion processing or empathy), including the middle temporal gyrus, ventromedial prefrontal, anterior cingulate and orbitofrontal cortex (Lamm and Singer, 2010; Molenberghs et al., 2016; Northoff and Bermpohl, 2004; Roy et al., 2012; Völlm et al., 2006). Further areas identified, include the insula [recognition and selection of salient events; (Menon and Uddin, 2010)], fusiform gyrus [face processing; (Kanwisher and Yovel, 2006)], right superior temporal sulcus [linked to the observation of socially relevant bodily cues; (Allison et al., 2000; Lee et al., 2012)], precuneus, parahippocampal gyrus and hippocampus [episodic memory retrieval; (Cavanna and Trimble, 2006; Spreng et al., 2009)].

Through use of conjunction analyses we observed shared and distinct activation patterns when further investigating affective and cognitive ToM trials, which is in line with past evidence (Bodden et al., 2013; Hynes et al., 2006; Schlaffke et al., 2015; Sebastian et al., 2012a). Brain regions that were implicated during both affective and cognitive ToM included bilateral temporal pole, temporoparietal junction, right superior temporal sulcus, anterior cingulum, precuneus, bilateral inferior frontal gyri and parts of the dorsomedial prefrontal cortex.

Activation was greater for affective as compared to cognitive trials within the anterior part of the precuneus extending into the posterior cingulate cortex, as well as within the cuneus and orbitofrontal cortex. This pattern remained when analyzing only the story portion of the trials, indicating that passive viewing of the CAToon stories may be sufficient to induce affective mentalizing. A distinct activation of the orbitofrontal and ventromedial prefrontal cortex in affective ToM is in line with literature emphasizing its role for affective processing (Hynes et al., 2006; Molenberghs et al., 2016; Schlaffke et al., 2015; Shamay-Tsoory et al., 2010). Similarly, the posterior cingulate, has been linked to empathetic perspective taking (Schlaffke et al., 2015; Völlm et al., 2006). In contrasts to previous findings (Schlaffke et al., 2015), we have not detected distinct activation in basal ganglia for affective compared to cognitive ToM. This may result from different task designs, as Schlaffke et al. (2015) measured affective and cognitive ToM through the use of the same set of images, but different questions, while CAToon included distinct trials for each condition.

Areas with increased activation during cognitive versus affective trials included posterior parts of the precuneus, parahippocampal gyrus, hippocampus and right insula. However, removing the explicit decision phase from the model, the right insula and inferior frontal gyrus did not remain significant. Both regions have been linked to decision making processes (Hartwright et al., 2016; Paulus et al., 2005), which may explain why a shortened model, excluding the decision phase, no longer results in activation increases of these areas.

4.3. Callous-unemotional traits, empathy and mentalizing

We observed a negative association between callous-unemotional traits and neural activation within the right temporoparietal junction in adults. The right temporoparietal junction is most commonly implicated when inferring about thoughts, beliefs and emotional states (Molenberghs et al., 2016). Within the limited literature investigating the relationship between callous-unemotional traits and neural correlates of mentalizing, our findings support those establishing a negative link (Lockwood et al., 2013; Sebastian et al., 2012b) between callous-unemotional traits and mentalizing skills. They may thus be in line with evidence suggesting that adults with higher levels of callous-unemotional traits are more likely to disregard others' feelings (Scheepers et al., 2011) or more likely to display deficient affective perspective taking (Lui et al., 2016). However, past findings are inconclusive, with some reporting a positive association (Gao et al., 2019) or those missing to find a significant connection (O'Nions et al., 2014). Notably, levels of callous-unemotional traits displayed within our adult group did not correlate with an increase in motion during fMRI task performance.

4.4. Neural correlates of mentalizing using CAToon in children

After successfully evaluating the neural correlates associated with CAToon task performance in adults (**Study 2**), **Study 3** further assessed the neural correlates in a group of children (ages 7–13 years), thus testing feasibility in an initial fMRI study of children using CAToon. We observed activation of the mentalizing network comparable to findings in our adult study. Activation clusters in the child group were similar, though seemed slightly less pronounced as reported in adults, which is in line with studies investigating adult and developmental populations during mentalizing (Fehlbaum et al., 2021; Richardson et al., 2018). An increase in neural activation was observed in areas including bilateral temporoparietal junction, medial prefrontal cortex and precuneus (in line with (Gweon et al., 2012; Richardson and Saxe, 2020b)). Our findings provide first evidence for the feasibility of employing CAToon as an fMRI task in children. An effort in developing and subsequently sharing age-appropriate neuroimaging tasks may further replicability and reproducibility of findings (Klapwijk et al., 2019). To provide

opportunity for others using CAToon or our findings further, the task and all T-maps reported in the manuscript are made openly available at https://www.jacobscenter.uzh.ch/en/research/developmental_neuroscience/downloads/catoon.html.

