Developmental Science

Developmental Science 13:2 (2010), pp 331–338 DOI: 10.1111/j.1467-7687.2009.00888.x

PAPER

Online usage of theory of mind continues to develop in late adolescence

Iroise Dumontheil,[1] Ian A. Apperly[2] and Sarah-Jayne Blakemore[1]

1. *Institute of Cognitive Neuroscience, University College London, UK*
2. *School of Psychology, Birmingham University, UK*

Abstract

The development of theory of mind use was investigated by giving a computerized task to 177 female participants divided into five age groups: Child I (7.3–9.7 years); Child II (9.8–11.4); Adolescent I (11.5–13.9); Adolescent II (14.0–17.7); Adults (19.1–27.5). Participants viewed a set of shelves containing objects, which they were instructed to move by a 'director' who could see some but not all of the objects. Correct interpretation of critical instructions required participants to use the director's perspective and only move objects that the director could see. In a control condition, participants were asked to ignore objects in slots with a grey background. Accuracy improved similarly in both conditions between Child I and Adolescent II. However, while performance of the Adolescent II and Adult groups did not differ in the control condition, the Adolescent II group made more errors than the adults in the experimental condition. These results suggest that theory of mind use improves between late adolescence and adulthood. Thus, while theory of mind tasks are passed by age 4, these data indicate that the interaction between theory of mind and executive functions continues to develop in late adolescence.

Introduction

Theory of mind – the ability to attribute mental states such as beliefs, desires and intentions – has been the subject of much research in developmental psychology and, more recently, in neuroscience. A large body of research indicates that theory of mind develops in the first few years in typically developing children: basic perspective taking emerges in the first 18 months (Sodian, Thoermer & Metz, 2007), understanding false belief by 4 years (Wellman, Cross & Watson, 2001; or younger: Onishi & Baillargeon, 2005; Surian, Caldi & Sperber, 2007) and second order metarepresentation by 6 or 7 (Perner & Wimmer, 1985). This early development of theory of mind sits uncomfortably with the finding from a large number of neuroimaging studies that brain regions critically involved in mental state attribution, in particular medial prefrontal cortex and lateral temporo-parietal regions, continue to develop both structurally (Giedd, Blumenthal, Jeffries, Castellanos, Liu, Zijdenbos, Paus, Evans & Rapoport, 1999; Shaw, Kabani, Lerch, Eckstrand, Lenroot, Gogtay, Greenstein, Clasen, Evans, Rapoport, Giedd & Wise, 2008; Sowell, Thompson, Holmes, Jernigan & Toga, 1999; Sowell, Thompson, Leonard, Welcome, Kan & Toga, 2004) and functionally (Blakemore, den Ouden, Choudhury & Frith, 2007; Wang, Lee, Sigman & Dapretto, 2006; Moriguchi, Ohnishi, Mori, Matsuda & Komaki, 2007; see Blakemore, 2008, for review) in the second and third decades of life. The protracted development in adolescence and early adulthood of the brain regions involved in theory of mind might be expected to affect mental state understanding.

There are several possible explanations for the lack of evidence of theory of mind development beyond early childhood. First, the tasks that have been used to test theory of mind in early development are not appropriate for testing older children and adolescents. Since most theory of mind tasks are passed by 5 years, ceiling effects might be obscuring the observation of any further development. Second, tasks typically directly enquire about children's representations of another person's mental states; they do not tap into how theory of mind is used to drive decisions and actions in everyday life. We hypothesized that, while theory of mind *per se* might not develop beyond early childhood (though see e.g. Chandler, Boyes & Ball, 1990; Kuhn, in press; Robinson & Apperly, 1998), the interaction between theory of mind and other cognitive processes such as executive functions continues to mature into adolescence. In order to test this hypothesis, and avoid ceiling effects in performance, we adapted a task developed by Keysar, Barr, Balin and Brauner (2000) and Keysar, Lin and Barr

Address for correspondence: Iroise Dumontheil, UCL Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK; e-mail: i.dumontheil@ucl.ac.uk

(2003) that showed that even adults have difficulty using theory of mind to guide behaviour.

