

Social perspective taking is associated with self-reported prosocial behavior and regional cortical thickness across adolescence

Christian K. Tamnes¹, Knut Overbye², Lia Ferschmann², Anders M. Fjell², Kristine B. Walhovd², Sarah-Jayne Blakemore³, and Iroise Dumontheil⁴

¹ Department of Psychology, University of Oslo, Oslo, Norway

² Center for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo, Oslo, Norway

³ Institute of Cognitive Neuroscience, University College London, London, UK

⁴ Department of Psychological Sciences, Birkbeck, University of London, UK

Corresponding author: Christian K. Tamnes, Department of Psychology, University of Oslo, PO Box 1094 Blindern, 0317 Oslo, Norway; Email: c.k.tamnes@psykologi.uio.no

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgments: This study was supported by grants from the Research Council of Norway (to CKT, AMF, KBW), the University of Oslo (to CKT, AMF, KBW), The Royal Society (to SJB), Jacobs Foundation (to SJB), and The Wellcome Trust (to SJB). The authors thank all participating individuals and families, and the psychology undergraduate students that collected data from high school students.

Abstract

Basic perspective taking and mentalising abilities develop in childhood, but recent studies indicate that the use of social perspective taking to guide decisions and actions has a prolonged development that continues throughout adolescence. Here, we aimed to replicate this research and investigate the hypotheses that individual differences in social perspective taking in adolescence are associated with real-life prosocial and antisocial behavior and differences in brain structure. We employed an experimental approach and a large cross-sectional sample (n=293) of participants aged 7-26 years old to assess age-related improvement in social perspective taking usage during performance of a version of the Director task. In subsamples, we then tested how individual differences in social perspective taking were related to self-reported prosocial behavior and peer relationship problems on the Strengths and Difficulties Questionnaire (SDQ) (n=184) and to magnetic resonance imaging (MRI) measures of regional cortical thickness and surface area (n=226). The pattern of results in the Director task replicated previous findings by demonstrating continued improvement in use of social perspective taking across adolescence. The study also showed that better social perspective taking usage is associated with more self-reported prosocial behavior, as well as to thinner cerebral cortex in regions in the left hemisphere encompassing parts of the caudal middle frontal and precentral gyri and lateral parietal regions. These associations were observed independently of age, and might partly reflect individual developmental variability. The relevance of cortical development was additionally supported by indirect effects of age on social perspective taking usage via cortical thickness.

Keywords: brain structure; cerebral cortex; development; MRI; social cognition; theory of mind

Introduction

Fundamental aspects of perspective taking emerge in the second year of life (Moll & Tomasello, 2006; Sodian, Thoermer, & Metz, 2007), while theory of mind or mentalising, the ability to attribute mental states such as beliefs, desires and intentions, is thought to develop in early childhood. Classic theory of mind tasks that examine the explicit understanding of other people's false beliefs or second-order beliefs are typically passed around age 4 and 7 years, respectively (Perner & Wimmer, 1985; Wellman, Cross, & Watson, 2001). More recent studies however, indicate that the use of social perspective taking, that is the ability to take the perspective of another individual into account in a communicative context and use this information to guide decisions and actions, continues to develop across childhood and adolescence (Begeer et al., 2016; Dumontheil, Apperly, & Blakemore, 2010; Mills, Dumontheil, Speekenbrink, & Blakemore, 2015; Symeonidou, Dumontheil, Chow, & Breheny, 2016). In the present study, we aimed to replicate this finding and to extend our understanding of social perspective taking across multiple levels of individual differences data by relating it to real-life social behavior, on the one hand, and cerebral cortex structure, on the other.

Prolonged development of social perspective taking has been found in studies using variants of the Director task (Apperly et al., 2010; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003). This is an experimental paradigm in which participants view sets of shelves containing objects, which they are instructed to move by an avatar (the 'Director') who can see some but not all of the objects. Correct interpretation of critical instructions in the experimental condition requires participants to use the Director's perspective and to move only objects that the Director can see. In a control condition, participants are asked to ignore certain objects according to a simple visual rule, specifically to move objects only in clear shelf slots and ignore objects in slots with a gray background. Dumontheil et al. (2010) tested a large sample of female participants in the age-range 7-27 years using this task and found that accuracy in the perspective taking condition continued to improve between adolescence and adulthood. As successful perspective taking in this task also involves inhibiting one's own perspective and integrating one's goals with the context, this prolonged development might reflect interactions between perspective taking and developing executive functions (Dumontheil, Apperly, et al., 2010). Two smaller studies using versions of the Director task have replicated the finding of continued development of social perspective taking usage across adolescence (Humphrey & Dumontheil, 2016; Symeonidou et al., 2016).

In summary, several studies have demonstrated that, contrary to earlier assumptions that mentalising stops developing in early childhood, the ability to use someone else's perspective to guide ongoing behavior is still developing throughout adolescence. A next step in this theoretical framework is to ask how perspective taking, as measured in a lab-based experimental task, relates to real world social behavior in adolescence. As social perspective taking is necessary to understand that someone else might think and feel differently than you do, it is thought to be a key building block of both sympathy and empathy, and to in turn foster prosocial behavior (Decety, Barta, Uzefovsky, & Knafo-Noam, 2016). Conversely, poor social perspective taking ability might lead to social maladjustment and peer problems. In a recent study, Sierksma et al. (2014) interviewed children about helping situations in vignettes that varied in the recipient's need for help and in the costs to the helper. The results showed that, when both need and costs were high, social perspective taking ability, measured via a separate task requiring understanding of a false evaluation of another character, was positively associated with stronger moral indignation against a character refusing to help another. There is a large body of evidence for a small positive association between children's theory of mind scores and concurrent measures of prosocial behavior, and this association appears to hold for both cognitive and affective theory of mind and for different subtypes of prosocial behavior (Imuta, Henry, Slaughter, Selcuk, & Ruffman, 2016). Longitudinal studies with children additionally support a mediational hypothesis of an indirect path from theory of mind to subsequent lower peer rejection and higher peer acceptance, via improvements in prosocial behavior (Caputi,

Lecce, Pagnin, & Banerjee, 2012), and suggest that aspects of theory of mind performance inversely predict later reactive and proactive aggression (Austin, Bondu, & Elsner, 2017). A hypothesis tested in the current study is that individual differences in social perspective taking usage will be associated with individual differences in real-life pro- and antisocial behavior in adolescence, a period of life when our social world becomes more complex and we hone our skills at navigating increasingly manifold and intimate relationships (Blakemore & Mills, 2014). Two previous experimental studies have shown links between social perspective taking and behavioral measures of social behavior. Specifically, these studies found associations between performance on the Director task and trust and reciprocity towards others in the Trust Game (Fett et al., 2014), and between self-reported perspective-taking skills and age-related increases in non-costly prosocial behavior towards friends (Güroğlu, van den Bos, & Crone, 2014), respectively. However, less is known about how adolescents' social perspective taking, as measured experimentally by the Director task, relates to naturally occurring social behavior, which we aimed to investigate in the current study.

The next aim of the current study was to investigate how individual differences in social perspective taking usage relate to individual difference in brain structure. Prolonged development of use of social perspective taking is consistent with neuroimaging studies of brain development. Both structural (Giedd et al., 2015; Mills, Lalonde, Clasen, Giedd, & Blakemore, 2014; Tamnes et al., 2013) and functional (Blakemore & Robbins, 2012; Braams & Crone, 2017; Flannery, Giuliani, Flournoy, & Pfeifer, 2017; Flournoy et al., 2016; Luna, Padmanabhan, & O'Hearn, 2010; Overgaauw, van Duijvenvoorde, Gunther Moor, & Crone, 2015; Sebastian et al., 2012) magnetic resonance imaging (MRI) studies indicate that brain regions critically involved in social cognition, including dorso-medial prefrontal and lateral temporo-parietal cortices, and/or executive functions, including lateral prefrontal and anterior cingulate cortices, show protracted developmental changes. For example, towards the end of the teenage years, cortical gray matter volume reductions exceeding the average rate are seen primarily in medial prefrontal, lateral prefrontal and lateral temporo-parietal regions (Tamnes et al., 2013). Surface-based cortical reconstruction software also allows for the ability to measure not only cortical volume, but also its separate components thickness and surface area (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). Thickness is defined as the estimated distance between the outer and inner boundary of the cortical sheet and area is defined as the estimated expansion or contraction of points on the surface (Mills & Tamnes, 2014). Although it is believed that, at birth, cortical surface area is largely determined by the number of cortical columns and cortical thickness by the number of cells within a columns (Geschwind & Rakic, 2013), the biological processes that drive later individual and developmental differences are not known. However, longitudinal studies document that these distinct structural properties show unique developmental trajectories across different stages of life (Lyall et al., 2015; Storsve et al., 2014), including across adolescence (Raznahan et al., 2011; Tamnes et al., 2017; Vijayakumar et al., 2016; Wierenga, Langen, Oranje, & Durston, 2014). Although some disagreements exist between available studies regarding the precise development across adolescence, a recent multisite study, which included four independent longitudinal datasets, showed consistent, widespread, and regionally variable nonlinear decreases in cortical thickness and comparably smaller steady decreases in surface area (Tamnes et al., 2017).

