The cognitive demands of remembering a speaker's perspective and managing common ground size modulate 8- and 10-year-olds' perspective-taking abilities

Zhao, Lin; Wang, J Jessica; Apperly, Ian

DOI:
10.1016/j.jecp.2018.05.013

The cognitive demands of remembering a speaker's perspective and managing common ground size modulate 8- and 10-year-olds' perspective-taking abilities

Lin Zhao[a *]

J. Jessica Wang[a,b]

Ian A. Apperly[a]

[a] School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

[b] Department of Psychology, Lancaster University, Lancaster, LA1 4YF, United Kingdom

*Corresponding author email: zhaolin.bas@gmail.com

Abstract

Using "theory of mind" to successfully accommodate differing perspectives during communication requires much more than just acquiring basic theory of mind understanding. Evidence suggests that children's ability to adopt a speaker's perspective continues to develop through childhood to adolescence till adulthood (e.g., Dumontheil, Apperly, & Blakemore, 2010). The present study examined the cognitive factors that could account for variations in children's abilities to use a speaker's perspective during language comprehension, and whether the same factors contribute to age-related improvements. Our study incorporated into a commonly-used communication task two types of memory demands which are frequently present in our everyday communication but have been overlooked in the previous literature: remembering a speaker's perspective, and the amount of common ground information. Findings from two experiments demonstrated that both 8- and 10-year-olds committed more egocentric errors when each of these memory demands was high. Our study also found some supporting evidence for the age-related improvement in children's perspective use, as 10-year-olds generally committed fewer egocentric errors compared to 8-year-olds. Interestingly, there was no clear evidence that the memory factors that affected children's perspective use in our experiments were also the factors that drove age-related improvement.

Key words: common ground; theory of mind; perspective taking; cognitive factors; age-related development; referential communication

Introduction

We take into account other people's perspectives in order to successfully navigate through our everyday social interaction. Especially when in communicative situations where our own perspective differs from another interlocutor's, we have to draw upon "*common ground*", an important concept in the psycholinguistic literature which refers to the set of mutual knowledge and beliefs shared between speakers and listeners (Clark & Marshall, 1978, 1981), to guide our communication.

Perspective taking can be seen as a common instance of "theory of mind" use. Until relatively recently, a large number of developmental studies on the research area of theory of mind had focused on answering the question of when children develop this understanding of others' mental states and perspectives (e.g., Astington & Gopnik, 1991; Perner, Leekam, & Wimmer, 1987). Classic accounts suggest that children are egocentric, and incapable of understanding others' perspectives before around 7 years (Piaget, 1959; Piaget & Inhelder, 1956). More recent evidence suggests that theory of mind concepts develop significantly earlier, between 2 and 7 years (e.g., Wellman, Cross, & Watson, 2001), and may even be present in infants (e.g., Kovács, Téglás, & Endress, 2010; Moll & Tomasello, 2006; Onishi & Baillargeon, 2005; Sodian, Thoermer, & Metz, 2007). In contrast, recent research on older children and adults, who have clearly developed a basic understanding of mental states, shows that they are still egocentrically biased especially when facing communicative situations where differing perspectives have to be taken into account (e.g., Dumontheil, Apperly, & Blakemore, 2010; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003). This suggests that utilising theory of mind abilities in communication requires much more than just acquiring the necessary theory of mind concepts.

Much work on perspective taking has used communication games, in which participants have to follow or produce instructions to another interlocutor whose perspective differs from their own. For example, in a commonly-used "director task" (e.g., Keysar et al., 2000), participants were asked to follow a director's instruction to move the referred objects around a shelf. Some objects were blocked from the director's perspective and thus only remained visible in participants' *privileged ground*. Other objects were visible to both the director and participants, and thus in their *common ground*. Critical instructions required participants to only move the matching objects that were visible to both themselves and the director. A large number of studies employing this director task on healthy adults have demonstrated that adults suffer from interference of their own privileged perspective, and this egocentrism is revealed by selecting a distractor which is invisible to the director, by looking at the matching distractor before they reach the target object in common ground, or by taking longer time to select the target referent when the distractor is present compared to it being absent (Apperly et al., 2010; Keysar et al., 2000, 2003; Wu & Keysar, 2007). These findings accord with the suggestions that adults suffer from "realist bias" (Mitchell, Robinson, Isaacs, & Nye, 1996), or "curse of knowledge" (Birch & Bloom, 2007), which refers to the phenomenon that adults' judgement of another person's belief is often biased by their privileged knowledge of the reality.

Children, who are less experienced communicators than adults, may suffer more from this egocentric bias when required to accommodate differing perspectives in the director task. Although eye movement data from Nadig and Sedivy (2002) suggested that even 6-year-old children were sensitive to their interlocutor's limited perspective and were able to use this perspective information from the early stage of language processing, egocentrism did not completely evaporate in their study. Indeed, the first experiment in Nadig and Sedivy (2002) suggested that children speakers were not as reliable as adult speakers at providing adjectives to fully disambiguate the referents when there were two similar candidates in the common

ground. Epley, Morewedge, and Keysar (2004) demonstrated that 4- to 12-year-old children acted more egocentrically in the task, as they reached for hidden referents more frequently compared to the adult participants. They provided a plausible explanation that adults and children appeared to not differ in the initial processing stage during which they both processed information egocentrically, but differed in the later adjustment stage during which adults were more capable to correct their egocentric bias and accommodate their interlocutor's differing perspective than children. Moreover, Epley et al. (2004) also observed an incremental improvement in performance as children's age increased. Convergingly, Dumontheil et al. (2010; see also Symeonidou et al., 2016) found that children's ability to accommodate a speaker's limited perspective continued to develop through childhood to adolescence till adulthood. Similarly, Frick, Möhring, and Newcombe (2014) used an experimental paradigm where two agents saw the same sets of objects from two different visuo-spatial perspectives from which participants' own perspective could also differ. This study found that the capacity to inhibit egocentric choices and to recognize what agents saw from their own perspectives gradually developed between 5 to 8 years of age. These findings suggest that the use of information about others' perspectives goes through prolonged developmental improvements even after the acquisition of the basic theory of mind competence.

In the present study, we first sought to identify cognitive factors that could affect children's abilities to infer and utilize a speakers' perspective during language comprehension. Secondly, we considered whether the same factors would contribute to the age-related improvements in their perspective-taking performance. Most studies examining the relationship between cognitive capacities and perspective use in language comprehension have been conducted on adults, but the findings are currently mixed. For example, Brown-Schmidt (2009) using a variant on the director task with temporary referential ambiguity found evidence that individuals with more enhanced inhibitory control capacities were more capable to ignore a

distractor in the privileged ground, whereas no correlation was found between individual differences in working memory and their perspective-taking performance. In Lin, Keysar, and Epley (2010), participants with low working memory capacities showed enhanced degrees of egocentrism when performing the director task; and the relatively high working memory load from a secondary task reduced participants' abilities to use a speaker's perspective. However, using a similar version of the director task, Cane, Ferguson, and Apperly (2016) showed that individual differences in working memory did not predict listeners' ability to adopt the director's perspective. Overall, research on adults on the relationship between executive functions and perspective use in language comprehension are rather mixed and inconsistent.