Keysar *et al.* (2000, 2003) report that adults frequently fail to use their conceptual competence for theory of mind in an online communication game in which they need to take account of a speaker's perspective. Keysar and colleagues designed a referential communication task in which participants viewed a 4 × 4 grid. The grid contained various objects in different slots, and participants were instructed by a 'director' (a confederate) to move certain objects around the grid (see Figure 1a for a schematic of the experimental setup based on our own stimuli). Certain slots in the grid were occluded, thus the director could see some but not all of the objects visible to the participant. Critical instructions required the participant to use the information about the director's perspective to interpret instructions. Although clearly capable of understanding that the director has a different perspective, adult participants frequently failed to use this information when interpreting the director's instructions. This can be considered as evidence that humans are prone to egocentric bias. Evidence from a group of 4- to 12-year-olds suggests that children are more prone to such egocentric errors than adults (Epley, Morewedge & Keysar, 2004b).

To investigate the development of theory of mind use between late childhood and adulthood, we adapted Keysar *et al.*'s Director task so that it was suitable for children and presented on a computer. We tested the ability of 179 female participants aged between 7 and 27 years. In a control No-Director condition, participants were instructed to ignore objects in grey slots. Thus, both Director and No-Director conditions involved online inhibition of a prepotent response of moving the object that best fits the instruction from the participant's perspective, as well as general task demands such as rule following, working memory and so on. Thus, the two conditions were designed to be matched in terms of executive functions. The critical difference between conditions was that, in the Director condition, participants were instructed to take into account which objects the Director could and could not see, whereas in the No-director condition, participants were instructed to take into account the colour of the slot the object was in. Therefore, the only difference between conditions was that the Director condition involved the interaction between theory of mind (taking into account the director's perspective) and executive functions (inhibiting the egocentric bias and performing the appropriate motor action). Accuracy and response times were measured in all conditions. Based on the findings that the neural circuitry for theory of mind is developing during adolescence, we predicted that accuracy would improve with age in the Director condition over and above improvements in memory and inhibition abilities inherent to the No-Director condition.

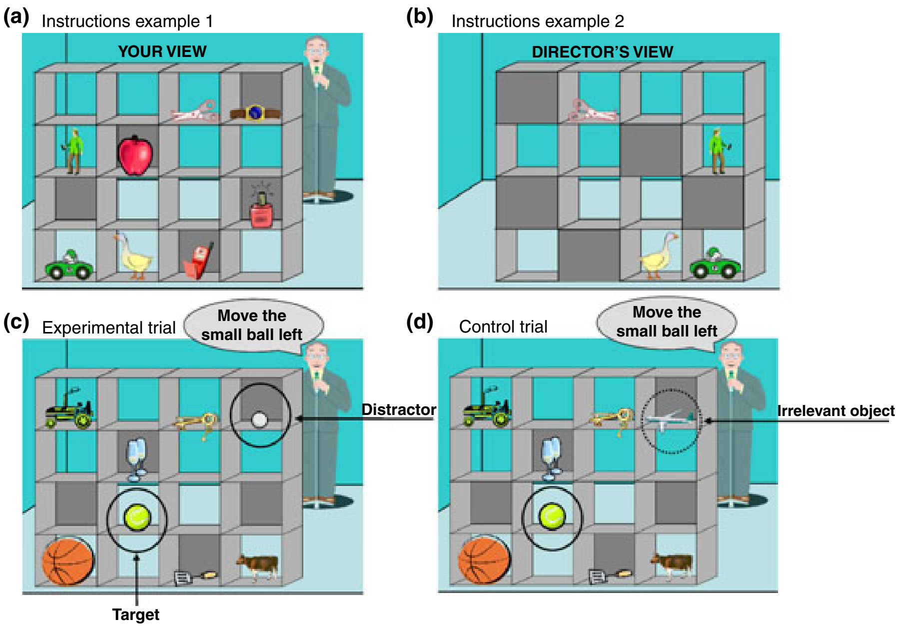

Figure 1 *(a–b) Images used to explain the Director condition to the participants: subjects were shown an example of their view (a) and the corresponding director's view (b) for a typical stimulus with four objects in occluded slots that the director could not see (e.g. the apple). (c–d) Example of an Experimental (c) and a Control trial (d) in the Director condition. The participant heard the verbal instruction: 'Move the small ball left' from the director. In the Experimental trial (c), if the participant ignored the director's perspective, she would choose to move the distractor ball (golf ball), which is the smallest ball in the shelves but which cannot be seen by the director, instead of the larger ball (tennis ball) shared by both the participant's and the instructor's perspective (target). In the Control trial (d), an irrelevant object (plane) replaces the distractor item.*