An increasing number of studies address the cortical foundations of cognitive development (Jernigan, Brown, Bartsch, & Dale, 2016; Walhovd, Tamnes, & Fjell, 2014). However, studies investigating associations between brain structure and social cognition during childhood and adolescence are scarce, and only a limited number of studies have linked individual differences in brain structure to individual differences in social cognition in adults (Kanai & Rees, 2011). A small number of studies have used social network size (e.g. on Facebook) as a proxy for assessing social cognitive functioning, and have found associations with the size of the amygdala (Von Der Heide, Vyas, & Olson, 2014) and temporal cortex in adults (Kanai, Bahrami, Roylance, & Rees, 2012).

Another study found that anthropomorphic attribution was associated with grey matter volume in the left temporo-parietal junction in adults (Cullen, Kanai, Bahrami, & Rees, 2014). Here, we investigate the relations between individual differences in social perspective taking usage and brain structure in adolescence, to improve our understanding of the sources of variation in social cognition during this period of development.

In the present study, we aimed to: 1) test the reproducibility of the previously reported pattern of age-related improvements in use of social perspective taking across adolescence (Dumontheil, Apperly, et al., 2010; Humphrey & Dumontheil, 2016; Symeonidou et al., 2016); 2) investigate the relationship between individual differences in social perspective taking usage and self-reported real-life social behavior; 3) investigate the relationship between individual differences in social perspective taking usage and structure of the cerebral cortex. We hypothesized that social perspective taking usage would show continued age-related improvement across adolescence. We also predicted that better social perspective taking, independent of age, would be associated with more reported prosocial behavior and fewer reported peer relationship problems, as well as with relatively more mature cortical structure, i.e. lower thickness and possibly lower surface area, in regions involved in mental state attribution and/or executive functions. These predictions were based on the idea that age-independent associations reflect, at least to some extent, individual developmental variability (Jernigan, Baare, Stiles, & Madsen, 2011).

Materials and Methods

Participants

Participants aged 7–26 years were recruited through advertisements and local schools in Oslo, Norway, and originally participated in in one of two longitudinal projects - *Neurocognitive Development* (Tamnes et al., 2010) or the Norwegian Mother and Child Cohort Neurocognitive study (Krogsrud et al., 2014) - or a student research project. Written informed consent was obtained from a parent of all participants under 16 years of age and from participants 12 years of age and older, while participants under 12 years of age gave oral assent. The Regional Committee for Medical and Health Research Ethics Norway approved the study (2009/200: Nevrokognitiv utvikling i skolealder – oppfølgingsstudie).

Participants aged 16 years or older, and parents of participants under 16 years, completed standardized health interviews regarding each participant to ascertain eligibility. All participants were required to be fluent Norwegian speakers, have normal or corrected-to normal vision and hearing, not have any injury or disease known to affect central nervous system function, including diagnosed neurological or psychiatric illness or serious head trauma, and not use psychoactive drugs known to affect central nervous system functioning.

Three hundred and two participants satisfied these criteria. Nine participants were excluded based on behavioral criteria defined in the Director task, as described below. This yielded a sample of 293 participants (164 females) aged 7.1–26.7 years (mean = 16.9, SD = 5.1). The age for females (mean = 16.9 years, SD = 5.3) and males (mean = 16.9 years, SD = 4.8) was not significantly different ($t = 0.03$, $p = .980$). The sample had a mean IQ of 110.1 (SD = 11.2, range = 79–141, missing data for 23 participants) as estimated by a Norwegian version of the Wechsler Abbreviated Scale of Intelligence (WASI) two-subtest form (Wechsler, 1999), including the Vocabulary and Matrix Reasoning subtests. Age groups were created to allow for direct comparison with previous studies reporting on age-related differences in performance on versions of the Director task (Dumontheil, Apperly, et al., 2010; Symeonidou et al., 2016): Children ($n=60$, 7.1–11.3 years, 38 females); Adolescents ($n=108$, 11.7–17.9 years, 53 females); and Adults ($n=125$, 18.0–26.7 years, 73 females).

Of the 293 participants in the full sample, 184 (63%, 101 females) were included in the analyses testing for associations between Director task performance and self-reported behavior, as measured by the Strengths and Difficulties Questionnaire (SDQ) (63 were below 12 years old and were not asked to complete the self-report SDQ, 21 older adolescents participated in a student research project where SDQ was not include in the protocol, and 25 had missing data). The participants in the SDQ sample were 12.2–26.1 years (mean = 18.9, SD = 3.4).

Finally, of the 293 participants in the full sample, 226 (77%, 122 females) were included in the analyses testing for associations with structural properties of the cerebral cortex (40 children were part of a research project that used a different MRI protocol (Krogsrud et al., 2014) and were thus not included, 21 older adolescents participated in a student research project that did not include scanning, and six were excluded during post-processing quality control (QC) as described below). The resulting MRI sample was aged 8.5–26.7 years (mean = 18.3, SD = 4.2).

Experimental task

Social perspective taking usage was assessed by a version of the Director task, originally adapted from Keysar et al. (Keysar et al., 2000; Keysar et al., 2003) by Apperly et al. (2010), and modified for the present study. E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA) was used for stimulus presentation and response logging. All participants first carried out the experimental Director condition before the control No-Director condition of the task.

Standardized instructions were read to the participants before each condition. For the Director condition (**Figure 1**), participants were shown an example stimulus. It was explained that, on each trial, they will be shown a set of shelves containing various objects in different slots and that the man standing on the other side of the shelves (the ‘Director’) will ask the participant to move specific objects to the basket. Emphasis was placed on the fact that the Director had a different perspective to the participant, by explaining and showing that some of the slots are occluded and that the participant, but not the Director, can see the objects in these slots. Participants were shown an illustration of the Director’s view of the same stimulus and it was reiterated that the Director cannot see the objects in the occluded slots and that the participant will have to think about this when performing the task. The task administrator then showed an example of an object that both the participant and the Director could see (“car”), and an example of an object that the participant, but not the Director, could see (“apple”). The participant was then asked to give a different example of an object that only she or he, and not the Director, could see, and an object that both could see. Instructions were repeated if needed. Participants were asked to respond as accurately and quickly as possible by pointing and clicking with the computer mouse and were then given three practice trials.

In Critical trials during the experimental Director condition, participants were required to take account of the Director’s perspective and the correct response was to select the ‘target’ object, which could be seen by the Director, and was the best fit for his instruction if his visual perspective was taken into account. For example in **Figure 2** top left panel when the Director asks “move the small glasses”, the correct response would be to move the glasses with the yellow frame, i.e. the second smallest glasses. If participants ignored the Director’s perspective they would select the ‘distractor’ object, the glasses with the red frame, which were the smallest in the shelves but not visible to the Director. In Control trials, the arrangement of the objects in the shelves was identical to that in the Critical trials, except that an irrelevant object replaced the distractor object (e.g. the truck in **Figure 2** top right panel). In Filler trials, instructions referred only to objects in clear slots, i.e. visible to both the participant and the Director (e.g. “move the tiger”).

Before the start of the control No-Director condition, new instructions were read while participants were shown two examples of stimuli without the Director on the other side of the shelves present. It was explained that some slots in the shelves have gray back panels, while others are clear, and that the participant in each trial will be asked to move specific object to the basket, but that they should only move objects in clear slots. It was stressed that the participant should ignore objects in the slots with a gray background. Each participant was asked to give examples of objects in both types of slots, and was then asked to select an object as they would in a Critical trial to demonstrate that they understood what was required of them. The No-Director trials were identical in every way to the Director trials except that the Director on the other side of the shelves was not present, and instead of having to take into account the Director's perspective, participants were instructed to follow the rule of ignoring all objects in slots with a grey background. Critical, Control and Filler trials were included in the No-Director condition. For example in **Figure 2's** lower left panel when instructed to "move the small ball", the correct response would be to select the second smallest ball, the yellow ball, and ignore the 'distractor' object, the white ball, which was the smallest ball in the shelves, but was in a slot with a gray background. In Control trials, an irrelevant object replaced the distractor object (e.g. the airplane in **Figure 2's** lower right panel). Thus, Critical trials of both the Director and the No-Director condition involved inhibition of a prepotent response of moving the object that best fit the instruction from the participant's perspective, as well as general task demands. The conditions critically differed in whether the participants were instructed to take into account another's perspective or to follow a simple visual rule. Control trials included the requirements of Critical trials to process relative size or position information from an auditory instruction, but did not require participants to take into account the perspective of the Director or inhibit a dominant response. Filler trials served to reduce the saliency of the key trials of interest.

For each participant, the version of the Director and No-Director conditions administered were randomly selected from six alternative versions in order to counterbalance the order of different trial types and stimuli configurations across participants. In both conditions, participants were shown on the computer screen cartoon pictures of a 4 × 4 set of shelves containing eight different objects and five slots with gray backgrounds (occluded slots). Each shelf-object configuration was first presented for 2 seconds, and then three successive auditory instructions were presented, corresponding to two Filler trials and one Control trial or two Filler trials and one Critical trial. Each of these trials lasted for 6 seconds. The instructions were played through the computer speakers and asked the participant to move one of the eight objects, either by only referring to the object name (Filler trials) or by the object name in combination with size (small/large) or relative horizontal position (top/bottom) information (Control and Critical trials). Compared to the developmental study by Dumontheil et al. (2010), where the instructions in the task asked the participants to move specific objects left, right, up or down, our modified version of the task only asked participants to move the object into a basket by clicking it. Thus, the aspect of the task requiring directional decisions, a potential confound, was eliminated. In total, there were 8 Critical trials, 8 Control trials, and 32 Filler trials in each condition (Director and No-Director). Each condition lasted approximately 5.5 minutes.