Little research has been done to directly examine the cognitive factors that affect children's ability to use a speaker's perspective in referential communication. The most relevant developmental evidence comes from Nilsen and Graham (2009), where a positive correlation was found in children aged three to five between their inhibitory control skills and their perspective-taking abilities in language comprehension. A similar tendency was also observed for the measure of working memory capacities, but it did not reach significance. A more recent study conducted by Wang, Ali, Frisson, and Apperly (2016) found that 10-year-olds performed less egocentrically than 8-year-olds in the director task overall, and both age groups committed more egocentric errors when a director's instruction sentence was complex (e.g., "nudge the large jar one slot up") rather than simple (e.g., "nudge the large jar"). However, both 8- and 10-year-olds were affected by the cognitive demands of integrating complex messages with the speaker's limited perspective to the same degree. This finding suggests that the age-related improvement in children's perspective-taking performance could not be explained by the development of executive capacities that might be necessary for meeting this demand. Studies using other communication tasks have also demonstrated the important role of memory in children's performance in referential communication. For

example, Dahlgren and Sandberg (2008) found that for both children with autism spectrum and typically developing children, their free recall capacities were positively correlated to their performance in a referential communication task. Another study found that preschoolers' executive functions, compared to IQ, were more positively correlated to their pragmatic communicative skills during a semi-structured conversation with an experimenter (Blain-Brière, Bouchard, & Bigras, 2014). Specifically, children with higher working memory capacities were found to be more likely to generate contingent answers and produce clear utterances to their listeners.

The inconsistent conclusions from research on adults and the lack of relevant research on children on this topic have led us to question what and how memory capacities constrain children's ability to use a speaker's perspective to guide their communication. In order to answer these questions, we would need to identify the memory loads commonly embedded in our everyday communication and to find ways to implement these processes into the laboratory director task to tax participants' memory resources while they accommodate the director's perspective. Apperly et al. (2010) proposed that perspective taking can be decomposed into three sub-processes, i.e., computing perspective information, holding perspective information in mind and using perspective information. Therefore, the first embedded memory factor we identified was the demand of holding in mind another person's perspective for later use. The director task typically presented participants with a shelf on which privileged ground objects were displayed in slots with coloured background and common ground objects were displayed in open slots. The visual cues (i.e., coloured backgrounds), which indicated the perspective content of the director, usually remained available when the director was referring to objects on the shelf. We speculate that this experimental design in principle allows and even encourages participants to compute the director's perspective rapidly and efficiently online during comprehension of her instruction. Our speculation is consistent with evidence that

Level-1 visual perspectives (about whether something can be seen by someone or not) can be calculated rapidly and relatively automatically (Qureshi, Apperly, & Samson, 2010; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). In this case, the demands on memory for tracking the director's perspective in the traditional director tasks would be greatly reduced, or even eliminated. However, in most cases of our everyday fast-moving communicative situations, the content of our communicative partners' perspective is not continuously present, and thus requires us to remember and store relevant perspective information in our memory rather than simply computing it online. Therefore, in the present study we removed the visual cues indicating the director's perspective after an initial shelf-preview period. This manipulation forced participants to infer and store relevant perspective information before the point of need.

Secondly, the varied amount of perspective information that we need to infer, hold in mind and later use would lead to a varied memory load embedded in everyday communication. In the director task, as the director is only able to mention objects from the common ground, it is plausible that participants are mainly keeping track of the objects in the common ground as the set of potential referents. However, the number of objects in the common ground (and thus the potential memory load) varied across previous studies. More critically, studies on younger children tended to use simple shelf arrays whereas studies on older children and adults tended to use more complex shelf arrays. For example, Nadig and Sedivy (2002) on children aged between 5 to 6 years used 2x2 shelves with one object in each slot, and Nilsen and Graham (2009) on children aged from 3 to 5 years used 3 x 3 shelves with one object in each of the four corners of the shelf. In both studies, one out of four objects was occluded in the privileged ground and the other three objects were in the common ground. However, many studies on older children, adolescents and adults used 4x4 shelves (Apperly et al., 2010; Cane et al., 2016; Keysar et al., 2000, 2003) and even 5x5 shelves (Epley et al., 2004) all with at least two objects

in the privileged ground and with a total number of seven to nine objects presented on the shelf. This variation in the number of items in the common ground and privileged ground across existing studies may account for some variations in children's and adults' perspective-taking performance and its relationship with working memory. The present study systematically manipulated this factor in the director task.

In the current study, we investigated the role of memory in 8- and 10-year-old children's perspective-taking performance during language comprehension. We chose 8- and 10-year-old children as our subjects on two bases. First, on almost any contemporary account, children in this age range are believed to have the full range of basic ToM concepts. In particular, the Level-1 perspective-taking necessary for the director task is typically observed in children as young as 2.5 years (Flavell, Everett, Croft, & Flavell, 1981) or even in infants (Luo & Baillargeon, 2007). The second is that previous studies (Dumontheil et al., 2010; Symeonidou, Dumontheil, Chow, & Breheny, 2016; Wang et al., 2016) have shown robust age-related changes in performance on the director task across this age range.

Two key memory-related manipulations were employed in our study. *Firstly*, children were assigned to two different versions of the director task: a hidden perspective version and a visible perspective version. The hidden perspective version and the visible perspective version were designed to be minimally different, with the only difference being: in the hidden perspective version, the visual cues to the director's perspective were removed after an initial 5000ms-preview-time, thus children had to infer and hold in mind what was in the common ground and/or in the privileged ground until the point of need; whereas in the visible perspective version, the visual cues still remained available after the 5000ms preview time, thus the relevant perspective information could be inferred efficiently at the point of need. The preview time was fixed at 5000ms as pilot work had determined that this gave enough encoding

time for most children to achieve adequate levels of accuracy, providing room for us to observe effects of the other factors we were manipulating. *Secondly*, we also systematically manipulated the relative size of common ground and privileged ground in the director task. In Experiment 1, the number of objects visible to both participants and the director was systematically manipulated from 3, 5 to 7, and the number of objects occluded from the director was held constant as 2, leaving the overall number of displayed objects varying from 5, 7 to 9. In Experiment 2, the overall number of objects was held constant as 12, and the number of objects in the common ground was systematically manipulated from 3, 5, 7 to 9, leaving the number of object occluded from to the director to vary from 9, 7, 5 to 3. The number of objects in common ground vs privileged ground was manipulated in this way following pilot work and previous studies using the director task (Wang, Cane, Ferguson, Frisson, & Apperly, submitted; Zhao, Wang, & Apperly, under review). The number of objects in common ground only started from 3 because this was the minimum number of objects that allowed for two experimental instructions using different nouns.

A study with similar manipulation has been conducted on healthy adults (Zhao et al., under review). When adult participants were required to retain information about the director's perspective and to integrate this encoded information with the director's verbal instructions at the point of need, they were more prone to commit egocentric errors, and their degrees of egocentrism were independently affected by the number of items in the common ground, and also by their working memory capacity. Based on these results from adults, we hypothesized that children would be more prone to make egocentric errors in the hidden perspective version than in the visible perspective version of the director task. Secondly, we predicted that when the size of common ground was greater, participants would need to devote more cognitive resources to managing the increasing size of common ground, and would thus become more prone to egocentric errors and slower to respond. Thirdly, we expected 10-year-olds to make

fewer egocentric errors than 8-year-olds on the experimental trials of the director task. Critically, if 10-year-olds were indeed better perspective takers than 8-year-olds, then it would be of interest to examine whether 10-year-olds and 8-year-olds were disproportionally affected by the memory loads from the hidden perspective version of the task and the increasing size of common ground. If a factor influences egocentrism less in older children than in younger children, then this factor could account for the developmental improvement in children's abilities to use a speaker's perspective information during communication. Specifically, the observation of interaction effects between age and memory-related factors (i.e., task version, common ground size) in children's degree of egocentrism would be positive evidence for such an account.

