Korean Preschoolers' Advanced Inhibitory Control and Its Relation to Other Executive Skills and Mental State Understanding

Seungmi Oh and Charlie Lewis
Lancaster University

This study assessed executive function and mental state understanding in Korean preschoolers. In Experiment 1, forty 3.5- and 4-year-old Koreans showed ceiling performance on inhibition and switching measures, although their performance on working memory and false belief was comparable to that of Western children. Experiment 2 revealed a similar advantage in a sample of seventy-six 3- and 4-year-old Koreans compared with sixty-four age-matched British children. Korean children younger than 3.5 years of age showed ceiling effects on some inhibition measures despite more stringent protocols and the link between executive function and mental state understanding was not as strong as in the British sample. The results raise key questions about the nature and development of the executive system and its relation to social understanding.

Over the past 20 years, increasing attention has been paid to the rapid developments in preschoolers' cognitive and social skills, under the banners of "executive function" and "theory of mind." Executive function can be defined as the set of higher order cognitive skills that are responsible for conscious and effortful control of thought and behavior. It encompasses a range of cognitive activities such as "planning, flexible strategy employment, impulse control, and organized search" (Welsh, Pennington, & Groisser, 1991, p. 132). "Theory of mind," broadly speaking, refers to the understanding of people's mental states (Perner, 1991; Wellman, 1990). In recent years, the link between these two areas of cognitive functioning has become central to the exploration of each individual area (e.g., Schneider, Schumann-Hengsteler, & Sodian, 2005). In this article, we explore whether links found in Western cultures can be generalized to a non-Western setting.

The multifaceted nature of executive function has been demonstrated in studies with adults (Miyake et al., 2000), school-age children (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003), and preschoolers (Hughes, 1998a). These studies investigated three core components: inhibition, working memory, and switching. One aim of the present study was to explore the relation among these three constituent executive skills in preschoolers. Some data on adults suggest that the best model fitting the three component skills is one in which a separateness between each is maintained but they are linked by an overall latent executive function variable (Miyake et al., 2000). However, such networks of executive skills have been identified in children and adults only in some Western cultures.

A second aim is to explore the links between these executive skills and mental state understanding. Numerous studies in the West have found that preschool children go through a parallel development in these two areas of cognitive functioning (e.g., Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Frye, Zelazo, & Palfai, 1995; Hala, Hug, & Henderson, 2003; Hughes, 1998a; Perner, Lang, & Kloo, 2002). These findings have led to much debate about the underlying mechanism between the two skills. Some have asserted that understanding of internal states facilitates self-control (e.g., Perner & Lang, 2000). According to this view, some degree of mental state understanding is itself a precursor to later developments in the executive system.

In contrast, others have argued that advances in different aspects of executive control are necessary or even sufficient for the development of mind awareness. There are three major positions within this perspective: the working memory, cognitive complexity and control (CCC), and inhibition accounts. The first holds that working memory skills provide a key ingredient for false belief understanding as the individual

---

has to hold in mind the contrasting aspects of events in relation to current reality (Davis & Pratt, 1995; Gordon & Olson, 1998). However, the correlation between working memory and mental state understanding does not hold when factors like children's general intelligence are taken into account (Carlson et al., 2002).

The CCC theory proposed by Frye and Zelazo argues that false belief is one example of a problem of understanding rules embedded within an overall framework. Support for this view comes from studies using a card sort procedure, the dimensional change card sort (DCCS), derived from the Wisconsin Card Sorting Test, which correlates highly with false belief (Frye et al., 1995). Frye et al. argue that this association is caused by the fact that the same conditional "if… if… then…" rule structure can be applied to false belief and switching tasks, like the DCCS (for a review, see Frye, 1999).