Method and materials

Participants

One hundred and seventy-nine female volunteers between the ages of 7.3 and 27.5 were recruited for this study. Children and adolescents were recruited from two London schools for girls, while adults were recruited from the UCL Psychology Department volunteer database. All participants spoke English as their first language. The child and adolescent participants were divided according to age into four groups of similar *N* to the adult group. Verbal ability was measured in children using the British Picture Vocabulary Scale II scores (BPVS II; Dunn, Dunn, Whetton & Burley, 1997), which is quick to administer, and in adults using the vocabulary subtest of the WASI (Wechsler, 1999). Data from two adolescents were excluded from the analysis: one had a verbal IQ score of less than 75; the other did not respond to any trials of one experimental condition. Table 1 presents details of all participants whose data were included in the analyses.

There was no significant difference between the verbal IQs of the groups (one-way ANOVA: $F(4, 172) = 0.61$, $p > .6$). Informed consent was obtained from the primary caregiver of each child and adolescent participant, and from the adults, and the study was approved by the local ethics committee.

Design

This experiment used a mixed-design with two within-subjects factors (Condition: Director, No-Director; and Trial type: Control, Experimental) and one between-subjects factor (Age group: Child I, Child II, Adolescent I, Adolescent II, Adults).

A computer simulation based on the task designed by Keysar *et al.* (2000) was used (task developed by Apperly, I.A., Carroll, D.J., Samson, D., Humphreys, G.W., Qureshi, A. & Moffatt, G., personal communication). The program was written using E-Prime version 2.0 (Psychology Software Tools, Inc.) and was presented on a laptop. The stimuli showed a 4 × 4 set of shelves containing eight different objects (see Figure 1 for example of stimuli). In the *Director* condition, five slots were occluded from the view of the 'director', who stood on the other side of the shelves and therefore viewed the shelves from behind (see Figure 1). The participant was asked to listen to instructions given by the director (heard through computer speakers). In each trial, the director instructed the participant to move one of the eight objects in a particular direction. Using a computer mouse, participants were required to click on the object they thought the director was referring to and to drag it into the appropriate slot on the shelves.

Experimental instructions required participants to take account of the director's perspective (see Figure 1c). 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 access was taken into account. For example in Figure 1c when the director asks to move the small ball left, the correct response would be to move the tennis ball, which is the smaller of the two balls visible to the director. If participants ignored his perspective they would select the 'distractor' object, which was invisible to the director. In Figure 1c, the incorrect response would be to move the golf ball, which is the smallest ball in the display, but which is invisible to the director. In the trials with the *Control* instruction, the arrangement of the objects in the shelves was identical to that in the Experimental instruction trials, except that an irrelevant object replaced the distractor object (e.g. the plane on Figure 1d). *Filler* trial instructions referred only to objects in clear slots, i.e. visible to both director and participant. For example on Figure 1c the director could ask to move the tractor right. The order of the Filler, Control and Experimental trials was counterbalanced between subjects.

In the *No-Director* condition, participants were told that the director had gone and they would hear instructions to move objects again and that these instructions would refer only to items in the clear slots; thus, objects in slots with a grey background should be ignored. The No-Director trials were identical in every way to the Director trials except that, instead of having to take into account the director's perspective, participants had to follow the rule of ignoring all objects in slots with a grey background. Experimental, Control and Filler trials were included in the No-Director condition, and trial order was counterbalanced between subjects.

Two sets of eight different shelf–object configurations were created, each presented once with an occluded distractor object (Experimental trial) and once with an irrelevant object (Control trial). One set was presented in the Director condition, the other in the No-Director condition, thus the stimuli were not repeated for individual subjects. The sets were counterbalanced across subjects. Each stimulus was presented for 2 seconds before the first auditory instruction was given. Three auditory instructions were given per stimulus and each lasted 2.2 seconds, and participants were given an additional 3.6 seconds to make their

Table 1 *Age and verbal IQ (BPVS II in children, WASI in adults) of the five groups of participants (all female)*

		Age (years)			Verbal IQ		
Groups	*N*	Mean	*SD*	Range	Mean	*SD*	Range
Child I	35	8.9	0.7	7.3–9.7	117.0	9.6	98–153
Child II	36	10.6	0.5	9.8–11.4	116.7	9.5	92–143
Adolescent I	35	12.7	0.8	11.5–13.9	117.1	15.6	90–158
Adolescent II	35	15.3	1.2	14.0–17.7	114.3	18.7	87–156
Adults	36	22.8	2.3	19.1–27.5	119.5	15.1	85–138

response. Each display was presented with two Filler instructions and one Control or Relational instruction. In total there were thus eight Control trials, eight Experimental trials, and 48 Filler trials in each condition (Director and No-Director). The order of stimulus presentation was counterbalanced between participants. Each condition lasted approximately 5.5 minutes.