Experiment 1

Methods

Participants

Thirty-nine 8-year-old children (mean age 8.17 years, age range 7.54 to 8.71 years, 18 males and 21 females) and fifty-seven 10-year-old children (mean age 10.13 years, age range 9.53 to 10.79 years, 28 males and 29 females) from three primary schools located in Birmingham and neighboring area took part in the study. Children in each age group in each school participated in our study individually by the order of an alphabetical class list, and were assigned to either the hidden perspective version or the visible perspective version alternately. Nineteen 8-year-olds took part in the hidden perspective version; another 20 took part in the visible perspective version. Twenty-nine 10-year-olds took part in the hidden perspective version, and another 28 took part in the visible perspective version. There was no difference between the ages of children assigned to the two task versions within each age group, $p = .546$ for the 8-year-old group and .751 for 10-year-old group respectively.

Design and procedure

A 2 x 2 x 2 x 3 mixed design was constructed with age (8- and 10-year-olds) and task version (hidden perspective and visible perspective version) as two between-subject factors, and condition (experimental and control conditions) and magnitude of common ground (3, 5 and 7 common ground objects, referred to as "magnitude") as two within-subject factors

Our computer-based referential communication task was introduced as an engaging computer game to children. Each child took part in the experiment individually. Children were greeted warmly, instructed step-by-step using a PowerPoint presentation and then tested individually outside of their classrooms by the experimenter. The whole experiment, consisting of an instruction phase, a practice phase and a test phase, lasted for 15-20 minutes for each individual. The experiment was presented on a 15.6" Samsung laptop. The practice and test trials were presented with Experiment Builder (SR Research Ltd, Mississauga, Ontario, Canada). Accuracy and response time data were automatically recorded by the Experiment Builder software.

Instruction phase. In the instruction phase, a 4 x 4 shelf image was presented on the computer screen. A number of objects were placed on the shelf. A female director, who was introduced to children as Sally, was standing behind the shelf and facing the participants. Some slots were blocked by green squared backgrounds from Sally's perspective. The experimenter explained to children that Sally could not see the objects in the blocked slots, because she was standing on the other side of the shelf and the green backgrounds were blocking her view. An image of the back of the self was shown to children in order to highlight Sally's limited perspective. Children were then asked five check questions about whether Sally could see certain objects. We included 3 objects from open slots and 2 objects from blocked shots into

the check questions in order to make sure that children fully understood Sally's limited perspective. Children were only allowed to carry on if they answered all five questions correctly, otherwise the experimenter would again explain how Sally's perspective was limited. If children correctly answered the check questions, they were further instructed that if Sally could not see objects in the blocked slots, then she could not know about those objects, and therefore would not ask them to move any of those objects. Children's understanding of the instructions was checked by another five questions on whether Sally could know about certain objects (3 from open slots and 2 from blocked slots), and three additional questions on whether Sally could ask them to move certain objects (2 from open slots and 1 from blocked slots). Every participating child managed to answer all these questions correctly.

After checking children's understanding of Sally's limited knowledge of the shelf, one example image of the shelf (see the left image in Figure 1) was shown to children. Children assigned to the hidden perspective version were given the instruction that "you have 5 seconds to remember which objects Sally does and doesn't know about. After 5 seconds, all of the slots will be blocked by a green background". In contrast, children assigned to the visible perspective version were told that "you have 5 seconds to look at what object Sally does and doesn't know about. After 5 seconds, Sally will start to give you some instructions." After five seconds, children in the hidden perspective version were shown the shelf image with a covering green background (the right image in Figure 1) and those in the visible perspective version were shown the original image (the left image in Figure 1).

Each child was first given an example instruction "nudge the ball one slot up", and was shown by experimenter how to drag the ball one slot up and then drop it using the computer mouse. They were then asked to practice and carry out the same "drag and drop" action. Children were then given the next instruction "nudge the big hat one slot down" and were prompted to point out the two hats that Sally knew about (one big and one small). If children

wrongly pointed to a hat in the blocked slot, the experimenter would point to the image and say, "look here, Sally doesn't know about this hat, because this one is in the green slot and it's blocked off". The same instruction would then be repeated until children got it right. After identifying the two hats Sally knew about, children were asked by the question "so which one is the big hat that Sally is talking about". If children responded correctly, the experimenter would ask children the reason why they chose that hat, to make sure that the correct response was not driven by unexpected strategies. If children responded incorrectly, the experimenter would explicitly tell children about the way in which Sally's limited perspective constrains reference, by saying "Sally couldn't see this red hat, so she does not know that this red hat is on the shelf, and therefore she can't be talking about this red hat. Instead, Sally must be talking about this blue hat, which she can see." At the end of the instruction phase, another shelf image with a different array of objects was presented to children as a comprehension check. A critical instruction "nudge the big bowling pin one slot right" was delivered. Children who failed the comprehension check and performed at floor level were excluded prior to the analysis (which will be explained in details in the results section).

Figure 1. An example image of the shelf presented in the instruction phase. In the hidden perspective version, the left image was presented during the 5000ms preview time, and the

right image was presented after it. In the visible perspective version, the left image was the image presented both during and after the 5000ms preview time.

Test phase. Children undertook either the hidden perspective or the visible perspective task version they were assigned to. For each task version, two practice shelf images were first presented with 3 instructions each. Twenty-four test shelf images were then presented with 2-3 instructions each. Twelve shelf images corresponded to the experimental condition, and the other 12 shelf images corresponded to the control condition. In the experimental condition, the object which best fitted a critical instruction from the participants' point of view was always in the participants' privileged ground, thus made it a "distractor" from the target referent. Participants had to use the information about the perspective difference between the director and participants to select the correct referents. Control trial images were minor adaptations of each corresponding experimental trial image, whereby the competing object in the participants' privileged ground (i.e., distractor) was replaced by an irrelevant object, which did not compete with the target referent as a potential referent. As illustrated in Figure 2, the pink balloon in the privileged ground in the experimental condition was replaced by the broccoli when in the control condition. Therefore, in the control condition, the use of the director's perspective was not a prerequisite for correctly selecting the target referent. The shelf image in the control condition was presented at least 6 shelf images apart (which equals to at least 12 instructions apart) from its corresponding shelf image in the experimental condition. Twenty-four shelf images were grouped into 4 blocks, each block was mixed of experimental and control trials. Participants could take breaks between blocks.

Experimental condition **Control condition**

Figure 2. Example images for an experimental condition trial (left) and its corresponding control condition trial (right).

The shelf image in the experimental condition and the shelf image in its corresponding control condition shared the same set of 2-3 verbal instructions, one of which was a critical instruction, and the others were fillers. The position of the critical instruction varied between the first and third in the sequence. Each instruction was made of individually pre-recorded words in order to prevent participants from using co-articulation to identify an object prior to the onset of the noun. The structure of the critical instructions was "nudge the [scalar adjective] [noun] one slot [directional word]" (e.g., nudge the small ball one slot up). The structure of the fillers was either "nudge the [scalar/normal adjective] [noun] one slot [directional word]" or "nudge the [noun] one slot [directional word]". Directional words included "left" "right" "up" and "down". Children were told that the directional words "left" and "right" referred to their own left and right sides, and labels indicating children's left and right sides were presented across the top of the laptop screen during their participation. For critical instructions, only "up" and "down" directional words were used.

Children were instructed to respond as accurately and as quickly as possible. If children did not respond within 8500ms after the onset of the adjective or noun, then the trial would timeout, bringing up the next instruction or the next shelf image.