CAToon adds to previous tasks used in children by measuring affective and cognitive aspects of mentalizing. However, it has to be highlighted that the two conditions are not as well isolated (e.g., Sebastian et al., 2012a had distinct cognitive and affective trials) or matched (e.g., Schlaffke et al., 2015 implementing the same images for cognitive and affective trials, but different questions asked), since for CAToon cognitive and affective ToM both include people and affective elements. While naturalistic (e.g., cognitive and affective ToM are social processes, rarely isolated from humans or fully free of affect in real life), this implementation results in a certain confound rendering it challenging to fully isolate individual aspects. Future studies could further test individual stimuli ratings by content, which may then be associated with neural activation in each condition to test the influence of different stimulus characteristics. CAToon was employed as an explicit task, including in-scanner responses. However, first analyses reveal its potential as a passive viewing task, which might be more appropriate for very young participants.

4.5. Study limitations

Several limitations have to be noted. In **Study 1** (behavioral evaluation in children), CAToon was presented either in a one-on-one setting or, for older children, in groups. This procedure ensured appropriate understanding and task conduction for younger children, but limits comparability across all age groups. It is notable that younger children performed still lower than children ages five and up. It is also mentionable that there was no strict timing or time limit for the image presentation and decision phase within the behavioral study, which is not possible when running CAToon within an fMRI setting.

Additionally, since characters' emotional expressions were included in cognitive ToM scenarios, the affective and cognitive scenarios presented in CAToon are less clearly distinct as compared to previous paradigms (e.g., Sebastian et al., 2012a). While the correct solving of the CT and AT trials is designed to rely on the inference about the targeted mental states (i.e., intentions in cognitive and emotions in affective trials), a direct investigation about the cognitive process underlying participants' answers (i.e., understanding what inferences they make during certain trials) should be investigated in future studies (e.g., collecting subjective responses of participants' reasoning for the answer selection). However, neural activation in adults provides initial evidence of the conditions eliciting activation in established areas (e.g., increased recruitment of ventromedial prefrontal and orbitofrontal cortex during affective trials)."

A deliberate choice to include two correct answers within the affective trials was made in order to be able to investigate positive/negative expectancy in later studies (e.g., of children with and without disruptive behaviors or prior maltreatment). This might be considered as a caveat as it makes the direct comparison of behavioral performance between the different trial types challenging. From a neural perspective, it might be hypothesized that the inclusion of two possible correct answers may require children to evaluate their response even more, thus increasing the need for mentalizing. However, such an effect will have to be further evaluated.

We would also like to note, that while there are variations in behavioral task performance across all ages, there are no significant improvements after the age of 5 years when measured outside the fMRI environment (**Study 1**), or after the age of 7 years when using CAToon inside the fMRI environment (**Study 3**). In future studies, the use of an additional, established behavioral measure is recommended in order to establish whether children's performance and neural activation is clearly associated. Also, CAToon task stimuli were not tested in different cultures, limiting generalizability and highlighting opportunities for future investigation. Due to the small sample size for brain-behavioral correlations (e.g., comparison of neural activation with callous-unemotional traits/empathy) such findings must be interpreted with caution (Cremers et al., 2017).

Data availability

The CAToon task, including manual, training and fMRI task are freely available at:
https://www.jacobscenter.uzh.ch/en/research/developmental_neuroscience/downloads/catoon.html

All T-maps associated with data from the present manuscript can be found at:
https://www.jacobscenter.uzh.ch/en/research/developmental_neuroscience/downloads/catoon.html

Funding

This study was funded through a Jacobs Foundation Early Career Research Grant (Nr. 2016201713) and an early career research grant by the University of Basel (both to NMR).

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

The authors thank all the families, children and adults that have participated in this research. We thank Anja Stirnimann, Johannes Hopf and Maria Irena Dobrowolska for their help with the families during neuroimaging and behavioral sessions. We also thank the radiology team, particularly Tanja Haas, Philipp Madörin, and Dr. Francesco Santini, for their support during the neuroimaging session.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.dcn.2021.100959.