Procedure

Standardized instructions were read to the participants and they were shown an example stimulus. For the Director condition it was explained that, on each trial, the director would instruct the participant which object to move and where to move it. Emphasis was placed on the fact that the director had a different perspective to the participant by showing participants an example of the director's view of the shelves (see Figure 1a and 1b). Each participant was asked to give an example of an object that only she, and not the director, could see (i.e. in an occluded slot), and an object that both she and the director could see (i.e. in a clear slot), to demonstrate that she understood that the director had a different perspective from hers. All participants performed this correctly, indicating that they had understood the instructions and that they knew the director could not see all the objects, and they were not given further feedback regarding the requirement to take the director's perspective into account. Before the start of the No-Director condition, new instructions were read and participants were shown an example of a No-Director stimulus and asked to give an example of an object that was in a slot with a grey background. Participants were then asked to move an object as they would in the experiment to demonstrate that they understood what was required of them. All participants were tested individually in a quiet room. All participants carried out the Director condition before the No-Director condition in order to prevent participants from applying the strategy provided in the No-Director condition to the Director condition.

Data analysis

Mean accuracy and median response times in correctly responded trials were calculated for each participant in each Condition (Director/No-Director) and Trial type (Control/Experimental). Mixed model repeated measures ANOVA with two within-subject factors and one between-subject factor (Age group: five levels) were performed on group mean accuracy and group mean response times. Bonferroni-corrected post-hoc independent or paired *t*-tests were performed to investigate further significant main effects and interactions. Possible floor and ceiling effects in accuracy were investigated by comparing performance in Experimental and Control trials in each condition and each age group using paired *t*-tests. Statistical analysis results are provided with standard *p*-values and effect sizes: Cohen's *d* for *t*-tests (Cohen, 1969), $d = 0.2$, 0.5 and 0.8 correspond respectively to small, medium and large effect sizes (Cohen, 1992); and partial eta squared (η_p^2) for *F*-tests, which is the proportion of the effect plus error variance that is attributable to the effect (Cohen, 1973).

Results

Accuracy data

Participants made fewer than 3% errors in Filler trials on average, and the data for these trials were not analysed. The mean accuracy in the critical (Director, Experimental) condition reflected a range of accuracies across subjects, as predicted on the basis of previous work on adults, rather than a bimodal distribution with participants either doing the task well or failing it completely.

A 2 × 2 × 5 mixed model repeated measures ANOVA with Condition (Director/No-Director), Trial type (Control/Experimental) and Age group (Child I, Child II, Adolescent I, Adolescent II and Adults) was performed on accuracy (see Figure 2). All main effects were significant: participants made more errors in Experimental than Control trials ($F(1, 172) = 684.04$, $p < .001$, $\eta_p^2 = .799$); more errors in the Director than the No-Director condition ($F(1, 172) = 553.25$, $p < .001$, $\eta_p^2 = .763$); and accuracy changed with age ($F(4, 172) = 8.94$, $p < .001$, $\eta_p^2 = .172$). There was a significant interaction between Trial type and Age group ($F(4, 172) = 7.45$, $p < .001$, $\eta_p^2 = .148$), between Condition and Age group ($F(4, 172) = 750.10$, $p = .015$, $\eta_p^2 = .068$) and between Condition and Trial type ($F(1, 172) = 548.76$, $p < .001$, $\eta_p^2 = .761$). The three-

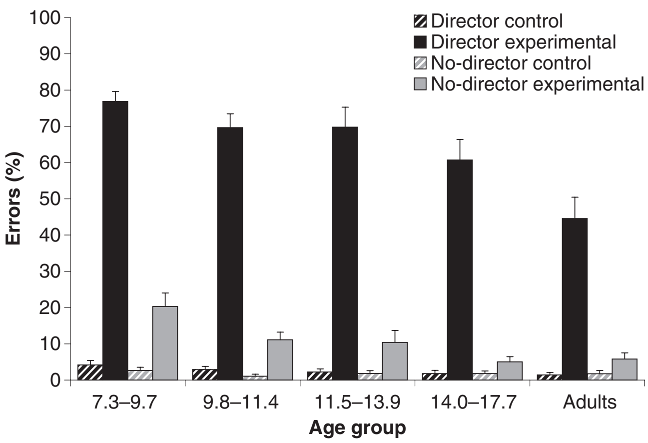

Figure 2 *Average percentage errors (mean + SE) in Control and Experimental trials in the Director and No-Director conditions for each age group.*

way interaction was also significant ($F(4, 172) = 2.52$, $p = .043$, $\eta_p^2 = .055$) and was explored further by looking at Experimental and Control trials separately.