Results

We performed both a comprehension check (i.e. whether children responded correctly to a critical question at the end of instruction phase), and a floor performance check (i.e., whether children responded correctly on more than 2 trials out of 12 trials) on participating children's performance. The purpose of adopting these two criteria together for exclusion was to exclude children whose low performance was highly likely due to the lack of understanding of the current experiment instructions rather than egocentrism. For children who failed to pass the comprehension check question in the instruction phase, it was unclear whether this failure occurred only accidently or it truly reflected insufficient understanding of the task instructions. For children who failed to perform significantly above floor, it was unclear whether their low performance was the true reflection of their insufficient understanding of the task instructions, or was induced by the cognitively demanding nature of the current experimental manipulation. Therefore, we used both criteria to exclude participants. Two 8-year-olds (who were not included into the 39 eight-year-old participating children reported in the participants section) were excluded prior the analysis due to both the failure to pass the comprehension check and the failure to perform significantly above floor.

Proportion egocentric errors

Only data from critical trials were entered into our analysis. A response was considered as correct when children selected a target object which was visible to both the director and

themselves. An egocentric error referred to the selection of the distractor in the privileged ground rather than the target object in the common ground. A non-egocentric error referred to the selection of other objects or spaces other than the target and the distractor. The causes of the non-egocentric errors (which constituted 5.92% of the data from the hidden perspective version in 8-year-olds, 7.92% of the data from the visible perspective version in 8-year-olds, 4.17% of the data from the hidden perspective version in 10-year-olds, and 3.72% of the data from the visible perspective version in 10-year-olds) were difficult to interpret and were not the central interest of the current study, therefore non-egocentric errors were excluded prior to the following analysis. Trials with response timeout (1.82%) were also excluded from the analysis.

The proportion egocentric errors in the experimental condition and the control condition were calculated separately for each magnitude condition for each participant. In the experimental condition, the proportion egocentric errors of individual participants ranged from 0 to 1 in a given magnitude condition and from 0 to .917 overall. In the control condition, the mean proportion egocentric errors for each age group in each task version were all below .005. Due to the floor level of proportion egocentric errors in the control condition and the unequal variance between the experimental and control conditions, it was very questionable to include the factor of condition (experimental and control) into an omnibus analysis. Therefore, a 3-way mixed ANOVA was conducted for the experimental condition only, with age (8-year-old, 10-year-old) and task version (hidden, visible) as two between-subject factors, and common ground magnitude (3,5,7) as a within-subject factor. The pattern observed here in the 2 x 2 x 3 ANOVA also held for the analysis of proportion egocentric errors when we included condition (experimental, control) as a factor in a 2 x 2 x 2 x 3 mixed ANOVA.

There was a significant main effect of common ground magnitude, $F(2, 184) = 14.026$, $p < .001$, $\eta p^2 = .132$, with a significant linear trend only, $F(1, 92) = 21.568$, $p < .001$, $\eta p^2 = .190$

(the quadratic trend was not significant, p = .116). This showed that the rates of egocentric errors increased proportionally as the common ground size grew. A significant main effect of age was observed, $F(1, 92) = 9.314$, $p = .003$, $\eta p^2 = .092$, with 10-year-olds outperforming 8-year-olds. A significant main effect of task version was also observed, $F(1, 92) = 4.133$, $p = .045$, $\eta p^2 = .043$, with children committing higher rates of egocentric errors in the hidden perspective version than in the visible perspective version. A significant interaction between task and age was observed, $F(1, 92) = 7.434$, $p = .008$, $\eta p^2 = .075$, and also a significant interaction between magnitude and age, $F(2, 184) = 3.627$, $p = .029$, $\eta p^2 = .038$. There was also a marginally significant interaction between magnitude and task version, $F(2, 184) = 2.740$, $p = .058$, $\eta p^2 = .031$. Most importantly, a significant 3-way interaction of magnitude x task version x age was observed, $F(2, 184) = 7.740$, $p = .001$, $\eta p^2 = .078$ (see Figure 3).

To follow up the significant 3-way interaction of magnitude x task version x age, two 2-way mixed ANOVAs were conducted for 8-year-olds and 10-year-olds separately, with task version and magnitude as the two factors. For 8-year-olds, there was a significant main effect of magnitude, $F(2, 74) = 10.179$, $p < .001$, $\eta p^2 = .216$, with a significant linear trend only, $F(1, 37) = 16.118$, $p < .001$, $\eta p^2 = .303$ (quadratic trend, $p = .444$). This showed that the proportion egocentric errors 8-year-olds committed increased proportionally with the increasing common ground size. A significant main effect of task version was also found, $F(1, 37) = 8.771$, $p = .005$, $\eta p^2 = .192$, as 8-year-olds made more egocentric errors in the hidden perspective version compared to the visible perspective version. Critically, a significant interaction between magnitude and task version was also found, $F(2, 74) = 6.442$, $p = .003$, $\eta p^2 = .148$. Two sets of one-way repeated measures ANOVAs with magnitude as the within-subject factor were conducted for hidden and visible perspective versions, respectively. For the hidden perspective version, a main effect of magnitude was observed, $F(2, 36) = 13.420$, $p < .001$, $\eta p^2 = .427$, with a significant linear trend only, $F(1, 18) = 16.753$, $p = .001$, $\eta p^2 = .482$ (quadratic trend, $p = .195$).

For the visible perspective version, only a marginally significant main effect of magnitude was observed, $F(2, 38) = 2.473$, $p = .098$, $\eta p^2 = .115$.

For 10-year-olds, a 2-way mixed ANOVA was also conducted with task version and magnitude as factors. There was a significant main effect of magnitude, $F(2, 110) = 3.252$, $p = .042$, $\eta p^2 = .056$, with a marginally significant linear trend, $F(1, 55) = 3.778$, $p = .057$, $\eta p^2 = .064$ (quadratic trend, $p = .120$). No main effect of task version was observed, $F(1, 55) = .315$, $p = .577$, $\eta p^2 = .006$. No significant interaction was found between task version and magnitude, $F(2, 110) = 1.245$, $p = .292$, $\eta p^2 = .022$.

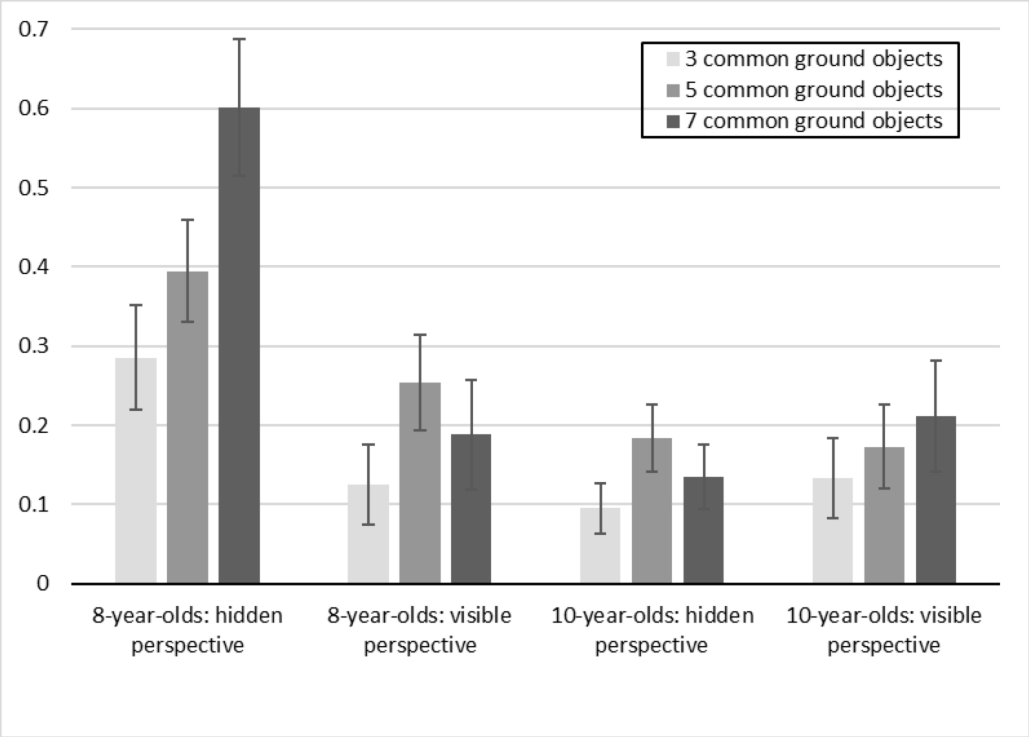

Figure 3. Proportions of egocentric errors on experimental trials in Experiment 1. Error bars show standard errors.