Behavioral criteria were used to exclude participants with performance indicating that they had not understood the instructions of the task, or had suboptimal motivation or task focus. Specifically, participants with 0% accuracy for any trial type in either of the two conditions were excluded. This led to the exclusion of nine participants, seven of whom were excluded based on performance on critical trials in the experimental Director conditions, and two of which were excluded based on performance on critical trials in the control No-Director condition. All behavioral data reported are based on the remaining 293 participants. Accuracy (percentage errors) and intra-individual median response times (RTs) in correct trials were calculated for each participant in each condition (Director/No-Director) and trial type (Critical/Control/Filler). Additionally, we computed the

difference in percentage errors on Director Critical trials and on No-Director Critical trials, to be used in the analyses testing for associations with self-reported behavior and cortical structure, as described below.

Behavioral questionnaire

The SDQ self-report version was used to assess participants' behavior (R. Goodman, Meltzer, & Bailey, 1998). The SDQ is a well-validated and clinically broadly used questionnaire which asks about 25 attributes, rated on a three-point Likert scale, equally divided between five scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Recent studies indicate that SDQ is not only suitable for distinguishing clinical and healthy groups of children but is also a valid continuous measure of child and adolescent mental health across the full range of variation (A. Goodman & Goodman, 2009). For the current study, we used only the prosocial behavior and peer relationship problems scales.

Image acquisition

MRI acquisition was done with a 3.0T Siemens Skyra (Erlangen, Germany) with a 24-channel coil. Three-dimensional T1-weighted MP-RAGE sequences with the following parameters were used for volumetric and cortical surface analyses: repetition time = 2300 ms; echo time = 2.98 ms; inversion time = 850 ms; flip angle = 8°; bandwidth = 240 Hz/pixel; field of view = 256 mm; and scan time = 9:50 minutes. The sequence consisted of 176 sagittal slices with a voxel size of 1.0 x 1.0 x 1.0 mm.

Image processing

Volumetric segmentation and cortical reconstruction and was performed with the FreeSurfer image analysis suite version 5.3, which is documented and freely available for download online (<http://surfer.nmr.mgh.harvard.edu/>). The technical details of these procedures are described in prior publications (Dale et al., 1999; Fischl, 2012; Fischl et al., 2002; Fischl et al., 1999). Briefly, the processing includes motion correction, removal of non-brain tissue, automated Talairach transformation, segmentation of subcortical volumetric structures, intensity normalization, tessellation of surfaces, automated topology correction, and surface deformation to optimally place tissue borders. Cortical thickness maps for each subject were obtained by calculating the distance between the cortical gray matter and white matter surface at each vertex (surface point) (Fischl & Dale, 2000). Cortical surface area (white matter surface) maps were computed for each subject by calculating the area of every triangle in the tessellation. The triangular area at each location in native space was compared with the area of the analogous location in registered space to give an estimate of expansion or contraction continuously along the surface ("local arealization") (Fischl et al., 1999). The maps produced are not restricted to the voxel resolution of the original data and are thus capable of detecting submillimeter differences. In addition to screening of all images immediately after data acquisition and rescanning if needed and possible, all processed scans were visually inspected in detail as part of the QC procedure. Before statistical analyses, surface maps for cortical thickness and area were smoothed with a Gaussian kernel of full-width at half maximum of 15 mm.

Statistical analysis

For the full sample (n=293, 7.1–26.7 years), participants on average made only 1.6% and 1.4% errors in Filler trials in the Director and the No-Director condition, respectively, and the data from these trials were not included in further analyses. For both accuracy and RT, a 2 × 2 × 3 mixed ANOVA with Condition (Director, No-Director) and Trial type (Critical, Control) as within-subject factors and Age group (Children, Adolescents, Adults) as between-subject factors was performed. ANOVAs on separate trial types, conditions or age groups and independent samples t-tests between age groups were performed as follow-up analyses to further investigate significant interaction and main effects.

We then tested for associations between performance on the Director task and both self-reported behavior and structure of the cerebral cortex. Our task performance measure of interest for these analyses was the difference in percentage errors on Director Critical trials and on No-Director Critical trials. This measure was chosen in order to identify individual differences in social perspective taking, while controlling for some general and executive function task demands. First, for a subsample of adolescents and young adults ($n=184$, 12.2–26.1 years), we used general linear models (GLMs) in SPSS with each of the two SDQ scales of interest (prosocial behavior, peer relationship problems) as the dependent variable, sex as a fixed factor, and age and task performance as covariates.

Second, for the MRI sample (226 participants, 8.5–26.7 years old), we performed surface-based cortical analyses on a vertex-wise (point-by-point) level using GLMs as implemented in FreeSurfer. Effects of task performance on both cortical thickness and surface area were tested. Initially, this was done while controlling only for sex to test for temporal co-occurrence of overall developmental trends in behavior and cortical structure. Such associations do however not necessarily imply that the variables are directly interrelated (Salthouse, 2011). The analyses were therefore then repeated while additionally controlling for age, as it is reasonable to hypothesize that such age-independent associations are mediated, at least to some extent, by developmental variability, i.e. variability among adolescents of similar age in the phase of brain maturation (Jernigan et al., 2011). The data were tested against an empirical null distribution of maximum cluster size across 10,000 iterations using Z Monte Carlo simulations as implemented in FreeSurfer (Hagler, Saygin, & Sereno, 2006; Hayasaka & Nichols, 2003) synthesized with a cluster-forming threshold of $p < .05$, yielding clusters fully corrected for multiple comparisons across the surfaces. Clusterwise corrected $p < .05$ was regarded as significant. Mean cortical thickness was then extracted from each significant cluster and we performed GLMs in SPSS with thickness as dependent variable, sex as fixed factor, and age and task performance as covariates in order to obtain effect size estimates. Note however that these are inflated because they are based on already-identified significant clusters. To test the indirect effect of age on Director task performance through cortical thickness, Hayes' PROCESS tool was used (v2.16.3; mediation model number 4; 10,000 bootstrap resamples) (Hayes, 2013). An indirect path is considered statistically significant if the associated 95% confidence interval (CI) does not include zero.

Results

Director task: Accuracy

A summary of task performance for the full sample is presented in **Table 1**. Average percentage errors for Critical trials and Control trials in the Director condition and the No-Director condition, respectively, for each of the three age groups are shown in **Figure 3**. All main effects were significant in a 2 (Condition: Director, No-Director) \times 2 (Trial type: Critical, Control) \times 3 (Age group: Children, Adolescents, Adults) mixed ANOVA on accuracy. Participants made more errors in the Director condition than in the No-Director condition ($F(1,290) = 69.63$, $p < .001$, $\eta_p^2 = .194$), more errors on Critical trials than on Control trials ($F(1,290) = 95.76$, $p < .001$, $\eta_p^2 = .248$), and accuracy differed with Age group ($F(2,290) = 42.90$, $p < .001$, $\eta_p^2 = .228$). There was a significant interaction between Condition and Trial type ($F(1,290) = 35.46$, $p < .001$, $\eta_p^2 = .109$), between Condition and Age group ($F(2,290) = 6.88$, $p = .001$, $\eta_p^2 = .045$), and between Trial type and Age group ($F(2,290) = 18.95$, $p < .001$, $\eta_p^2 = .116$). The three-way interaction was also significant ($F(2,290) = 5.91$, $p = .003$, $\eta_p^2 = .039$), and was explored further by looking at Critical and Control trials separately.

A 2 \times 3 mixed ANOVA performed on the Critical trials showed main effects of Condition ($F(1,290) = 58.34$, $p < .001$, $\eta_p^2 = .167$), with more errors in the Director condition, and Age group ($F(2,290) = 33.03$, $p < .001$, $\eta_p^2 = .186$), as well as a significant interaction between Condition and Age group ($F(2,290) = 7.14$, $p < .001$, $\eta_p^2 = .047$). The same analysis on the Control trials only showed significant

main effects of Condition ($F(1,290) = 11.48, p < .001, \eta_p^2 = .038$), again with more errors in the Director condition, and Age group ($F(2,290) = 21.71, p < .001, \eta_p^2 = .130$), but no significant interaction effect ($F(2,290) = 1.13, p = .323, \eta_p^2 = .008$).