A 2 × 5 repeated measures ANOVA performed on the Experimental trials only showed a significant effect of Condition ($F(1, 172) = 585.54$, $p < .001$, $\eta_p^2 = .773$), with more errors in the Director condition than in the No-Director condition. There was also a main effect of Age group ($F(4, 172) = 8.49$, $p < .001$, $\eta_p^2 = .165$), and a significant interaction between Age group and Condition ($F(4, 172) = 2.99$, $p = .02$, $\eta_p^2 = .065$). A similar 2 × 5 repeated measures ANOVA performed on the Control trials revealed no significant main effect (Condition: $F(1, 172) = 1.59$, $p > .2$; Age group: $F(4, 172) = 1.45$, $p > .2$) or interaction ($F(4, 172) = .59$, $p > .6$).

Performance in Experimental trials was explored further. A 2 × 4 repeated measures ANOVA performed on the two child and two adolescent groups and Condition (Director/No-Director) revealed no Condition by Age group interaction ($F(3, 137) = .12$, $p > .9$), suggesting that the 2 × 5 interaction effect was driven by the adults. However, the main effects of Condition ($F(1, 137) = 563.84, p < .001$, $\eta_p^2 = .805$), and Age group ($F(3, 137) = 4.79$, $p = .003$, $\eta_p^2 = .095$) were significant. Post-hoc *t*-tests on accuracy averaged between the two conditions indicated that only the Child I (7.3–9.7 years old) and Adolescent II groups (14.0–17.7) differed significantly ($t(68) = 3.87$, $p < .001$, $d = .93$), with Child I participants making more errors. A 2 × 2 repeated measures ANOVA performed on the Adolescent II group (aged 14.0–17.7) and the adults showed a significant main effect of Condition ($F(1, 69) = 152.65, p < .001$, $\eta_p^2 = .689$) as well as a Condition by Age group interaction ($F(1, 69) = 4.91$, $p = .03$, $\eta_p^2 = .066$). There was no significant main effect of Age group ($F(1, 69) = 2.78, p = .1$). Follow-up *t*-tests showed that the adolescent participants made marginally significantly more errors than adults in the Director ($t(69) = 1.98$, $p = .052$, $d = .47$), but not in the No-Director condition ($t(69) = .34$, $p > 0.3$).

Two additional analyses were performed. First, we addressed the issue of whether participants were performing at floor in the Director Experimental trials. If the children were never taking the Director's perspective into consideration, one could expect them to make a similar number of 'errors' in the Experimental trials, i.e. *not* selecting the Distractor item, to the number of errors made in the Control trials. The Control trials were thus used as a baseline, and the percentage of correct responses in the Experimental trials was compared to the percentage of errors in the Control trials of the Director condition. All age groups showed significant differences between the two types of trials (all *p*s $< .001$, $d > 1.1$), suggesting no group performed at floor in the Director Experimental trials.

Second, we addressed the issue of whether participants were performing at ceiling in the No-Director Experimental trials. Accuracy in Experimental trials was compared to accuracy in No-Director Control trials. All age groups showed significant differences between these two types of trial (all *p*s $< .05$, $d > .45$), suggesting that no group performed at ceiling in No-Director Experimental trials.

Response time data

The median response times (RTs) were calculated from correct responses for each subject. Subjects with no correct response in one of the conditions were omitted from the analysis (resulting group sizes: Child I $n = 29$, Child II $n = 31$, Adolescent I $n = 25$, Adolescent II $n = 26$, Adults $n = 31$). Mean RTs were calculated for each group (see Figure 3). Note that the same analyses performed on both correct and incorrect trials together showed similar results.