Response time

Only trials with correct responses were included in the response time analysis. As the error rates among the experimental trials were uneven among different task versions in different age groups, the numbers of data points across conditions varied. The response time data therefore could not be interpreted with confidence, as it was likely for the results to be inflated by the variance induced by insufficient data points. For this reason, we only reported the descriptive statistics for the response time (see Table 1).

Table 1: Means and standard errors of response time (ms) for 8- and 10-year-olds in Experiment 1.

Age	8-year-olds								10-year-olds							
Task	Hidden perspective				Visible perspective				Hidden perspective				Visible perspective			
Condition	Experimental		Control		Experimental		Control		Experimental		Control		Experimental		Control	
Magnitud	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
3-CG	3815	246	3409	138	3609	138	3314	128	3376	142	3325	96	3370	128	3412	116
5-CG	3763	185	3587	136	3810	151	3588	128	3524	162	3464	103	3824	156	3528	121
7-CG	4078	239	4002	219	4132	209	4052	179	3980	147	3734	125	4174	176	3587	96

Summary

Overall, children performed more egocentrically when they were under the enforced memory load in the hidden perspective version, and when the common ground size was greater. As expected, 10-year-olds committed fewer egocentric errors compared to 8-year-olds. Critically, a 3-way interaction was observed between the effects of task version, common ground size and age on egocentrism. In 8-year-olds, we found a systematic linear effect of common ground size on egocentrism in the hidden perspective version but not in the visible perspective version; while for 10-year-olds, there was limited evidence suggesting that common ground size had a systematic linear effect on their degree of egocentrism, and no difference was found between the hidden perspective and visible perspective versions. This interaction with age demonstrated that younger children were disproportionally affected by the

common ground size when they were under the enforced memory load than when such memory load was absent. The result pattern from Figure 3 was also consistent in suggesting that the main effect of task version, age and the size of common ground was mainly driven by the distinctive performance of 8-year-old children in the hidden perspective version. Altogether, this suggests that our memory manipulations might have partially accounted for the age-related improvements between 8- and 10-year-olds' perspective-taking performance.

However, two confounding variables were associated with the experimental design in Experiment 1. The first was that although we were manipulating the number of objects in the common ground, the number of objects in the privileged ground was held constant as 2, which left the total number of objects to increase with the size of common ground. The greater the common ground size, the greater the whole object array size, and the more complex the object array. Frick et al. (2014) found that the visual complexity of an array of objects influenced 5- to 8-year-old children's visuo-spatial perspective-taking performance, so it was unclear from Experiment 1 whether 8-year-old children's degrees of egocentrism in the hidden perspective version varied as a function of the size of common ground as we intended to examine, or alternatively as a function of the overall array size or visual complexity. Secondly, the size of privileged ground was held constant as 2 in quantity. If children in the hidden perspective version were being strategic, then they could choose to only encode the two objects in the privileged ground regardless of the number of objects presented in the common ground. Therefore, the cognitive load of encoding and holding in mind relevant perspective information could have been greatly reduced and kept invariable across different magnitude conditions. Given that only 8-year-old children in the hidden perspective version were affected by the increasing common ground size, it is plausible that the 8-year-olds attempted to encode the objects in the common ground or the whole object array, whereas 10-year-olds may have

identified this efficient encoding strategy and encoded the objects in the privileged ground instead.

In Experiment 2, we aimed to disentangle the effects of common ground size from the confounding effect of object array size and to rule out the use of potential memory strategy. We systematically varied the number of objects in the common ground between 3, 5, 7 and 9 whilst holding the total number of objects constant as 12, leaving the number of objects in the privileged ground to vary between 9, 7, 5 and 3.

Experiment 2

Methods

Participants

Sixty-two 8-year-old children (mean age 7.91 years, age range 7.17 to 8.81 years, 30 males and 32 females) and 71 10-year-old children (mean age 9.97 years, age range 9.18 to 10.81 years, 36 males and 35 females) from three primary schools located in the Birmingham area took part in the study. Children in each age group in each school participated in our study individually following the order of an alphabetical class list, and were assigned to either the hidden perspective version or the visible perspective version alternatively. Twenty-nine 8-year-olds took part in the hidden perspective version; another 33 took part in the visible perspective version. Thirty-four 10-year-olds took part in the hidden perspective version, and another 37 took part in the visible perspective version. There was no difference between the ages of children assigned to the hidden perspective version and those assigned to the visible perspective version within each age group, p =.380 for 8-year-olds, and p = .725 for 10-year-olds. Six additional 8-year-olds and 5 additional 10-year-olds were excluded prior the analysis due to

both their failure to pass a comprehension check at the end of the instruction phase and also the failure to perform significantly above floor level.

Design and procedure

A 2 x 2 x 2 x 4 mixed design was constructed with age (8- and 10-year-olds) and task version (hidden perspective and visible perspective versions) as between-subject factors, and condition (experimental, control) and the magnitude of common ground (3, 5, 7, and 9 common ground objects, referred to as "magnitude") as within-subject factors.

The design and procedures were identical to that of Experiment 1 with the following exceptions. All shelf images contained 12 objects in total, while the number of objects in the common ground varied between 3, 5, 7, and 9, thus leaving the number of objects in the privileged ground to vary between 9, 7, 5, and 3. A total of 32 shelf images were presented. Sixteen shelf images were used for the experimental condition with 4 shelf images for each magnitude condition, and the other 16 shelf images were for the control condition also with 4 for each magnitude condition.[1] Each shelf image was accompanied by 2 verbal instructions, one of which was a critical instruction, and the other instruction was a filler instruction. The

[1] In the hidden perspective version, a technical error occurred in one of the control trials, whereby an image corresponding to its matching experimental condition was presented in the 5000ms preview time before a correct image onset and remained onscreen during the verbal instructions. Therefore, the data from this control trial in both the hidden perspective and visible perspective versions were excluded prior to the analysis, in order to make fair comparisons between children's performance in these two task versions.

critical instruction could occur as either the first or the second instruction in a sequence. Thirty-two shelf images were grouped into 4 blocks, and each block had 8 shelf images, mixed of experimental and control trials. The shelf image in the control condition was presented at least 8 shelf images apart from its corresponding shelf image in the experimental condition. One practice block consisting of 3 shelf images was presented prior to the test phase, each shelf image was accompanied by 2 verbal instructions.

Results

Proportion of egocentric errors

Non-egocentric errors (which constituted 10.46% of the data from the hidden perspective version in 8-year-olds, 11.14% of the data from the visible perspective version in 8-year-olds, 9.11% of the data from the hidden perspective version in 10-year-olds, and 7.85% of the data from the visible perspective version in 10-year-olds) and response timeouts (2.79% overall) were excluded prior to the analysis. In the experimental condition, the proportion egocentric errors of individual participants ranged from 0 to 1 in a given magnitude condition and from 0 to .854 overall. In the control condition, the mean proportion egocentric errors for each age group in each magnitude condition of each task version were all below .011.