References

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Do visual perspective tasks need theory of mind? Neuroimage 30 (3), 1059–1068.
Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues: role of the STS region. Trends Cogn. Sci. (Regul. Ed.) 4 (7), 267–278.
Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat. Rev. Neurosci. 7 (4), 268.
Andershed, H.A., Kerr, M., Stattin, H., Levander, S., 2002. Psychopathic Traits in Non-referred Youths: a New Assessment Tool.
Baron-Cohen, S., Leslie, A.M., Frith, U., 1985. Does the autistic child have a "theory of mind"? Cognition 21 (1), 37–46.
Blakemore, S.-J., 2012. Development of the social brain in adolescence. J. R. Soc. Med. 105 (3), 111–116.
Bodden, M.E., Kübler, D., Knake, S., Menzler, K., Heverhagen, J.T., Sommer, J., et al., 2013. Comparing the neural correlates of affective and cognitive theory of mind using fMRI: Involvement of the basal ganglia in affective theory of mind. Adv. Cogn. Psychol. 9 (1), 32.
Brett, M., Anton, J.L., Valabregue, R., Poline, J.B., 2002. Region of interest analysis using an SPM toolbox. In 8th international conference on functional mapping of the human brain 16 (2), 497.
Brod, G., Bunge, S.A., Shing, Y.L., 2017. Does one year of schooling improve children's cognitive control and alter associated brain activation? Psychol. Sci. 28 (7), 967–978.
Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A.R., Langner, R., Eickhoff, S. B., 2012. Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct. Funct. 217 (4), 783–796. https://doi.org/10.1007/s00429-012-0380-y.
Carlson, S.M., Moses, L.J., Breton, C., 2002. How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. Infant Child Dev. 11 (2), 73–92.
Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. Brain 129 (3), 564–583.