A 2 × 2 × 5 ANOVA performed on the RTs showed that all main effects were significant: RTs were slower in the No-Director than in the Director condition ($F(1, 137) = 37.92$, $p < .001$, $\eta_p^2 = .217$), slower in Control than in Experimental trials ($F(1, 137) = 8.82$, $p = .004$, $\eta_p^2 = .061$), and changed with age ($F(4, 137) = 4.56$, $p = .002$, $\eta_p^2 = .118$). Post-hoc *t*-tests indicated that Child I participants (7.3–9.7 years old) were slower than Adolescent II participants (14.0–17.7 years old; $t(53) = 3.66$, $p = .006$, $d = .986$) and than the adults ($t(58) = 3.65$, $p = .006$, $d = .943$); no other between-group difference reached significance. There were interactions between Condition and Trial type ($F(1, 137) = 4.46$, $p = .037$, $\eta_p^2 = .032$), with a greater RT difference between Experimental and Control trials in the Director than in the No-Director condition, and between Condition and Age group ($F(4, 137) = 6.22$, $p < .001$, $\eta_p^2 = .154$). This latter interaction effect was investigated further. RTs did not change with age in the Director condition ($F(4, 137) = .99$, $p > 0.4$), whereas

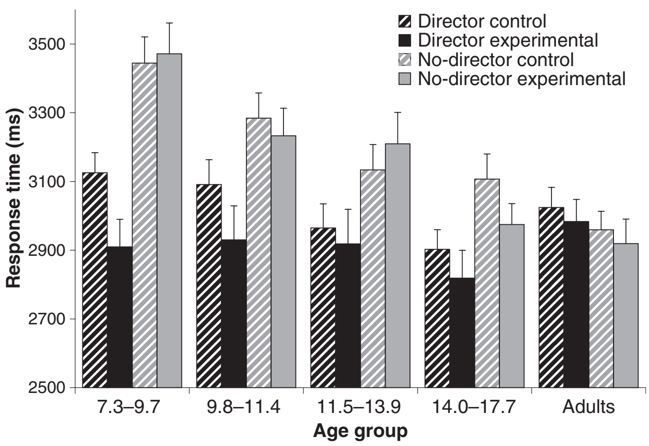

Figure 3 *Average response times (mean + SE) from correct trials only in Control and Experimental trials of the Director and No-Director conditions for each age group.*

there was a significant effect of Age group in the No-Director condition ($F(4, 137) = 8.82$, $p < .001$, $\eta_p^2 = .205$). Post-hoc *t*-tests showed that, in the No-Director condition, the Child I group responded more slowly than the Adolescent II group ($t(53) = 4.13$, $p = .001$, $d = 1.12$) and than the adults ($t(58) = 5.42$, $p < .001$, $d = 1.40$), and the Child II group responded more slowly than the adults ($t(60) = 3.60$, $p = .007$, $d = .913$). The two remaining interactions were not significant: Trial type by Age group ($F(4, 137) = 1.05$, $p = .4$), and the three-way interaction between Condition, Trial type and Age group ($F(4, 137) = 1.59$, $p = .2$).

Discussion

In this study, we tested the ability of a large sample of children, adolescents and young adults (aged 7–27) to use information about another person's perspective when following their instructions. Critical trials required participants to use information about the director's perspective, i.e. which objects he could see and which he could not, to interpret his instructions and respond appropriately by inhibiting their egocentric bias. In the control No-Director condition, participants were instructed to ignore objects in particular locations. Thus, in the critical Director trials, participants had to take into account the fact that the director was unable to see (and therefore could not be referring to) certain objects, whereas in the No-Director condition, participants were given an explicit rule to facilitate performance. Both Director and No-Director conditions required a variety of executive functions. In addition, the Director condition also required level-1 perspective taking, the ability to represent what another person can see (Flavell, Everett, Croft & Flavell, 1981). This ability is a core component of theory of mind since, to predict and explain other person's behaviour, we make inferences about their knowledge or beliefs on the basis of their visual access. In the Director condition, participants used information derived from level-1 perspective taking to infer what the Director knew, and therefore what object he could be referring to, and then had to perform the appropriate motor action towards that object.

In the current computerized task, all participants were able to describe which objects the director could and could not see when directly prompted during the practice. This demonstrated that all participants were able to achieve level-1 perspective taking, as would be expected. However, during the testing phase, a large proportion of participants in all groups frequently failed to use information about the director's perspective to interpret his instruction and move the appropriate object. These findings are consistent with a real-life version of the task (Keysar *et al.*, 2000, 2003), in which participants had to pass objects in a set of shelves to a real director whose view of some of the objects was physically obscured. The data fit with other suggestions that adult perspective taking is subject to egocentric or 'reality' bias (e.g. Birch & Bloom, 2004, 2007; Epley, Keysar, Van Boven & Gilovich, 2004a; Mitchell, Robinson, Isaacs & Nye, 1996).