The data from the experimental condition were submitted to a 2 x 2 x 4 mixed ANOVA with age and task version as two between-subject factors, and magnitude of common ground (3, 5, 7, 9 common ground objects) as a within-subject factor. There was a significant main effect of the magnitude of common ground, $F(3, 387) = 14.980$, $p < .001$, $\eta p^2 = .104$, with a significant linear trend only, $F(1, 129) = 38.001$, $p < .001$, $\eta p^2 = .228$ (quadratic trend, $p = .374$, and cubic trends, $p = .227$). Proportion egocentric errors increased in a systematic linear way with the increase of the size common ground. A significant main effect of task version was also observed, $F(1, 129) = 24.587$, $p < .001$, $\eta p^2 = .160$, with participants in hidden perspective

version committing more egocentric errors than those in the visible perspective version. A marginally significant main effect of age was also observed, $F(1, 129) = 3.489$, $p = .064$, $\eta p^2 = .026$. In general, 10-year-olds showed a tendency to outperform 8-year-olds, but this tendency did not reach statistical significance. A marginally significant interaction between age and task version was also observed, $F(1, 129) = 3.194$, $p = .076$, $\eta p^2 = .024$. No other significant 2-way interaction was found, *p*s > .100. The 3-way interaction was also non-significant, $F(3, 387) = 1.699$, $p = .167$, $\eta p^2 = .013$. The results can be seen in Figure 4[2].

[2] Experiment 2 was first run with just 92 of these children. Analysis of data from this sample yielded the same results pattern to the results we presented in this paper, i.e., significant main effects of task version and common ground size, marginally significant effect of age, and marginally significant interaction between age and common ground. To check whether the marginally significant effects in our initial sample were due to insufficient power, we extended the sample size by 50% to gain the full sample reported here, but this did not change the pattern of results. The marginally significant effects seemed to be rather robust as they remained marginal after the sample size was enlarged. Since this additional data collection was informed by our original analyses, a downward correction of the acceptable p value might have been warranted to account for the increased risk of Type 1 errors. However, we note that such a correction would not alter the inferences we are drawing here.

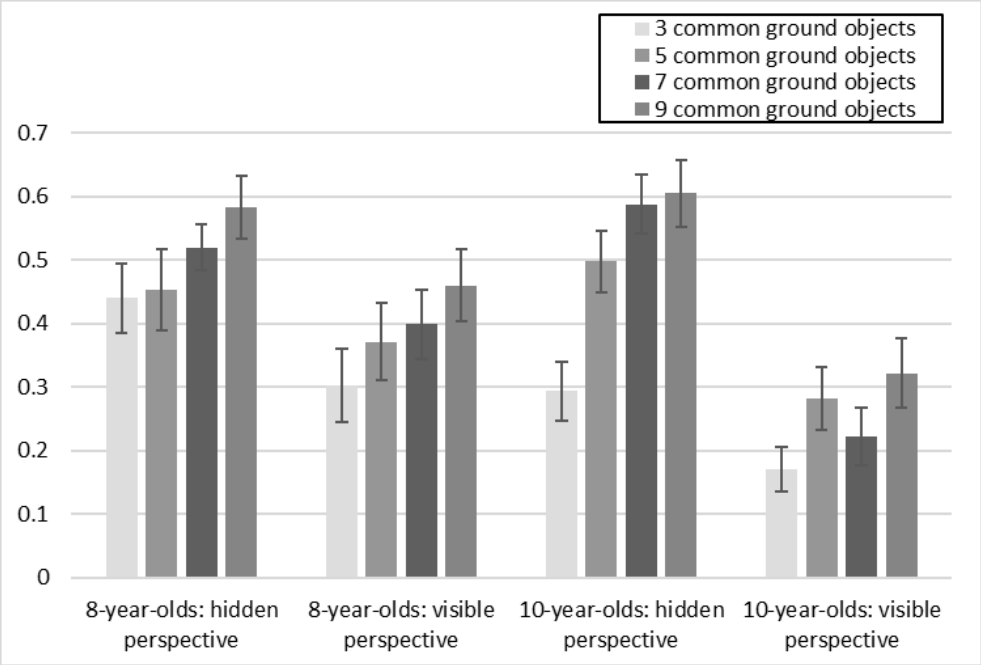

Figure 4. Proportions of egocentric errors on experimental trials in Experiment 2. Error bars show standard errors.

Given the importance of these results to our conclusions, Bayesian analyses were conducted using JASP 0.8.0.1 (JASP Team, 2016) in order to quantify the extent to which data support a model without the Age factor against any model with this variable. We included Magnitude, Task Version and their interaction as nuisance variables, i.e. these variables were included in all models including the null model. The Bayes factor (BF_{01}) represents the degree to which the data are more likely to favour the null model with only the nuisance variable, compared to the models with both the nuisance variables and any effect of age, compared to the null model. Results showed that data are 1.071 times more likely to have occurred under the null model than the model with the main effect of age, which suggested that there was no strong favour towards either of the two models. Likewise, data are 1.120 times more like to have occurred under the null model than the model with the main effect of Age and interaction of Age x Version. Data are also 0.902 times more like to have occurred under the null model than the model with the main effect of Age and interaction of Magnitude x Version. Except for

these three models, the inclusion of other effects all led to higher BF_{01}S, which were larger than 6, indicating that data was much more likely to have occurred under the null model than under these more complex models. From the Bayesian analyses, it was unclear whether the marginally significant main effect of age and marginally significant interaction of Age x Version observed in the normal ANOVA represented any meaningful effects. Therefore, these marginally significant effects need to be interpreted with caution.

Table 2. Bayesian Repeated Measures ANOVA table

Models	P(M)	P(M|data)	BF_M	BF_{01}	Error %
Null model (incl. Magnitude, Version, subject)	0.091	0.224	2.885	1.000	n/a
Age	0.091	0.209	2.642	1.071	2.617
Age + Age*Magnitude	0.091	0.021	0.212	10.780	3.551
Age + Age*Version	0.091	0.200	2.499	1.120	3.086
Age + Age*Magnitude + Age*Version	0.091	0.020	0.199	11.465	4.908
Magnitude*Version	0.091	0.037	0.384	6.052	2.332
Age + Magnitude*Version	0.091	0.248	3.304	0.902	85.286
Age + Age*Magnitude + Magnitude*Version	0.091	0.003	0.034	65.180	2.993
Age + Age*Version + Magnitude*Version	0.091	0.034	0.351	6.611	6.609
Age + Age*Magnitude + Age*Version + Magnitude*Version	0.091	0.003	0.033	68.935	4.082
Age + Age*Magnitude + Age*Version + Magnitude *Version + Age*Magnitude*Version	0.091	9.717e -4	0.010	230.431	3.690

Note: All models include Magnitude, Version, subject.

Response time

As for Experiment 1, only the descriptive statistics of the response time for Experiment 2 were reported in Table 3.

Table 3. Means and standard errors of response time (ms) for 8 and 10-year-olds in Experiment 2.