Clegg, J., Hollis, C., Mawhood, L., Rutter, M., 2005. Developmental language disorders–a follow-up in later adult life. Cognitive, language and psychosocial outcomes. J. Child Psychol. Psychiatry 46 (2), 128–149.
Cremers, H.R., Wager, T.D., Yarkoni, T., 2017. The relation between statistical power and inference in fMRI. PLoS One 12 (11), e0184923.
Daseking, M., Petermann, U., Petermann, F., 2007. Intelligenzdiagnostik mit dem HAWIK-IV. Kindheit und Entwicklung 16 (4), 250–259.
Davis, M.H., 1980. A Multidimensional Approach to Individual Differences in Empathy.
Döhnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., Sommer, M., 2012. Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. Neuroimage 60 (3), 1652–1661.
Dufour, N., Redcay, E., Young, L., Mavros, P., L, Moran, J., M, Triantafyllou, C., et al., 2013. Similar brain activation during false belief tasks in a large sample of adults with and without autism. PLoS One 8 (9), e75468.
Fehlbaum, L., Borbás, R., Paul, K., Raschle, N., 2021. Early and late neural correlates of mentalizing: ALE meta-analyses in adults, children, adolescents. doi:10.31219/osf.io/d8fq2.
Field, A., 2013. Discovering Statistics Using IBM SPSS Statistics. Sage.
Frith, U., 2001. Mind blindness and the brain in autism. Neuron 32 (6), 969–979.
Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 358 (1431), 459–473.
Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. Neuron 50 (4), 531–534.
Frith, C.D., Frith, U., 2008. Implicit and explicit processes in social cognition. Neuron 60 (3), 503–510.
Gallagher, H.L., Frith, C.D., 2003. Functional imaging of 'theory of mind'. Trends Cogn. Sci. (Regul. Ed.) 7 (2), 77–83.
Gao, Y., Rogers, J.C., Pauli, R., Clanton, R., Baker, R., Birch, P., et al., 2019. Neural correlates of theory of mind in typically-developing youth: Influence of sex, age and callous-unemotional traits. Sci. Rep. 9 (1), 1–12.
Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. Int. J. Endocrinol. Metab. 10 (2), 486–489. https://doi.org/10.5812/ijem.3505.
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.
Hartwright, C., Hansen, P.C., Apperly, I.A., 2016. Current knowledge on the role of the inferior frontal gyrus in theory of Mind-A commentary on schurz and Tholen (2016). Cortex 85, 133–136.
Huijbers, W., Van Dijk, K.R., Boenniger, M.M., Stirnberg, R., Breteler, M.M., 2017. Less head motion during MRI under task than resting-state conditions. Neuroimage 147, 111–120.
Hyde, D.C., Simon, C.E., Ting, F., Nikolaeva, J.I., 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.
Hynes, C.A., Baird, A.A., Grafton, S.T., 2006. Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. Neuropsychologia 44 (3), 374–383.
Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. Biol. Sci. 361 (1476), 2109–2128.
Klapwijk, E., van den Bos, W., Tamnes, C.K., Mills, K., Raschle, N., 2019. Opportunities for Increased Reproducibility and Replicability of Developmental Cognitive Neuroscience.
Korkmaz, B., 2011. Theory of mind and neurodevelopmental disorders of childhood. Pediatr. Res. 69 (5 Pt 2), 101R–108R.
Lamm, C., Singer, T., 2010. The role of anterior insular cortex in social emotions. Brain Struct. Funct. 214 (5–6), 579–591.
Lee, S.M., Gao, T., McCarthy, G., 2012. Attributing intentions to random motion engages the posterior superior temporal sulcus. Soc. Cogn. Affect. Neurosci. 9 (1), 81–87.
Lillard, A.S., Kavanaugh, R.D., 2014. The contribution of symbolic skills to the development of an explicit theory of mind. Child Dev. 85 (4), 1535–1551.
Lockwood, P.L., Sebastian, C.L., McCrory, E.J., Hyde, Z.H., Gu, X., De Brito, S.A., Viding, E., 2013. Association of callous traits with reduced neural response to others' pain in children with conduct problems. Curr. Biol. 23 (10), 901–905. https://doi.org/10.1016/j.cub.2013.04.018.
Lui, J.H., Barry, C.T., Sacco, D.F., 2016. Callous-unemotional traits and empathy deficits: mediating effects of affective perspective-taking and facial emotion recognition. Cogn. Emot. 30 (6), 1049–1062.
Mahy, C.E., Moses, L.J., Pfeifer, J.H., 2014. How and where: theory-of-mind in the brain. Dev. Cogn. Neurosci. 9, 68–81.
Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct. Funct. 214 (5–6), 655–667.
Molenberghs, P., Johnson, H., Henry, J.D., Mattingley, J.B., 2016. Understanding the minds of others: a neuroimaging meta-analysis. Neurosci. Biobehav. Rev. 65, 276–291.
Moraczewski, D., Chen, G., Redcay, E., 2018. Inter-subject synchrony as an index of functional specialization in early childhood. Sci. Rep. 8 (1), 1–12.
Müller, U., Liebermann-Finestone, D.P., Carpendale, J.I., Hammond, S.I., Bibok, M.B., 2012. Knowing minds, controlling actions: the developmental relations between theory of mind and executive function from 2 to 4 years of age. J. Exp. Child Psychol. 111 (2), 331–348.
Northoff, G., Bermpohl, F., 2004. Cortical midline structures and the self. Trends Cogn. Sci. (Regul. Ed.) 8 (3), 102–107.
O'Connell, S.M., 1995. Empathy in chimpanzees: evidence for theory of mind? Primates 36 (3), 397–410.
O'Nions, E., Sebastian, C.L., McCrory, E., Chantiluke, K., Happe, F., Viding, E., 2014. Neural bases of Theory of Mind in children with autism spectrum disorders and children with conduct problems and callous-unemotional traits. Develomental Science 17 (5), 786–796. https://doi.org/10.1111/desc.12167.
Onishi, K.H., Baillargeon, R., 2005. Do 15-month-old infants understand false beliefs? Science 308 (5719), 255–258.
Paulus, M.P., Feinstein, J.S., Leland, D., Simmons, A.N., 2005. Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. Neuroimage 25 (2), 607–615.