The developmental results indicate that accuracy improved in a similar way in both Director and No-Director trials in early adolescence. However, in the No-Director condition, there was no further improvement in accuracy beyond adolescence (14–17.7 years), whereas in the Director condition, accuracy continued to improve between adolescence and adulthood. A possible floor effect in the Director condition in the younger children and a ceiling effect in the No-Director condition in older participants were tested for and did not appear to drive the critical interaction observed between the older adolescents and the adults. Following instructions in both the Director and No-Director conditions involves holding the task rule in mind over the whole block, and potentially inhibiting a prepotent response (towards the distractor object) on a trial-by-trial basis. The initial parallel improvement in accuracy with age observed in both conditions is in line with previous studies demonstrating development beyond childhood of certain executive functions, such as inhibition of a prepotent response (Casey, Trainor, Orendi, Schubert, Nystrom, Giedd, Castellanos, Haxby, Noll, Cohen, Forman, Dahl & Rapoport, 1997; Tamm, Menon & Reiss, 2002) and working memory (e.g. Anderson, Anderson, Northam, Jacobs & Catroppa, 2001; see Romine & Reynolds, 2005, for a review). However, the continued improvement of accuracy only in the Director condition suggests that the ability to take account of another person's perspective to direct appropriate behaviour is still improving in late adolescence, after working memory and response inhibition abilities recruited in this task have reached adult levels.

An improvement with age in the time taken for correct responses was observed in the No-Director condition and, together with the improvements in accuracy, is likely to reflect the maturation of the ability to inhibit a prepotent response (towards the object in the grey background) while holding information in mind (see Romine & Reynolds, 2005, for a review). There was no evidence of age-related changes in response time in the Director condition. Moreover, the time taken for correct use of the Director's perspective was relatively fast compared with the No-Director condition. This difference in response times may indicate that participants were faster to inhibit a response towards an object that was not seen by another person than towards an object arbitrarily selected by a memorized rule. Response times were calculated from correct trials only (though in fact similar patterns are observed if both correct and incorrect responses are combined). The difference in RT between the Director and No-Director conditions suggests that participants who were able to take into account the perspective of the director did this faster than when they were required to ignore objects in a grey background. This could indicate that, when answering correctly to Experimental trials of the Director condition, participants did not simply apply a rule similar to that of the No-Director condition, and that this strategy, possibly related to real-life properties of objects and occlusions, was more efficient, once in place, than the arbitrary rule given by the experimenter. This raises the interesting possibility that the reduction in error rate in the Director condition between older adolescents and adults was not due to increases in the efficiency of perspective-taking processes. Rather, it may be that these perspective-taking processes are relatively efficient throughout the developmental period studied here but that an important additional change is the *propensity* for participants to take account of a speaker's perspective. We speculate that changes in such higher-level strategies for the use of 'theory of mind' may be an important locus of development over and above improvements in the efficiency of basic theory of mind processes.

Our data extend previous developmental studies using similar paradigms. Epley and colleagues found that children aged 4 to 12 years are more prone to egocentric errors (ignoring the director's perspective in Experimental trials) than are adults (Epley *et al*., 2004b). In contrast, Nadig and Sedivy (2002) found that 6-year-olds' eye movements showed sensitivity to the director's perspective. Reasons for the discrepancy between these results include the fact that Nadig and Sedivy's task was much simpler, using a 2×2 rather than a 4×4 array, and also that the director's instructions were ambiguous if the speaker's perspective was ignored, which may have prompted children to take account of the speaker's perspective (see Keysar *et al*., 2003, for discussion). It would be interesting to record gaze behaviour in a more complex array, as was used in the current study and by Epley *et al*. (2004b), to investigate possible implicit processing (reflected in participants' eye movements) of the director's perspective during adolescence.