Follow-up analyses on the Critical trials in the two conditions separately showed a significant effect of Age group on accuracy in both the Director condition ($F(2,290) = 20.88, p < .001, \eta_p^2 = .126$) and the No-Director condition ($F(2,290) = 36.42, p < .001, \eta_p^2 = .201$). Independent samples t-tests for the Critical trials in the Director condition revealed that the child group made significantly more errors than both the adolescent group ($t = 3.45, p < .001, d = 0.66$) and the adult group ($t = 5.75, p < .001, d = 1.31$), and also that the adolescent group made significantly more errors than the adult group ($t = 2.73, p = .007, d = 0.41$). While for the Critical trials in the No-Director condition, the child group made significantly more errors than both the adolescent group ($t = 5.99, p < .001, d = 1.45$) and the adult group ($t = 5.54, p < .001, d = 1.24$), but there was no difference between the adolescent and the adult group ($t = -0.41, p = .683$). Additional analyses with age as a continuous variable showed very similar results (see Supplementary Material).

Director task: Response times

Average median RTs for correct Critical trials and Control trials in the Director condition and the No-Director condition, respectively, for each of the three age groups are shown in **Figure 4**. All main effects were significant in a 2 (Condition: Director, No-Director) \times 2 (Trial type: Critical, Control) \times 3 (Age group: Child, Adolescent, Adult) mixed ANOVA on RT. Participants were slower overall in the Director condition than in the No-Director condition ($F(1,290) = 9.15, p = .003, \eta_p^2 = .031$), and on Control trials than on Critical trials ($F(1,290) = 64.42, p < .001, \eta_p^2 = .182$), and RTs differed with Age group ($F(2,290) = 88.41, p < .001, \eta_p^2 = .379$). There was a significant interaction between Condition and Age group ($F(2,290) = 6.27, p = .002, \eta_p^2 = .041$), but no significant interactions between Trial type and Age group ($F(2,290) = 1.17, p = .312, \eta_p^2 = .008$), or Condition and Trial type ($F(1,290) = 2.82, p = .094, \eta_p^2 = .010$). The three-way interaction was not significant ($F(2,290) = 0.35, p = .708, \eta_p^2 = .002$).

Due to the interaction between Condition and Age group, the main effect of Condition on RTs was explored further in each Age group separately. In the follow-up 2 \times 2 mixed ANOVAs, the main effect of Condition was not significant in the child group ($F(1,70) = 0.82, p = .368, \eta_p^2 = .014$), but was significant in both the adolescent group ($F(1,96) = 10.23, p = .002, \eta_p^2 = .087$) and the adult group ($F(1,124) = 25.35, p < .001, \eta_p^2 = .170$). In both cases, participants were slower in the Director condition than in the No-Director condition (see Supplementary Material for analyses with age as a continuous variable).

Associations between task performance and self-reported behavior

Relationships between Director task performance and self-reported behavior on the SDQ were investigated with GLMs, with each of the two SDQ scales of interest (prosocial behavior, peer relationship problems) as dependent variable, sex as fixed factor, and age and the difference in percentage errors on Director Critical trials and on No-Director Critical trials as covariates. There was a significant small negative association between errors and prosocial behavior ($F = 4.42, p = .037, \eta_p^2 = .024$), such that participants who performed better on the Director task reported to show more prosocial behavior. In contrast, although there was a positive trend, the association between task errors and reported peer relationship problems was not significant ($F = 2.83, p = .094, \eta_p^2 = .015$).

Associations between task performance and structure of the cerebral cortex

Relationships between Director task performance and cerebral cortex structure were initially explored across the cortical surface with GLMs testing the effects of the difference in percentage errors on Director Critical trials and on No-Director Critical trials on both cortical thickness and

surface area, while controlling only for sex. After correction for multiple comparisons using cluster size inference, extensive bilateral fronto-parietal regions, including superior, lateral and medial prefrontal and lateral parietal cortices, as well as lateral temporal lobe regions in the left hemisphere showed positive associations between errors and cortical thickness (**Figure 5**). There were no significant effects in the other direction or on cortical surface area. We then repeated the analysis while additionally controlling for age, and these age-independent results showed two lateral regions in the left hemisphere with positive associations between errors and cortical thickness (**Figure 6**): A frontal cluster, which included parts of the caudal middle frontal and precentral gyri (1,454 mm², clusterwise $p = .038$, CI = .036 - .040), and a parietal cluster encompassing parts of the superior and inferior parietal lobules and the postcentral sulcus (1,997 mm², clusterwise $p = .006$, CI = .005 - .007). These positive associations indicate that better performance was related to thinner cortices in these regions, independently of sex and age. Again, there were no significant effects in the other direction or on cortical surface area. To assess the size of the age-independent effects, we performed GLMs with mean cortical thickness in each of the two significant clusters as dependent variable, sex as fixed factor, and age and errors specifically on Director Critical trials as covariates. The results showed a small effect size for task performance in the frontal cluster ($F = 4.21$, $p = .042$, $\eta_p^2 = .023$) and a somewhat larger effect in the parietal cluster ($F = 8.60$, $p = .004$, $\eta_p^2 = .046$).

Finally, the indirect effect of age on the difference in percentage errors between Director Critical trials and No-Director Critical trials via cortical thickness in the two identified clusters was tested in two mediation analyses using Hayes' bootstrapping method. The analyses revealed significant indirect effects of age on Director task performance via cortical thickness in the frontal cluster (indirect effect = -0.26, SE = 0.12, CIs = -0.552 to -0.089), and via cortical thickness in the parietal cluster (indirect effect = -0.35, SE = 0.13, CIs = -0.665 to -0.154), whereby older age was associated with lower cortical thickness in the two clusters, which was in turn associated with better Director task perspective taking performance.

Discussion

In this study, we tested for age-related differences in the ability to use information about another person's perspective when following instructions, and investigated whether this experimental measure of social perspective taking was related with self-reported real-life social behavior and with MRI-derived measures of the structure of the cerebral cortex. The behavioral results support previous findings of continued development of the use of social perspective taking across adolescence (Dumontheil, Apperly, et al., 2010). Further, independent of age, participants who performed better specifically on trials requiring social perspective taking reported more prosocial behavior and had thinner cerebral cortex in regions in the left hemisphere encompassing parts of the middle frontal gyrus and the lateral parietal lobe. There were also indirect effects of age on social perspective taking usage through cortical thickness in these regions.

We included a large cross-sectional sample ($n=293$) of participants ranging in age from 7 to 26 years and a slightly modified version of the computerized Director task to test the reproducibility of a previously reported pattern of age-related improvements in social cognition across adolescence (Dumontheil, Apperly, et al., 2010; Symeonidou et al., 2016). Accurate performance in the experimental condition of this task is thought to depend upon use of the ability to represent what another person can see, which is a core component of theory of mind (Dumontheil, Apperly, et al., 2010; Flavell, Everett, Croft, & Flavell, 1981). The results showed that for trials specifically requiring participants to take into account the Director's perspective in order to identify and select the instructed target objects in cartoon pictures of a set of shelves containing multiple different objects, children made more errors than adolescents, and adolescents made more errors than adults. In comparison, for trials that required participants to follow a simple visual rule, but which were

otherwise identical to the perspective taking trials, children made more errors than other group, but the accuracy of adolescents and adults did not differ.

In the original study on age-related differences in performance on the Director task, Dumontheil et al. (2010) reported on data from 177 female participants 7-27 years old and similarly found that, on critical trials in the experimental condition, but not in the control condition, both children and adolescents made more errors than adults. This main finding has been replicated in two smaller studies including both male and female participants: one with 65 participants aged 9-29 years (Symeonidou et al., 2016) and one with 90 participants aged 11-18 years (Humphrey & Dumontheil, 2016). Our results also replicate this main finding and thus support the conclusion that developmental changes in use of social perspective taking are still occurring across adolescence. A caveat is that the Director in our task was an older man. If social perspective taking usage is contingent on participants' relationship with the target, it is possible that younger people are less inclined to take the perspective of this older individual. However, a previous study used a Director task version with a younger Director (Symeonidou et al., 2016) and found similar developmental differences as the original study (Dumontheil, Apperly, et al., 2010).

Our results for response times showed that both adolescents and adults were slower in the Director condition than in the No-Director condition, and counterintuitively that participants overall also were slower on Control trials than on Critical trials. Response time is not a key measure of interest in the Director task (Humphrey & Dumontheil, 2016), and previous studies have reported conflicting findings. Inconsistent with the current results, Dumontheil et al. (2010) found slower responses in the No-Director condition than in the Director condition, but consistent with the current results, slower responses in Control trials than in Critical trials. In contrast, Humphrey and Dumontheil (2016) found no significant effects of condition or trial type, while Symeonidou et al. (2016) found a three-way interaction between condition, trial type and age.

There were some notable differences between the paradigm used in our study and that used in previous studies. Compared to previous studies, we used a modified task with simplified instructions, which did not require participants to make directional decisions. Possibly as a function of this, as the instructions were simpler, the error rates were on average much lower in our study than in previous studies (Dumontheil, Apperly, et al., 2010; Humphrey & Dumontheil, 2016; Symeonidou et al., 2016). Despite this difference, the overall pattern of age-related effects on accuracy was the same in our study as in previous studies. As the use of social perspective taking is a key component of theory of mind, these findings are also consistent with studies indicating ongoing development of mentalizing about both emotions and actions throughout adolescence (Keulers, Evers, Stiers, & Jolles, 2010; Sebastian et al., 2012; Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013). To understand the underlying factors in Director task performance, Symeonidou et al. (2016) analyzed eye-tracking data acquired during correct and incorrect trials separately and found that children, adolescents and young adults did not significantly differ in their online processing during perspective taking. This might suggest that the age-related differences in behavior are in the likelihood, rather than the nature, of perspective taking, i.e. they are quantitative rather than qualitative. Other studies have investigated the possibility that inhibitory control, by allowing participants to inhibit their own perspective in favor of another individual's perspective, may underlie developmental changes in social perspective taking. These studies have found, both in adults and in developmental samples, a positive correlations between inhibition and perspective taking (Nilsen & Graham, 2009). For instance, one study found that inhibitory control, as measured by Go/Nogo task performance, partly accounted for Director task accuracy in adolescents (Symeonidou et al., 2016), although this finding was not replicated in a later study with a smaller age range (Humphrey & Dumontheil, 2016).