Age	8-year-olds								10-year-olds							
Task	Hidden perspective				Visible perspective				Hidden perspective				Visible perspective			
Condition	Experimental		Control		Experimental		Control		Experimental		Control		Experimental		Control	
Magnitude	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
3 CG	3702	151	4119	130	3853	116	4221	168	3333	119	3871	126	3980	127	3852	124
5 CG	4217	220	4030	155	4262	182	3981	167	4023	204	3702	107	4105	131	3850	100
7 CG	3980	176	3971	181	4114	179	3689	137	3767	154	3645	112	4145	116	3632	93
9 CG	4194	148	4133	190	4203	173	4075	132	4122	158	4148	132	4245	166	3855	82

Summary

Experiment 2 provided convergent evidence with Experiment 1 on the significant roles of the number of items in common ground and the need to retain these objects in memory while interpreting the director's instructions. When the total number of objects was kept constant, the size of common ground still systematically affected children's perspective-taking performance: the greater the common ground size, the more egocentric errors children committed. Moreover, we found that children were more prone to select the distractor when they were under the enforced memory load than when such memory load was not enforced. Ten-year-olds showed a tendency to be more successful at identifying the target referent in the common ground compared to 8-year-olds, but the difference was only marginally significant. Moreover, we observed a marginally significant interaction between age and task version, which was mostly driven by 10-year-olds performing less egocentrically than 8-year-olds in the visible perspective version, but equally egocentrically in the hidden perspective version. Detailed discussion of the implications of these findings will follow in the next section.

General discussion

The aim of the present study was to examine whether memory factors could explain the variations in children's perspective-taking performance. What was novel in this study was that we adapted a widely-used referential communication task (i.e., director task) to capture the memory demand of holding in mind someone's perspective, which is commonly embedded in our everyday communication but was overlooked in the previous literature. Eight- and 10-year-old children were instructed to take a speaker's perspective either under enforced memory load

(i.e., the hidden perspective version) or under reduced memory load (i.e., the visible perspective version). In Experiment 1, the size of common ground was systematically varied with the size of privileged ground held constant as 2; in Experiment 2, the size of common ground was systematically varied with the total number of objects held constant as 12. Consistent with previous studies, egocentric effects were clearly observed in the measure of errors in our study (e.g., Apperly et al., 2010; Dumontheil et al., 2010; Keysar et al., 2003; Nilsen & Graham, 2009). The current study had three key objectives: (a) to evaluate the effect of holding in mind a speaker's perspective implemented in our newly-adapted director task on children's degree of egocentrism, (b) to assess the effect of common ground size on children's degree of egocentrism, (c) to examine whether 10-year-old children would perform less egocentrically compared to 8-year-olds in our experimental setting, and whether memory factors would explain this age-related reduction in egocentrism. The current findings will be discussed in further details below in relation to these three objectives.

Our study was the first attempt to directly measure the cost of holding in mind a speaker's perspective on participants' performance in the director task by removing the visual cues to the director's perspective after a brief preview time. Our findings from both experiments support the conclusion that children were more egocentric in the hidden perspective version than in the visible perspective version. The specific cognitive demands of each task can be decomposed as follows (Apperly et al., 2010). In the hidden perspective version, participants had to encode, retain and later integrate the director's perspective with her instructions in order to successfully resolve the reference. Failure to include the target referent into the memory record of potential common ground referents would lead to failure to correctly recognise the target referent. However, in the visible perspective version, participants could identify the director's perspective and integrate it with her verbal message at the point of need, but it was not necessary to store relevant perspective information for any longer than it took to infer and

integrate perspective information. The only difference between the two task versions was that there was an extra memory load of remembering and retaining encoded perspective information in the hidden perspective version. Therefore, it seems reasonable to conclude that the increase in egocentric errors arose from the difficulty children faced in managing this extra memory load in the hidden perspective version. This finding corresponds to the everyday situations in which communication requires us to *remember* who knows what, rather than simply *figure out* who knows what at the precise time we need this information. Our study operationalises a way to incorporate this memory demand which is frequently and naturally present in our everyday communication into the widely-used director task, and results confirm that this memory load could partially explain children's variations in perspective use in different real-life situations.

Secondly, our study systematically manipulated the size of common ground in two different ways. In Experiment 1, we manipulated the number of objects in the common ground between 3, 5 and 7 whilst keeping the number of objects in the privileged ground constant as 2. In Experiment 2, we manipulated the number of objects in the common ground between 3, 5, 7 and 9 whilst keeping the overall number of objects as 12. Two experiments consistently found that both 8- and 10-year-olds were less successful at accommodating a speakers' limited perspective in communication when there was a greater amount of common ground information. This effect of increasing common ground size on egocentrism is in line with a previous study showing the same effect of common ground size on adults' egocentrism under the enforced memory load of holding in mind a speaker's perspective (Zhao et al., under review). While common ground size co-varied with the total number of objects in Experiment 1, this was not the case in Experiment 2. Frick et al.'s study (2014) found that visual complexity of the layouts of object array had a detrimental effect on 5- to 8-year-olds' perspective-taking performance when their own perspective differed from another person. Our findings have advanced knowledge in this field by providing evidence that even when two object arrays were

of the same degree of visual complexity, the size of common ground could still affect children's degree of egocentrism. These findings provide support for the long assumed notion that listeners work on the basis of common ground to interpret a speaker's message (Clark & Marshall, 1978, 1981). The findings are consistent with Nadig and Sedivy (2002)'s findings that children do show considerable sensitivity to common ground information in referential communication, although our study is not designed to examine the time-course regarding this sensitivity to common ground. The findings also accord with Barr (2008)'s claim that people do anticipate items in common ground to be referred to, but also suffer from egocentric interference when they need to integrate linguistic input with perspective information. More importantly, our findings provide evidence about how the cognitive demand of managing common ground size might constrain people's ability to use another's perspective to guide communication.

Thirdly, we sought evidence for age-related improvement in children's ability to adopt a speaker's perspective in referential communication, and the possible cognitive factors that drove this age-related improvement. Results from Experiment 1 demonstrated that 10-year-olds outperformed 8-year-olds in the director task in the proportion of egocentric errors they committed. This finding is consistent with a number of previous studies suggesting that children's ability to use other people's mental states continues to grow over time after they have acquired the basic theory of mind competence (Dumontheil et al., 2010; Epley et al., 2004; Symeonidou et al., 2016), especially in accord with Wang et al. (2016) showing that 10-year-olds committed significantly fewer egocentric errors than 8-year-olds in the director task. Experiment 1 also witnessed a significant 3-way interaction among factors of task version, common ground size and age. The performance of 8-year-olds in the hidden perspective version was disproportionally affected by the increasing common ground size, compared to 8-year-olds in the visible perspective version and 10-year-olds in both versions of the task.

However, as discussed previously, this significant 3-way interaction might have been driven by older children's improved ability of identifying and employing the appropriate memory strategies when they were under the extra memory load of remembering the director's perspective, rather than older children's improved memory capacity of correctly remembering a larger amount of perspective information. This speculation is consistent with the suggestion that one key mechanism of the development of visuo-spatial working memory is that older children are better at recognising and adopting possible processing strategies (see Pickering (2001) for a review paper). After controlling for the possibility of being strategic in only encoding the objects in privileged ground, Experiment 2 found that 10-year-olds committed only marginally fewer egocentric errors compared with 8-year-olds. Moreover, only a marginally significant interaction was found between the factors of age and task version in Experiment 2. However, Bayesian analyses showed that data did not strongly favour either the null model or the model with effects of Age or/and Age x Version. Overall, the findings from Experiment 1 and Experiment 2 together provided some evidence for an age-related development in children's ability to accommodate a speaker's limited perspective during referential communication. While 10-year-olds performed significantly less egocentrically than 8-year-olds in Experiment 1, this age-related effect was marginally significant in Experiment 2. As our study only tested 8- and 10-year-old children, it is possible that this limited age range may have obscured the age-related development we intended to find. It would be sensible for future research to have a larger age-range when it comes to examining age-related development in perspective use.