This is the first time that an empirical study has shown evidence of such late development on a task that involves representing another person's mental states. There is a long history of research on the early development of theory of mind, which has consistently shown that false belief tasks are normally passed by age 4 or 5 (Wellman *et al*., 2001) or even earlier (Onishi & Baillargeon, 2005; Surian *et al*., 2007). Very few studies have investigated theory of mind performance development beyond early childhood (e.g. Perner & Wimmer, 1985). Here we suggest that the improvement until mid-adolescence in the capacity to meet the demands that both Director and No-Director conditions make on working memory and inhibitory control is followed during late adolescence by an additional age-related increase in participants' propensity to take account of a speaker's perspective to guide behaviour. We suggest that this developmental pattern reflects continuing maturation of the interaction between theory of mind and executive functions. While the current study cannot determine the cause of this late development, our data fit with recent neuroimaging studies showing that brain regions critically involved in mental state attribution, in particular, medial prefrontal cortex and lateral temporo-parietal regions, continue to develop both structurally (e.g. Giedd *et al*., 1999; Shaw *et al*., 2008) and functionally (see Blakemore, 2008, for review) during adolescence. A priority for future work is to determine how this neural development contributes to a gradual, and protracted, improvement in the use of theory of mind for everyday action, and whether this is due to changes in motivation for taking account of speakers' perspectives (e.g. as observed in chimpanzees; Bräuer, Call & Tomasello, 2007), or whether theory of mind use becomes slowly automatized and integrated with cognitive control systems, which may help participants resist interference from their own perspective (Apperly, Back, Samson & France, 2008).

References

Anderson, V.A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, **20**, 385–406.

Apperly, I.A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, **106**, 1093–1108.

Birch, S.A.J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, **8**, 255–260.

Birch, S.A.J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, **18**, 382–386.

Blakemore, S.J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, **9**, 267–277.

Blakemore, S.J., den Ouden, H., Choudhury, S., & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive and Affective Neuroscience*, **2**, 130–139.

Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition*, **10**, 439–448.

Casey, B.J., Trainor, R.J., Orendi, J.L., Schubert, A.B., Nystrom, L., Giedd, J.N., Castellanos, F.X., Haxby, J.V., Noll, D.C., Cohen, J.D., Forman, S.D., Dahl, R.E., & Rapoport, J.L. (1997). A developmental functional MRI study of prefrontal activation during performance of a go/no-go task. *Journal of Cognitive Neuroscience*, **9**, 835–847.

Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*, **50**, 370–395.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, **33**, 107–112.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, **112**, 155–159.

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd edn.). Windsor: NFER-NELSON.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004a). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, **87**, 327–339.

Epley, N., Morewedge, C.K., & Keysar, B. (2004b). Perspective taking in children and adults: equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, **40**, 760–768.

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.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., & Rapoport, J.L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, **2**, 861–863.

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.

Keysar, B., Lin, S., & Barr, D.J. (2003). Limits on theory of mind use in adults. *Cognition*, **89**, 25–41.

Kuhn, D. (in press). The importance of learning about knowing: creating a foundation for development of intellectual values. *Perspectives on Child Development*.

Mitchell, P., Robinson, E.J., Isaacs, J.E., & Nye, R.M. (1996). Contamination in reasoning about false belief: an instance of realist bias in adults but not children. *Cognition*, **59**, 1–21.

Moriguchi, Y., Ohnishi, T., Mori, T., Matsuda, H., & Komaki, G. (2007). Changes of brain activity in the neural substrates for theory of mind during childhood and adolescence. *Psychiatric and Clinical Neurosciences*, **61**, 355–363.

Nadig, A.S., & Sedivy, J.C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, **13**, 329–336.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, **308**, 255–258.

Perner, J., & Wimmer, H. (1985). 'John thinks that Mary thinks that ...': attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, **39**, 437–471.

Robinson, E.J., & Apperly, I.A. (1998). Adolescents' and adults' views about the evidential basis for beliefs: relativism and determinism re-examined. *Developmental Science*, **1**, 279–290.

Romine, C.B., & Reynolds, C.R. (2005). A model of the development of frontal lobe functioning: findings from a meta-analysis. *Applied Neuropsychology*, **12**, 190–201.

Shaw, P., Kabani, N.J., Lerch, J.P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J.L., Giedd, J.N., & Wise, S.P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, **28**, 3586–3594.

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.

Sowell, E.R., Thompson, P.M., Holmes, C.J., Jernigan, T.L., & Toga, A.W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, **2**, 859–861.

Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., & Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, **24**, 8223–8231.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, **18**, 580–586.

Tamm, L., Menon, V., & Reiss, A.L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry*, **41**, 1231–1238.

Wang, A.T., Lee, S.S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social Cognitive and Affective Neuroscience*, **1**, 107–121.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.

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.

Received: 21 October 2008
Accepted: 21 January 2009