The nature of the link between false belief and switching is still being debated for two reasons. First, there is disagreement about what the card sort test actually measures. The CCC account of Frye and Zelazo has been questioned by those who argue that it is not grasping rules in an all-or-none fashion that is crucial. Rather, children fail the DCCS because it establishes an attentional inertia for them toward following the first rule—in other words, it is a task of inhibitory control (Kirkham, Cruess, & Diamond, 2003; Towse, Redbond, Houston-Price, & Cook, 2000). Thus, when children are shown the correct action before they sort the cards (Towse et al., 2000) or if they are asked to give each card a label before sorting (Kirkham et al., 2003), then more 3-year-olds now sort correctly. Happaney and Zelazo (2003) interpreted the results of Kirkham et al. (2003) in terms of the CCC theory, as Zelazo and Jacques (1997) had done with reference to earlier inhibition tasks, and the debate has yet to be resolved.

Second, there is continuing discussion on the nature of any causal link between switching and "theory of mind." Perner (Perner & Lang, 2000; Perner, Stummer, & Lang, 1999) argues that the DCCS and false belief tasks need not be explained by the same embedded rules and, indeed, that within a CCC framework false belief could easily be passed by employing two simple, dissociated rules: *if* I am looking for the chocolate *then* there; *if* Maxi is looking for the chocolate *then* here. While this point is not universally agreed upon, the evidence on whether mental state understanding or a grasp of embedded rules emerges first is equivocal. There is some evidence for functional interdependence. For example, Kloo and Perner (2003) found that training of executive skills as measured by the DCCS (Frye et al., 1995) enhanced children's performance on the false belief

task and training of false belief understanding led to gains in children's DCCS performance.

A third perspective about the effects of executive control on theory-of-mind development holds that inhibitory control is a key to developing social understanding (Carlson & Moses, 2001; Carlson et al., 2002; Carlson, Moses, & Hix, 1998; Hala et al., 2003; Russell, 1996). Although in the 1990s inhibition was assumed to be a dissociable construct, there is now debate concerning whether this is the case (Simpson & Riggs, 2005a) and whether inhibition alone can explain false-belief reasoning. In the studies done by Carlson and Moses, in particular, correlations between inhibitory control and mental state understanding (usually false belief but also deception) remain significant even when other factors like language or working memory are taken into account. One type of task, conflict inhibition, is particularly important because, they argue it contains a strong working memory load. The associations between such conflict inhibition tasks and false beliefs hold even when taking other skills, like planning tasks, into account (Carlson, Moses, & Claxton, 2004). Thus, Moses and Carlson (2004) assert that children need to develop capacities both for reflecting upon their thought and behavior and for overcoming a salient but irrelevant action tendency in order to appreciate their own and others' perspectives on the world. The claim that executive skills are required for false belief understanding is supported by longitudinal (Hughes, 1998b) and microgenetic (Flynn, O'Malley, & Wood, 2004) data.

Whereas the evidence on the relation between inhibitory control and mental state understanding is growing, "the nature of that linkage remains to be specified" (Carlson et al., 2004, p. 300). The same holds for associations between false belief and both switching and working memory. Such links raise important questions about the basis of our understanding of these connections. The research we cite is founded upon a number of assumptions. First, the executive system develops in ways that are consistent with the data on older children and adults. Second, the executive system is relatively culture free and children in a diversity of societies will develop individual executive skills at the same rate and with the same overall relationships. Third, the links between executive and "theory-of-mind" capacities have some functional dependence and thus, by implication, will be relatively consistent over all cultures. It is these issues that this research set out to explore.

The studies on the relation between the two abilities in preschool children have been conducted mostly in the United States, Austria, Britain, and Canada. They have recently been extended to Africa

and Latin America (Chasiotis, Kiessling, Winter, & Hofer, 2006). Research on this topic in other cultures might contribute to a fuller understanding of the nature of the relationships described previously. This study looked at Korean preschool children's executive functioning and theory of mind. Far Eastern countries such as China, Japan, and Korea have a long tradition of Confucianism which emphasizes respect for elders and obedience to authority figures including teachers and parents. Although these countries have been exposed to Western influences, with their stress on individualism and horizontal interpersonal relationships, Confucian practices have been preserved. Asian cultures, for example, place greater stress upon academic training and less emphasis on the value of play in children, even among "Asian American" families (Parmar, Harkness, & Super, 2004). However, it is clear that Korean children in the United States are exposed to a mixture of the two cultures (Farver & Lee-Shin, 2000). Thus, this research was conducted in Korea.