Critically, although common ground size had a systematic effect on children's degree of egocentrism in both Experiment 1 and Experiment 2, no interaction involving common ground size was found in Experiment 2 after controlling for the aforementioned confounding variable. This finding may appear to be surprising, as previous studies have found that children's

successful performance on theory of mind tasks draws upon their executive functions including working memory and inhibitory control (Carlson, Moses, & Breton, 2002; Carlson, Moses, & Claxton, 2004; Davis & Pratt, 1995; Frye, Zelazo, & Palfai, 1995; Lagattuta, Sayfan, & Blattman, 2010; Lagattuta, Sayfan, & Harvey, 2014). Therefore, it seems plausible to assume that the increasing size of common ground might influence younger children's perspective-taking performance more strongly by placing an additional demand on their rather limited executive function capacities. However, our results from Experiment 2, though observing main effects of age and common ground size, provide no clear evidence for such an interaction between age and common ground size. This finding, though surprising, is consistent with Wang et al. (2016) who found that 8- and 10-year-olds were equally influenced by the complexity of the instructions they had to integrate with the director's limited perspective. These findings suggest that we cannot take it for granted that the cognitive factors that affect children's perspective-taking performance are also those that contribute to the process of children's age-related improvement in perspective use.

A reviewer suggested that participants' higher egocentric error rates in the experimental condition compared to the control condition might be due to the distractors being more visually salient than the control object (e.g., the biggest pink balloon on the image on Page 14 was somehow visually more salient than the green broccoli). The visual salience (e.g., size and colour) varied across different images used in our experiment. Importantly, for this to account for our effects it would have to be that the distractors on experimental trials were systematically more salient than on control trials (on average). While we did not pre-test stimuli for salience, nothing in our design should have led to a systematic bias for more salient distractors on experimental trials, and inspection of the actual stimuli suggests that experimental trial distractors did not seem to be systematically more salient than the irrelevant objects which replaced the distractors in the corresponding control condition. The images presented in our

experiment are available upon request. Nonetheless, future studies using this experimental paradigm could consider creating two sets of stimuli by balancing the visual salience of the distractor and the irrelevant object for each shelf image, e.g., by swapping the colours of the distractor and the irrelevant object.

To conclude, the present study incorporated into the commonly-used director task two types of memory demands which are frequently present in our everyday communication: the memory demand to hold in mind a speaker's perspective, and the amount of common ground information to keep track of. We examined the effects of these two memory demands on 8- and 10-year-old children's perspective-taking performance during language comprehension. Findings demonstrated that 8- and 10-year-olds' rates of egocentrism were modulated both by the demand of holding in mind a speaker's perspective, and by the varied amount of common ground information they needed to manage. While 10-year-olds were significantly less egocentric than 8-year-olds in Experiment 1, this effect was marginal in Experiment 2. No clear evidence was found for the memory-related factors to account for age-related improvement in children's perspective-taking performance. Nonetheless, the current findings imply that children will be more successful at accommodating a speaker's limited perspective when they can easily access information about what is shared between them and the speaker, compared to when they have to remember and later retrieve this information from their memory. The findings also suggest that the increments in the number of potential referents shared between listeners and speakers may cause children to be more egocentric when they need to use a speaker's limited perspective to identify a correct referent. This was supported by a previous finding that young children found it harder to learn how to request a target referent in a training task for language production when an object array size was larger (Matthews, Butcher, Lieven, & Tomasello, 2012). When talking to children, highlighting what is shared between conversational partners and narrowing down the scope of potential referents would help them

to better avoid falling back to using their own perspective to interpret what other people are referring to.

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. *The Quarterly Journal of Experimental Psychology*, *63*(6), 1201–1217. https://doi.org/10.1080/17470210903281582

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, *9*(1), 7–31. https://doi.org/10.1111/j.2044-835X.1991.tb00859.x

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, *109*(1), 18–40. https://doi.org/10.1016/j.cognition.2008.07.005

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*(5), 382–386.

Blain-Brière, B., Bouchard, C., & Bigras, N. (2014). The role of executive functions in the pragmatic skills of children age 4–5. *Frontiers in Psychology*, *5*.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*(5), 893–900. https://doi.org/10.3758/PBR.16.5.893

Cane, J. E., Ferguson, H. J., & Apperly, I. (2016). Using perspective to resolve reference: the impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Retrieved from https://kar.kent.ac.uk/57114/

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319.

Clark, H. H., & Marshall, C. (1978). Reference Diaries. In *Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing* (pp. 57–63). Stroudsburg, PA, USA: Association for Computational Linguistics. https://doi.org/10.3115/980262.980272

Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In *Elements of Discourse Understanding*. Cambridge University Press.

Dahlgren, S., & Sandberg, A. D. (2008). Referential communication in children with autism spectrum disorder. *Autism*, *12*(4), 335–348.

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology*, *47*(1), 25–31.

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*(2), 331–338. https://doi.org/10.1111/j.1467-7687.2009.00888.x

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

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*(1), 99.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds. *Frontiers in Psychology*, *5*, 386.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*(4), 483–527.

JASP Team, (first). (2016). JASP (Version 0.8.0.1) [Computer software].

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*(1), 32–38.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41. https://doi.org/10.1016/S0010-0277(03)00064-7

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834.

Lagattuta, K. H., Sayfan, L., & Blattman, A. J. (2010). Forgetting common ground: Six- to seven-year-olds have an overinterpretive theory of mind. *Developmental Psychology*, *46*(6), 1417–1432. https://doi.org/10.1037/a0021062

Lagattuta, K. H., Sayfan, L., & Harvey, C. (2014). Beliefs about thought probability: Evidence for persistent errors in mindreading and links to executive control. *Child Development*, *85*(2), 659–674.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, *46*(3), 551–556. https://doi.org/10.1016/j.jesp.2009.12.019

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*(3), 489–512.

Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two-and four-year-olds learn to adapt referring expressions to context: effects of distracters and feedback on referential communication. *Topics in Cognitive Science*, *4*(2), 184–210.

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), 1–21.

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603–613.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329–336.

Nilsen, E., & Graham, S. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220–249. https://doi.org/10.1016/j.cogpsych.2008.07.002

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*(2), 125–137. https://doi.org/10.1111/j.2044-835X.1987.tb01048.x

Piaget, J. (1959). *Judgment and reasoning in the child. Paterson, NJ: Littlefield, Adams, & Co.*

Piaget, J., & Inhelder, B. (1956). *The child's concept of space*. Routledge & Paul.

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory*, *9*(4-6), 423–432.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236. https://doi.org/10.1016/j.cognition.2010.08.003

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other

people see. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

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*(2), 199–204.

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.

Wang, J. J., Ali, M., Frisson, S., & Apperly, I. A. (2016). Language complexity modulates 8- and 10-year-olds' success at using their theory of mind abilities in a communication task. *Journal of Experimental Child Psychology*, *149*, 62–71.

Wang, J. J., Cane, J. E., Ferguson, H. J., Frisson, S., & Apperly, I. A. (submitted). *Why is it sometimes difficult to take account of a speaker's perspective?*

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 655–684.

Wu, S., & Keysar, B. (2007). The Effect of Culture on Perspective Taking. *Psychological Science*, *18*(7), 600–606. https://doi.org/10.1111/j.1467-9280.2007.01946.x

Zhao, L., Wang, J. J., & Apperly, I. A. (under review). *Remembering what she knows: The costs of holding in mind a speaker's perspective during a communication task.*