The main objective of the present study was to connect multiple levels of analysis by relating the experimental measure of social perspective taking obtained through the Director task to real life social behavior and to individual differences in brain structure during adolescence. For these analyses, we used the difference in percentage errors on critical trials in the Director condition and the control condition, respectively, as our measure of interest to specifically focus on social perspective taking, while controlling for general and executive function task demands.

First, supporting our hypothesis, we found that, independent of age and thus possibly indicative of developmental variability, there was a negative association between errors specifically on trials requiring use of social perspective taking and self-reported prosocial behavior. The strength of this relationship was small, but it fits with numerous studies documenting a small positive association between children's theory of mind scores and various measures of prosocial behavior (Imuta et al., 2016). We had also hypothesized that better use of social perspective taking would be associated with fewer reported peer relationships problems, and although there was such a trend, this association was not significant. Future studies with more in-depth assessment of social behavior, including reports from multiple informants or observational data for example, should investigate this further.

Few previous studies have focused on the pro- and antisocial behavioral relevance of perspective taking ability in adolescents. One notable exception is an experimental study by Fett et al. (2014), which found that individual differences in adolescents' social perspective taking, as measured with the Director task, was associated with social behavior, specifically behavioral measures of initial trust and reciprocity in the Trust Game. Another exception is a study by Güroğlu et al. (2014), which found that older adolescents compared to younger adolescents showed increased differentiation in prosocial behavior depending on the relation with the interacting partner in the task (friend, antagonist, neutral classmate, or anonymous peer). Furthermore, the age-related increase in non-costly prosocial behavior towards friends was mediated by self-reported perspective taking skills (Güroğlu et al., 2014). The current study adds to this literature, by showing that adolescents' use of social perspective taking to guide decisions and behavior is associated with more reported prosocial behavior.

Second, the current study provided novel results regarding the brain structural correlates of the use of social perspective taking. Specifically, we studied both cortical thickness and surface area, as these separate components of cortical structure have heterogeneous phylogenetic development and ontogenetic origins (Geschwind & Rakic, 2013) and distinct genetic influences and patterning (Chen et al., 2013), and critically, develop differently from childhood to adulthood (Raznahan et al., 2011; Tamnes et al., 2017; Vijayakumar et al., 2016; Wierenga et al., 2014). The results showed that, when age was not statistically controlled for, better performance specifically on trials requiring use of social perspective taking was associated with thinner cortex in widespread bilateral fronto-parietal and left hemisphere lateral temporal regions. Interestingly, among the regions showing the strongest associations were the medial prefrontal, lateral prefrontal and anterior cingulate cortices; brain regions that have been implicated in social cognition (Apps, Rushworth, & Chang, 2016; Kilford, Garrett, & Blakemore, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle, 2009) and/or executive functioning (Crone & Steinbeis, 2017; Paus, 2001; Yuan & Raz, 2014). Our results indicate temporal co-occurrence of developmental trends in use of social perspective taking and cortical thickness in these widespread regions, but these results are not sufficient evidence for directly linking the two. We therefore repeated the analyses while additionally controlling for age (together with sex), as such age-independent associations might partly be mediated by individual developmental variability.

Independent of age and sex, better performance specifically on trials requiring use of social perspective taking was associated with thinner cortex in the left hemisphere in parts of the caudal middle frontal and precentral gyri, and in a lateral parietal region covering parts of the superior and inferior parietal lobules and the postcentral gyrus. As the cerebral cortex generally, as well as in these regions specifically, decreases with age across adolescence (Tamnes et al., 2017), this might indicate that individuals with better ability to use social perspective taking have relatively more mature cortical structure in these regions. It should be noted that these results showed small effect sizes. However, age-independent relationships between behavioral and brain measures are typically moderate, likely related to the fact that there is much individual variance at both levels at any given age, and that the relationships may also fluctuate with age (Walhovd et al., 2014). Moreover, a central tenet is that the shape of brain developmental trajectories may be more strongly related to behavioral and functional characteristics than absolute brain measures at any given point during development (Giedd & Rapoport, 2010), and longitudinal studies should therefore be performed. Mediation analyses in the present cross-sectional sample did reveal indirect effects of age on social perspective taking usage via cortical thickness in these frontal and parietal regions. This supports the purported relevance of cortical development for development of social perspective taking.

While, as far as we know, the present study is the first to investigate the brain structural correlates of social perspective taking, results from fMRI studies of adults performing a version of the Director task have shown that using social perspective taking is associated with linked activation of lateral temporal cortices, and medial and lateral prefrontal regions, i.e. regions typically involved in both social cognition and executive functions (Dumontheil, Kuster, Apperly, & Blakemore, 2010; Hillebrandt, Dumontheil, Blakemore, & Roiser, 2013). Results from another fMRI study suggest that adolescents show greater activation in dorsal medial prefrontal cortex whenever social information is present, whereas adults only show such increased activation when the social information is relevant to task performance, and this might indicate a lesser functional specificity of this brain region in adolescence (Dumontheil, Hillebrandt, Apperly, & Blakemore, 2012). Complementing these findings, the current study link social perspective taking ability and brain structure. In contrast to the associations found between Director task performance and cortical thickness, no associations were found with cortical surface area. This might possibly relate to that this dimension of cortical structure changes less than cortical thickness across adolescence (Jernigan et al., 2016; Tamnes et al., 2017).

The current findings should be interpreted in light of the following issues. First, the results were obtained using cross-sectional data. The development of social perspective taking usage and its links with social behavior and brain structure should be further investigated with longitudinal data. Second, an important question is whether errors in the experimental condition of the Director task actually reflect failure to use social perspective taking, which involves some understanding of another person's preferences, goals, intentions etc., or selective attention (Rubio-Fernandez, 2017) or visuospatial manipulation failure (Fett et al., 2014). Studies of adults indicate that errors on this type of task do not arise simply as a result of failure to effectively switch perspectives (Apperly et al., 2010), but further studies on developmental samples comparing visuospatial processing abilities and performance on the Director task are called for. Third, and related to the previous issue, our results showed relationships between use of social perspective taking and cortical thickness in regions including the superior parietal lobule and the caudal middle frontal gyrus, regions known to show increased activity associated with visual perspective taking (Schurz, Aichhorn, Martin, & Perner, 2013) and mental rotation (Tomasino & Gremese, 2016). This also begs the question as to what degree the employed task really requires social perspective taking. However, a growing body of work links the visuospatial and the social aspects of perspective taking (Hamilton, Kessler, & Creem-Regehr, 2014) and it can thus be argued against a simple distinction between the two. Nonetheless, future neuroimaging studies are needed to investigate to what degree the two are dissociable in

terms of brain structure and function. Finally, the assessment of prosocial behavior and peer relationship problems was limited to brief self-report questionnaire scales.

There has recently been a call for more large-scale studies on individual differences in neurocognitive development (Foulkes & Blakemore, in press). The results of the current study, which used an experimental approach and a large cross-sectional sample of participants aged 7- 26 years, replicate the findings of earlier studies indicating continued development of use of social perspective taking across adolescence. Furthermore, within subsamples, the study yielded novel results linking individual differences in use of social perspective taking with a higher level of real-life prosocial behavior and with thinner and possibly more mature cerebral cortex in fronto-parietal regions.