There are good grounds for assuming that Korea is a natural laboratory for exploring the executive skills of preschoolers. Indeed, it is conceivable that its cultural practices may promote young children's capacity for self-regulation. Within Korea, traditional values are strongly reflected in preschool education despite "child-centered" influences from Western philosophies (Kwon, 2002). The national curriculum for all preschools of some 1,500 pages contains specific themes, topics, and materials for education for 3- to 5-year-olds (French & Song, 1998). Specifically, although teachers in nurseries or kindergartens share a belief that children should be intrinsically motivated and preschool education should be integrated rather than divided into several subjects, in practice, they enforce discipline, teach separate subjects (e.g., music, physical activities, language, and science experiments), and prefer whole-class activities to child-directed or small-group activities (Kwon, 2002).

The observations of French and Song (1998) of Korean preschool settings were that 3- and 4-year-olds spent up to 1 hr per day being instructed by a teacher, with only 3 or 4 in a group of 40 showing any signs of restlessness. The teacher offered information about a chosen topic. They presented a case study of one 16-min segment about bats that the teacher did not know was to be recorded. This conveyed 115 pieces of information on the topic. The children were asked questions as a group and their transcript showed that they often offered information to which the teacher replied. If the children became very vocal or distracted, the teacher said "Please look at . . . " to which the children chorused "your teacher!" Every 5 – 10 min, the teacher and children sang the attention-management song, "My eyes are for watching, ears are for listening . . . " By providing information and guiding the children's attention, the teacher aimed to foster both attention skills and executive processes: "The teacher models and provides support for children to practice these 'higher-order' metacognitive and comprehension monitoring skills as she asks questions that require inference, interpretation, and other cognitive processes . . . Opportunities for listening also offer children to practice and develop attention-management skills" (French & Song, 1998, p. 426).

Given the preceding observations of children in Korean preschool settings, we predicted that Korean children would display better performance on inhibitory measures than Western children. Our third aim was to confirm this prediction. If this pattern were found, then it would allow us to explore whether such skills would enhance performance in other executive skills (i.e., working memory and switching) and to test the unity of these skills found in earlier studies (Lehto et al., 2003; Miyake et al., 2000). We administered a well-established theory-of-mind procedure—the false belief task—and a variety of executive measures to Korean preschool children. As the evidence suggests that the executive system has multiple components, we examined children's performance across the range of skills described previously—inhibitory control, working memory, and switching—to examine if the relationships found in Western samples generalize to a culture in which self-control is accentuated.

Experiment 1

The general aim was to explore whether the relationships across individual executive skills and between them and mental state understanding in the West would be echoed in Korea. Nine executive and one theory-of-mind tasks were administered. Given the recent data on the relationship between inhibitory control and mental state understanding, our initial starting point was the nature of this link. The construct "inhibitory control" was divided into conflict and delay measures according to the classification proposed by Carlson and Moses (2001). The central focus was on conflict inhibition because Moses and Carlson (2004) claim that this requires key skills for false belief understanding: not just the inhibition of prepotent but irrelevant responses, but also the generation of novel responses conflicting with such prepotent ones. According to Carlson and Moses, delay inhibition measures simply require suppression of

salient responses and impose few working memory demands. We included these as a point of comparison with conflict inhibition.

At the same time, we were mindful of the debates over the nature of executive tests that we discussed in the Introduction. The classic test of switching, the DCCS, has been regarded by some as a measure of inhibition of attentional inertia. In this study, we included the DCCS and two measures of inhibitory control, the day/night and Luria's hand game, in part because this allowed us to compare performance on these tasks and their relation to false belief performance. In some studies, the correlation between the DCCS and false belief has been found to be notably stronger than the link between false belief and both Luria's hand game (Lang & Perner, 2002) and the go/no-go task (Perner et al., 2002).