Powell, J.L., Grossi, D., Corcoran, R., Gobet, F., Garcia-Finana, M., 2017. The neural correlates of theory of mind and their role during empathy and the game of chess: a functional magnetic resonance imaging study. Neuroscience 355, 149–160.
R Core Team, 2020. R: a Language and Environment for Statistical Computing. URL. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
Rakoczy, H., 2010. Executive function and the development of belief–desire psychology. Dev. Sci. 13 (4), 648–661.
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 https://doi.org/10.3791/1309.
Raschle, N., Zuk, J., Ortiz-Mantilla, S., Sliva, D.D., Franceschi, A., Grant, P.E., et al., 2012. Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. Ann. N. Y. Acad. Sci. 1252, 43–50. https://doi.org/10.1111/j.1749-6632.2012.06457.x.
Richardson, H., Saxe, R., 2020a. Development of predictive responses in theory of mind brain regions. Dev. Sci. 23 (1), e12863.
Richardson, H., Saxe, R., 2020b. Early signatures of and developmental change in brain regions for theory of mind. Neural Circuit and Cognitive Development. Elsevier, pp. 467–484.
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), 1–12.
Roebers, C.M., Röthlisberger, M., Cimeli, P., Michel, E., Neuenschwander, R., 2011. School enrolment and executive functioning: a longitudinal perspective on developmental changes, the influence of learning context, and the prediction of pre-academic skills. Eur. J. Dev. Psychol. 8 (5), 526–540.
Röthlisberger, M., Neuenschwander, R., Michel, E., Roebers, C.M., 2010. Exekutive funktionen: zugrundeliegende kognitive prozesse und deren korrelate bei kindern im späten vorschulalter. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie.
Roy, M., Shohamy, D., Wager, T.D., 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn. Sci. (Regul. Ed.) 16 (3), 147–156.
Ruby, P., Decety, J., 2001. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. Nat. Neurosci. 4 (5), 546.
Saxe, R., 2006. Uniquely human social cognition. Curr. Opin. Neurobiol. 16 (2), 235–239.
Saxe, R., 2010. The right temporo-parietal junction: a specific brain region for thinking about thoughts. Handbook of Theory of Mind, pp. 1–35.
Scheepers, F.E., Buitelaar, J.K., Matthys, W., 2011. Conduct Disorder and the specifier callous and unemotional traits in the DSM-5. Eur. Child Adolesc. Psychiatry 20 (2), 89–93. https://doi.org/10.1007/s00787-010-0149-x.
Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., et al., 2015. Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. Hum. Brain Mapp. 36 (1), 29–39.
Schnell, K., Bluschke, S., Konradt, B., Walter, H., 2011. Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. Neuroimage 54 (2), 1743–1754.
Scott, R.M., Baillargeon, R., 2017. Early false-belief understanding. Trends Cogn. Sci. (Regul. Ed.) 21 (4), 237–249.
Sebastian, C.L., Fontaine, N.M., Bird, G., Blakemore, S.J., Brito, S.A., McCrory, E.J., Viding, E., 2012a. Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. Soc. Cogn. Affect. Neurosci. 7 (1), 53–63. https://doi.org/10.1093/scan/nsr023.
Sebastian, C.L., McCrory, E.J., Cecil, C.A., Lockwood, P.L., De Brito, S.A., Fontaine, N.M., Viding, E., 2012b. Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. Arch. Gen. Psychiatry 69 (8), 814–822. https://doi.org/10.1001/archgenpsychiatry.2011.2070.
Senju, A., 2012. Spontaneous theory of mind and its absence in autism spectrum disorders. Neuroscientist 18 (2), 108–113.
Shamay-Tsoory, S.G., Aharon-Peretz, J., 2007. Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. Neuropsychologia 45 (13), 3054–3067.
Shamay-Tsoory, S.G., Harari, H., Aharon-Peretz, J., Levkovitz, Y., 2010. The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. Cortex 46 (5), 668–677.
Sharp, C., 2008. Theory of Mind and conduct problems in children: deficits in reading the "emotions of the eyes". Cogn. Emot. 22 (6), 1149–1158.
Southgate, V., Senju, A., Csibra, G., 2007. Action anticipation through attribution of false belief by 2-year-olds. Psychol. Sci. 18 (7), 587–592.
Spreng, R.N., Mar, R.A., Kim, A.S., 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J. Cogn. Neurosci. 21 (3), 489–510.
Surian, L., Caldi, S., Sperber, D., 2007. Attribution of beliefs by 13-month-old infants. Psychol. Sci. 18 (7), 580–586.

Thieba, C., Frayne, A., Walton, M., Mah, A., Benischek, A., Dewey, D., Lebel, C., 2018. Factors associated with successful MRI scanning in unsedated young children. Front. Pediatr. 6.
Uekermann, J., Kraemer, M., Abdel-Hamid, M., Schimmelmann, B., G, Hebebrand, J., Daum, I., et al., 2010. Social cognition in attention-deficit hyperactivity disorder (ADHD. Neurosci. Biobehav. Rev. 34 (5), 734–743.
Van Overwalle, F., Baetens, K., 2009. Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. Neuroimage 48 (3), 564–584.
Völlm, B.A., Taylor, A.N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al., 2006. Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. Neuroimage 29 (1), 90–98.
Walter, H., 2012. Social cognitive neuroscience of empathy: concepts, circuits, and genes. Emot. Rev. 4 (1), 9–17.
Walter, H., Ciaramidaro, A., Adenzato, M., Vasic, N., Ardito, R.B., Erk, S., Bara, B.G., 2009. Dysfunction of the social brain in schizophrenia is modulated by intention type: an fMRI study. Soc. Cogn. Affect. Neurosci. 4 (2), 166–176.
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.
Wiesmann, C.G., 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.
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.
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.