References

- Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology*, *63*, 1201-1217. doi:10.1080/17470210903281582
- Apps, M. A., Rushworth, M. F., & Chang, S. W. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron*, *90*(4), 692-707. doi:10.1016/j.neuron.2016.04.018
- Austin, G., Bondu, R., & Elsner, B. (2017). Longitudinal relations between children's cognitive and affective theory of mind with reactive and proactive aggression. *Aggressive Behavior*, *43*, 440-449. doi:10.1002/ab.21702
- Begeer, S., Bernstein, D. M., Assfalg, A., Azdad, H., Glasbergen, T., Wierda, M., & Koot, H. M. (2016). Equal egocentric bias in school-aged children with and without autism spectrum disorders. *Journal of Experimental Child Psychology*, *144*, 15-26. doi:10.1016/j.jecp.2015.10.018
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187-207. doi:10.1146/annurev-psych-010213-115202
- Blakemore, S. J., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, *15*, 1184-1191. doi:10.1038/nn.3177
- Braams, B. R., & Crone, E. A. (2017). Longitudinal changes in social brain development: Processing outcomes for friend and self. *Child Development*, *88*, 1952-1965. doi:10.1111/cdev.12665
- Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: the role of prosocial behavior. *Developmental Psychology*, *48*(1), 257-270. doi:10.1037/a0025402
- Chen, C. H., Fiecas, M., Gutierrez, E. D., Panizzon, M. S., Eyler, L. T., Vuoksimaa, E., . . . Kremen, W. S. (2013). Genetic topography of brain morphology. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 17089-17094. doi:10.1073/pnas.1308091110
- Crone, E. A., & Steinbeis, N. (2017). Neural perspectives on cognitive control development during childhood and adolescence. *Trends in Cognitive Science*, *21*(3), 205-215. doi:10.1016/j.tics.2017.01.003
- Cullen, H., Kanai, R., Bahrami, B., & Rees, G. (2014). Individual differences in anthropomorphic attributions and human brain structure. *Social Cognitive and Affective Neuroscience*, *9*(9), 1276-1280. doi:10.1093/scan/nst109
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, *9*, 179-194. doi:10.1006/nimg.1998.0395
- Decety, J., Barta, I. B., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B Biological Sciences*, *371*, 20150077. doi:10.1098/rstb.2015.0077

- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*, 331-338. doi:10.1111/j.1467-7687.2009.00888.x
- Dumontheil, I., Hillebrandt, H., Apperly, I. A., & Blakemore, S. J. (2012). Developmental differences in the control of action selection by social information. *Journal of Cognitive Neuroscience*, *24*, 2080-2095. doi:10.1162/jocn_a_00268
- Dumontheil, I., Kuster, O., Apperly, I. A., & Blakemore, S. J. (2010). Taking perspective into account in a communicative task. *Neuroimage*, *52*, 1574-1583. doi:10.1016/j.neuroimage.2010.05.056
- Fett, A. K., Shergill, S. S., Gromann, P. M., Dumontheil, I., Blakemore, S. J., Yakub, F., & Krabbendam, L. (2014). Trust and social reciprocity in adolescence--a matter of perspective-taking. *Journal of Adolescence*, *37*, 175-184. doi:10.1016/j.adolescence.2013.11.011
- Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*, 774-781. doi:10.1016/j.neuroimage.2012.01.021
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11050-11055. doi:10.1073/pnas.200033797
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., . . . Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341-355. doi:[https://doi.org/10.1016/S0896-6273\(02\)00569-X](https://doi.org/10.1016/S0896-6273(02)00569-X)
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, *9*, 195-207. doi:10.1006/nimg.1998.0396
- Flannery, J. E., Giuliani, N. R., Flournoy, J. C., & Pfeifer, J. H. (2017). Neurodevelopmental changes across adolescence in viewing and labeling dynamic peer emotions. *Developmental Cognitive Neuroscience*, *25*, 113-127. doi:10.1016/j.dcn.2017.02.003
- Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the level 1 - level 2 distinction. *Developmental Psychology*, *17*, 99-103. doi:DOI: 10.1037/0012-1649.17.1.99
- Flournoy, J. C., Pfeifer, J. H., Moore, W. E., Tackman, A. M., Masten, C. L., Mazziotta, J. C., . . . Dapretto, M. (2016). Neural reactivity to emotional faces may mediate the relationship between childhood empathy and adolescent prosocial behavior. *Child Development*, *87*(6), 1691-1702. doi:10.1111/cdev.12630
- Foulkes, L., & Blakemore, S. J. (in press). Studying individual differences in human adolescent brain development. *Nature Neuroscience*. doi:10.1038/s41593-018-0078-4
- Geschwind, D. H., & Rakic, P. (2013). Cortical evolution: judge the brain by its cover. *Neuron*, *80*(3), 633-647. doi:10.1016/j.neuron.2013.10.045
- Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron*, *67*, 728-734. doi:10.1016/j.neuron.2010.08.040
- Giedd, J. N., Raznahan, A., Alexander-Bloch, A., Schmitt, E., Gogtay, N., & Rapoport, J. L. (2015). Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*, *40*(1), 43-49. doi:10.1038/npp.2014.236
- Goodman, A., & Goodman, R. (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*, 400-403. doi:10.1097/CHI.0b013e3181985068
- Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry*, *7*, 125-130. doi:DOI: 10.1007/s007870050057
- Güroğlu, B., van den Bos, W., & Crone, E. A. (2014). Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior. *Frontiers in Psychology*, *5*, 291. doi:10.3389/fpsyg.2014.00291

- Hagler, D. J., Jr., Saygin, A. P., & Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *Neuroimage*, *33*, 1093-1103. doi:10.1016/j.neuroimage.2006.07.036
- Hamilton, A. F., Kessler, K., & Creem-Regehr, S. H. (2014). Perspective taking: building a neurocognitive framework for integrating the "social" and the "spatial". *Frontiers in Human Neuroscience*, *8*, 403. doi:10.3389/fnhum.2014.00403
- Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: random field and permutation methods. *Neuroimage*, *20*, 2343-2356. doi:<https://doi.org/10.1016/j.neuroimage.2003.08.003>
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach*. New York: Guilford Press.
- Hillebrandt, H., Dumontheil, I., Blakemore, S. J., & Roiser, J. P. (2013). Dynamic causal modelling of effective connectivity during perspective taking in a communicative task. *Neuroimage*, *76*, 116-124. doi:10.1016/j.neuroimage.2013.02.072
- Humphrey, G., & Dumontheil, I. (2016). Development of risk-taking, perspective-taking, and inhibitory control during adolescence. *Developmental Neuropsychology*, *41*(1-2), 59-76. doi:10.1080/87565641.2016.1161764
- Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, *52*(8), 1192-1205. doi:10.1037/dev0000140
- Jernigan, T. L., Baare, W. F., Stiles, J., & Madsen, K. S. (2011). Postnatal brain development: structural imaging of dynamic neurodevelopmental processes. *Progress in Brain Research*, *189*, 77-92. doi:10.1016/B978-0-444-53884-0.00019-1
- Jernigan, T. L., Brown, T. T., Bartsch, H., & Dale, A. M. (2016). Toward an integrative science of the developing human mind and brain: Focus on the developing cortex. *Developmental Cognitive Neuroscience*, *18*, 2-11. doi:10.1016/j.dcn.2015.07.008
- Kanai, R., Bahrami, B., Roylance, R., & Rees, G. (2012). Online social network size is reflected in human brain structure. *Proceedings of the Royal Society B Biological Sciences*, *279*(1732), 1327-1334. doi:10.1098/rspb.2011.1959
- Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*(4), 231-242. doi:10.1038/nrn3000
- Keulers, E. H., Evers, E. A., Stiers, P., & Jolles, J. (2010). Age, sex, and pubertal phase influence mentalizing about emotions and actions in adolescents. *Developmental Neuropsychology*, *35*(5), 555-569. doi:10.1080/87565641.2010.494920
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychological Science*, *11*, 32-38. doi:10.1111/1467-9280.00211
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25-41. doi:[https://doi.org/10.1016/S0010-0277\(03\)00064-7](https://doi.org/10.1016/S0010-0277(03)00064-7)
- Kilford, E. J., Garrett, E., & Blakemore, S. J. (2016). The development of social cognition in adolescence: An integrated perspective. *Neuroscience & Biobehavioral Reviews*, *70*, 106-120. doi:10.1016/j.neubiorev.2016.08.016
- Krogsrud, S. K., Tamnes, C. K., Fjell, A. M., Amlien, I., Grydeland, H., Sulutvedt, U., . . . Walhovd, K. B. (2014). Development of hippocampal subfield volumes from 4 to 22 years. *Human Brain Mapping*, *35*(11), 5646-5657. doi:10.1002/hbm.22576
- Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and Cognition*, *72*, 101-113. doi:10.1016/j.bandc.2009.08.005
- Lyall, A. E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., . . . Gilmore, J. H. (2015). Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. *Cerebral Cortex*, *25*(8), 2204-2212. doi:10.1093/cercor/bhu027

- Mills, K. L., Dumontheil, I., Speekenbrink, M., & Blakemore, S. J. (2015). Multitasking during social interactions in adolescence and early adulthood. *Royal Society Open Science*, 2(11), 150117. doi:10.1098/rsos.150117
- Mills, K. L., Lalonde, F., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). Developmental changes in the structure of the social brain in late childhood and adolescence. *Social Cognitive and Affective Neuroscience*, 9, 123-131. doi:10.1093/scan/nss113
- Mills, K. L., & Tamnes, C. K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience*, 9, 172-190. doi:10.1016/j.dcn.2014.04.004
- Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, 24, 603-613. doi:DOI: 10.1348/026151005X55370
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, 58, 220-249. doi:10.1016/j.cogpsych.2008.07.002
- Overgaauw, S., van Duijvenvoorde, A. C., Gunther Moor, B., & Crone, E. A. (2015). A longitudinal analysis of neural regions involved in reading the mind in the eyes. *Social Cognitive and Affective Neuroscience*, 10, 619-627. doi:10.1093/scan/nsu095
- Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, 2(6), 417-424. doi:10.1038/35077500
- Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." Attributions of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437-471. doi:[https://doi.org/10.1016/0022-0965\(85\)90051-7](https://doi.org/10.1016/0022-0965(85)90051-7)
- Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., . . . Giedd, J. N. (2011). How does your cortex grow? *Journal of Neuroscience*, 31, 7174-7177. doi:10.1523/JNEUROSCI.0054-11.2011
- Rubio-Fernandez, P. (2017). The director task: A test of theory-of-mind use or selective attention? *Psychonomic Bulletin & Review*, 24, 1121-1128. doi:10.3758/s13423-016-1190-7
- Salthouse, T. A. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin*, 137(5), 753-784. doi:10.1037/a0023262
- Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, 7, 712. doi:10.3389/fnhum.2013.00712
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, 42, 9-34. doi:10.1016/j.neubiorev.2014.01.009
- Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S. J., Brito, S. A., McCrory, E. J., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, 7, 53-63. doi:10.1093/scan/nsr023
- Sierksma, J., Thijs, J., Verkuyten, M., & Komter, A. (2014). Children's reasoning about the refusal to help: the role of need, costs, and social perspective taking. *Child Development*, 85(3), 1134-1149. doi:DOI: 10.1111/cdev.12195
- Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, 10, 199-204. doi:10.1111/j.1467-7687.2007.00580.x
- Storsve, A. B., Fjell, A. M., Tamnes, C. K., Westlye, L. T., Øverbye, K., Aasland, H. W., & Walhovd, K. B. (2014). Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: regions of accelerating and decelerating change. *Journal of Neuroscience*, 34(25), 8488-8498. doi:10.1523/JNEUROSCI.0391-14.2014

- Symeonidou, I., Dumontheil, I., Chow, W. Y., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology, 149*, 81-97. doi:10.1016/j.jecp.2015.11.007
- Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., . . . Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *Journal of Neuroscience, 37*(12), 3402-3412. doi:10.1523/JNEUROSCI.3302-16.2017
- Tamnes, C. K., Walhovd, K. B., Dale, A. M., Østby, Y., Grydeland, H., Richardson, G., . . . Alzheimer's Disease Neuroimaging Initiative. (2013). Brain development and aging: overlapping and unique patterns of change. *Neuroimage, 68*, 63-74. doi:10.1016/j.neuroimage.2012.11.039
- Tamnes, C. K., Østby, Y., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cerebral Cortex, 20*(3), 534-548. doi:10.1093/cercor/bhp118
- Tomasino, B., & Gremese, M. (2016). Effects of stimulus type and strategy on mental rotation network: An activation likelihood estimation meta-analysis. *Frontiers in Human Neuroscience, 9*, 693. doi:10.3389/fnhum.2015.00693
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping, 30*, 829-858. doi:10.1002/hbm.20547
- Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology, 38*, 114-125. doi:10.1080/87565641.2012.733786
- Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology, 19*, 615-629. doi:10.1080/09297049.2012.718324
- Vijayakumar, N., Allen, N. B., Youssef, G., Dennison, M., Yucel, M., Simmons, J. G., & Whittle, S. (2016). Brain development during adolescence: A mixed-longitudinal investigation of cortical thickness, surface area, and volume. *Human Brain Mapping, 37*(6), 2027-2038. doi:10.1002/hbm.23154
- Von Der Heide, R., Vyas, G., & Olson, I. R. (2014). The social network-network: size is predicted by brain structure and function in the amygdala and paralimbic regions. *Social Cognitive and Affective Neuroscience, 9*(12), 1962-1972. doi:10.1093/scan/nsu009
- Walhovd, K. B., Tamnes, C. K., & Fjell, A. M. (2014). Brain structural maturation and the foundations of cognitive behavioral development. *Current Opinion in Neurology, 27*(2), 176-184. doi:10.1097/WCO.0000000000000074
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: The Psychological Corporation.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development, 72*, 655-684. doi:DOI: 10.1111/1467-8624.00304
- Wierenga, L. M., Langen, M., Oranje, B., & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *Neuroimage, 87*, 120-126. doi:10.1016/j.neuroimage.2013.11.010
- Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: a meta-analysis of structural neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 180-192. doi:10.1016/j.neubiorev.2014.02.005

Figures

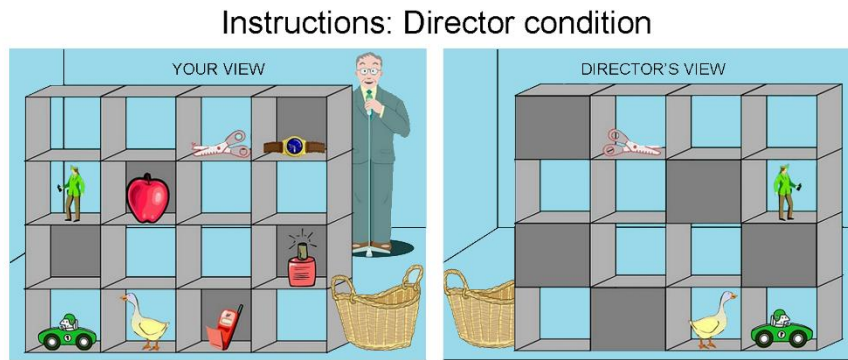


Figure 1. Director task instructions. Before the experimental Director condition of the task, participants were shown and explained images of their view and the corresponding Director's view of a stimulus configuration with an example of an object that both the participant and the Director can see ("car"), and an example of an object that the participant, but not the Director, can see ("apple").

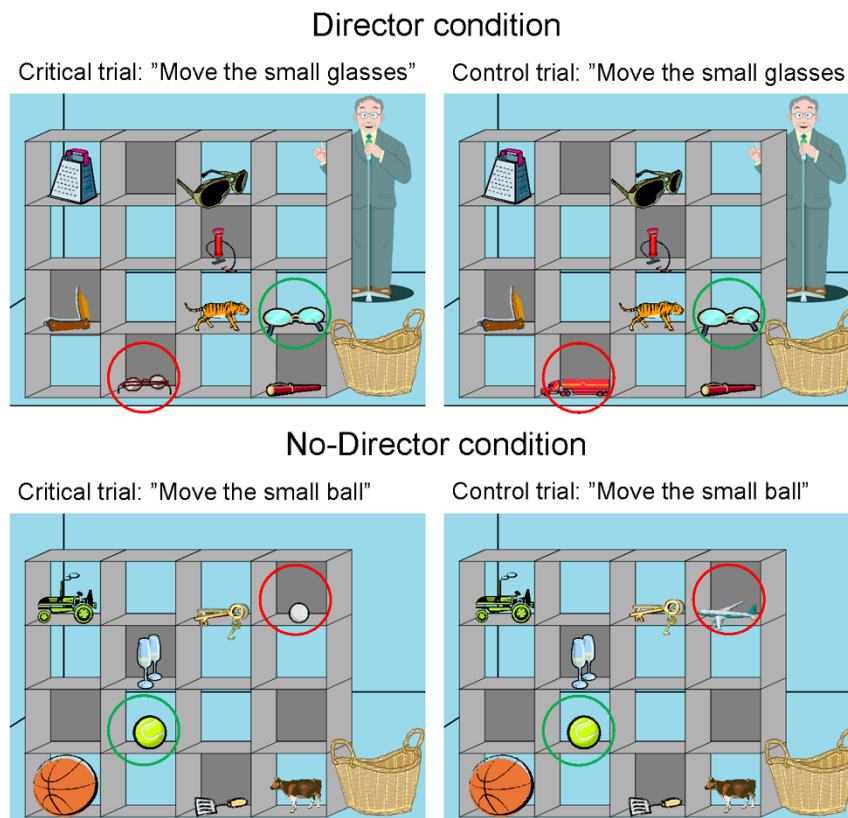


Figure 2. Director task. Top panel: Examples of a Critical trial and a Control trial in the experimental Director condition. Bottom panel: Examples of a Critical trial and a Control trial in the control No-Director condition. For illustration purposes the target stimulus is highlighted with a green circle, while the distractor/irrelevant object is highlighted with a red circle.

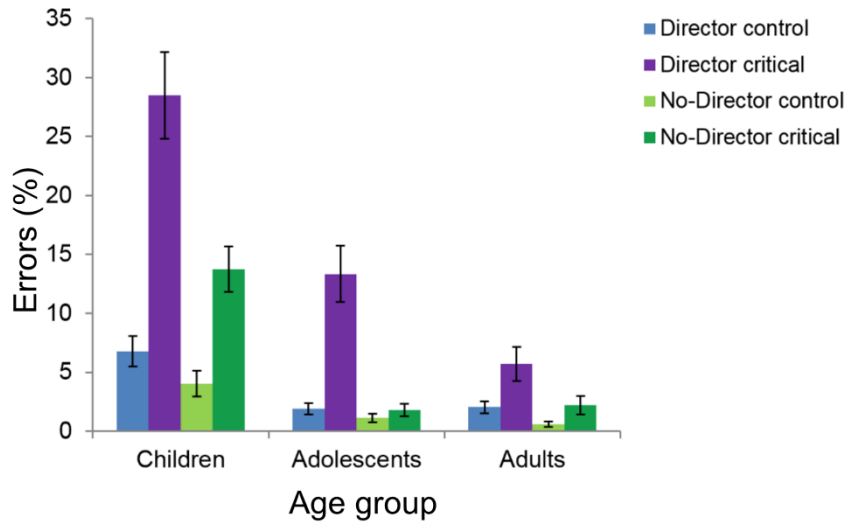


Figure 3. Director task performance: accuracy. Percentage errors (mean and standard errors) in Control trials and Critical trials in the Director condition and the No-Director condition for each age group. Children: 7.1–11.3 years (n=60), Adolescents: 11.7–17.9 years (n=108), and Adults: 18.0–26.7 years (n=125).

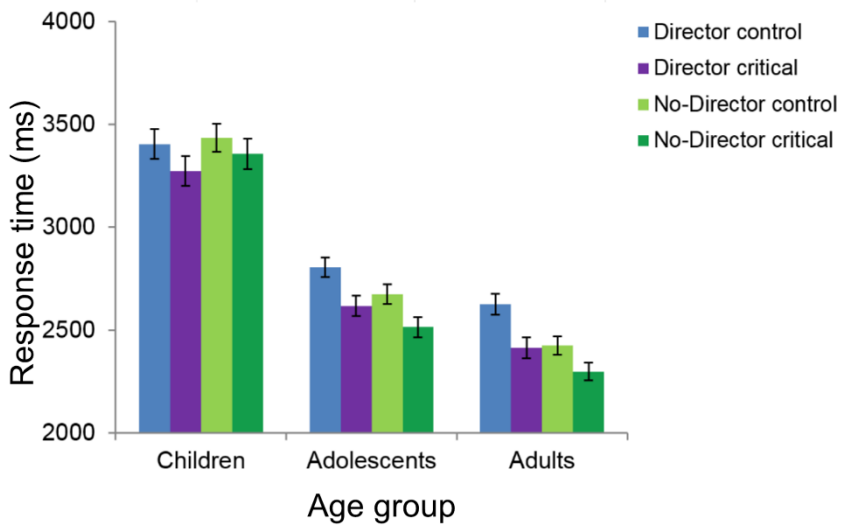


Figure 4. Director task performance: response time. Median response times (mean and standard errors) from correct trials only in Control trials and Critical trials in the Director condition and the No-Director condition for each age group. Children: 7.1–11.3 years (n=60), Adolescents: 11.7–17.9 years (n=108), and Adults: 18.0–26.7 years (n=125).

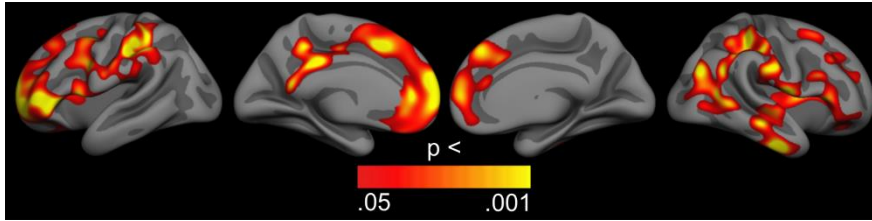


Figure 5. Associations between task performance and cortical thickness. GLMs were used to test the effects of the difference in percentage errors on Director Critical trials and on No-Director Critical trials on cortical thickness, while controlling for sex. The results were corrected for multiple comparisons using cluster size inference. Uncorrected p values within the corrected significant clusters are shown. All clusters showed positive effects, indicating that better performance was related to thinner cortices. No effects were seen in the opposite direction.

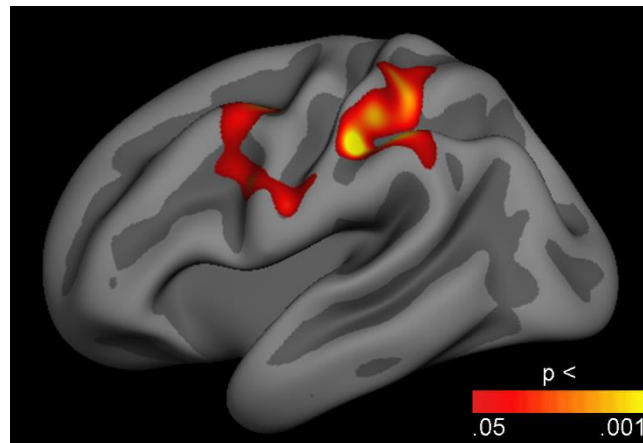


Figure 6. Age-independent associations between task performance and cortical thickness. GLMs were used to test the effects of the difference in percentage errors on Director Critical trials and on No-Director Critical trials on cortical thickness, while controlling for sex and age. The results were corrected for multiple comparisons using cluster size inference. Uncorrected p values within the corrected significant clusters are shown. Two clusters in the left hemisphere showed positive effects, indicating that better performance was related to thinner regional cortices. No effects were seen in the opposite direction.

Tables

Table 1. Task performance summary

	Errors (%)		Response time (ms)	
	Mean	SD	Mean	SD
Director filler trials	1.6	3.3	2211.1	353.5
Director control trials	3.0	6.9	2851.2	616.8
Director critical trials	13.2	23.9	2664.4	632.4
No-Director filler trials	1.4	2.7	2211.8	402.4
No-Director control trials	1.5	4.9	2723.9	633.4
No-Director critical trials	4.4	10.6	2594.1	648.2
Difference Director critical- No-Director critical	8.8	21.6		

n=293.

Supplemental Material

Materials and Methods

Statistical analysis

To test whether the age-related differences found in the analyses using age groups held when treating age as a continuous variable, additional analyses were performed. For both accuracy and RT, we ran a mixed ANOVA with Condition (Director, No-Director) and Trial type (Critical, Control) as within-subject factors and age (standardized) as a covariate. Follow-up analyses were performed to further investigate significant interaction and main effects.

Results

Director task: Accuracy

All main effects were significant in a mixed ANOVA on accuracy with Condition (Director, No-Director) and Trial type (Critical, Control) as within-subject factors and age as covariate. Participants made more errors in the Director condition than in the No-Director condition ($F(1,291) = 60.40, p < .001, \eta_p^2 = .172$), more errors on Critical trials than on Control trials ($F(1,291) = 69.49, p < .001, \eta_p^2 = .193$), and accuracy showed an age-related improvement ($F(1,291) = 80.02, p < .001, \eta_p^2 = .216$). There was a significant interaction between Condition and Trial type ($F(1,291) = 30.47, p < .001, \eta_p^2 = .095$), between Condition and age ($F(1,291) = 14.33, p < .001, \eta_p^2 = .047$), and between Trial type and age ($F(1,291) = 38.54, p < .001, \eta_p^2 = .117$). The three-way interaction was also significant ($F(1,291) = 10.71, p = .001, \eta_p^2 = .035$), and was explored further by looking at Critical and Control trials separately.

A mixed ANOVA performed on the Critical trials showed main effects of Condition ($F(1,291) = 50.42, p < .001, \eta_p^2 = .148$), with more errors in the Director condition, and age, with decreasing percentage errors with age ($F(1,291) = 64.04, p < .001, \eta_p^2 = .180$), as well as a significant interaction between Condition and age ($F(1,291) = 14.22, p < .001, \eta_p^2 = .047$). The same analysis on the Control trials only also showed significant main effects of Condition ($F(1,291) = 9.98, p = .002, \eta_p^2 = .033$) and age ($F(1,291) = 34.44, p < .001, \eta_p^2 = .106$), but no significant interaction effect ($F(1,291) = 0.51, p = .478, \eta_p^2 = .002$).

Analyses on the Critical trials in the two conditions separately showed a significant decrease in accuracy with age in both the Director condition ($F(1,291) = 44.46, p < .001, \eta_p^2 = .133$) and the No-Director condition ($F(1,291) = 49.47, p < .001, \eta_p^2 = .145$). A follow-up analysis on the difference in percentage errors on Director Critical trials and on No-Director Critical trials showed a significant effect of age ($F(1,291) = 14.22, p < .001, \eta_p^2 = .047$), i.e. a steeper slope corresponding to greater improvement in performance with age in Director Critical than No-Director Critical trials.

Director task: Response times

In a mixed ANOVA on RT with Condition (Director, No-Director) and Trial type (Critical, Control) as within-subject factors and age as covariate, the main effects of Condition ($F(1,291) = 18.81, p < .001, \eta_p^2 = .061$), Trial type ($F(1,291) = 79.57, p < .001, \eta_p^2 = .215$), and age ($F(1,291) = 172.07, p < .001, \eta_p^2 = .372$) were all significant. There was a significant interaction between Condition and age ($F(1,291) = 13.80, p < .001, \eta_p^2 = .045$), but no significant interactions between Trial type and age ($F(1,291) = 1.68, p = .196, \eta_p^2 = .006$), or Condition and Trial type ($F(1,291) = 3.39, p = .067, \eta_p^2 = .012$). The three-way interaction was not significant ($F(1,291) < 0.01, p = .996, \eta_p^2 < .001$). Due to the interaction between Condition and age, the main effect of age on RTs was explored further in each Condition separately. The results showed a significant effect of age on RT in both the Director condition ($F(1,291) = 105.43, p < .001, \eta_p^2 = .266$) and the No-Director condition ($F(1,291) = 193.60, p < .001, \eta_p^2 = .399$), with a steeper decrease in RT with age in the No-Director condition than in the Director condition.