One problem in discerning the relationships between switching and inhibitory control concerns the fact that the former tends to be measured by only one task. In order to explore the link further, we devised a task, fruit animal alternation, in which children have to apply two rules over alternate trials, naming the fruit on one card and the animal on the next where each card contains one item of each. In piloting, we became aware of the possible demands of other executive skills, particularly the working memory demands of recalling which rule to apply in each trial. Nevertheless, we felt it was worth including this test to attempt to broaden the repertoire of switching tasks available for testing preschoolers.

Research in Korean schools is not a commonplace activity. This had two implications. First, because the national curriculum imposes demands upon the teachers, giving little free time during which children can be spared to participate, we were required to do all testing in one session, despite the numerous tests. Second, we were mindful of the possibility that children not used to being tested might not demonstrate their optimal skills. As a result, we took pains to make the tasks as child friendly as we could. Thus, in between trials of the inhibitory control tasks, the experimenter occasionally deviated from the script of set procedures and said "good." When the rules were changed in two tasks (the DCCS and Luria's hand game), children were informed about the significance of the change in rule.

Method

Participants

Forty Korean children of 42–55 months (22 girls; $M = 47.8$, $SD = 3.94$) were recruited from one nursery school located in a middle-class area of Seoul. The majority of parents came from professional (i.e., upper middle class) families, though with a social mix representative of nursery schools as government statistics show that 43% of children attend centers of preschool education (McMullen et al., 2005). For the statistical analyses, participants were divided into two age groups: older 3-year-olds ($M = 44.35$ months, $SD = 1.60$, range $= 42–47$) and younger 4-year-olds ($M = 51.25$ months, $SD = 2.07$, range $= 48–55$).

Measures

Verbal ability. The vocabulary subscale of the Korean–Wechsler Preschool and Primary Scale of Intelligence (K–WPPSI; Park, Kwak, & Park, 1995) was administered. For the first three questions, children were shown the pictures of a cat, a tree, and a key and asked to label them. They gained 1 point for each question if they correctly answered. For the remaining 22 questions, children were required to define the words presented by the experimenter. For each question, they were scored according to the following scoring system: 2 points = *full understanding*, 1 point = *partial understanding*, and 0 point = *no understanding*. The perfect score was 47.

Executive Functioning Battery

Conflict inhibition measures. In the day/night task (Gerstadt, Hong, & Diamond, 1994), the experimenter presented eight cards depicting a picture of the sun and eight cards depicting a picture of the moon with some stars around it in a pseudorandom order. The children were asked to say "night" in response to sun cards and to say "day" in response to moon cards. Two practice trials were given and, following Gerstadt et al. (1994), if they performed successfully, these were counted as the first two test trials and the test phase continued from that point. If participants did not succeed on the first two practice trials, the experimenter reminded them of the rule and started again. The dependent measure was the number of correct responses. The experimenter sometimes said "well done" or "good" to keep the children engaged in the test procedure.

In Luria's hand game (Hughes, 1996), children were instructed to imitate one of two hand actions by the experimenter, a fist or point. Following six imitation trials (three fist and three finger trials in random order), children were required to produce conflicting hand actions. When the experimenter made a fist, the children had to point a finger and vice versa. Seven fist trials and seven finger trials

were presented in a pseudorandom order. Before the conflict phase began, the experimenter said, "If you copy me, you will lose this new game." The dependent measure was the number of correct trials.

Delay inhibition measures. In the tower-building task (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996), children were asked to take turns with the experimenter in building a tower. Twenty wooden blocks were used, with 10 allocated to the experimenter and children. The experimenter deliberately waited before placing each of her blocks until children explicitly signaled they were giving a turn to the experimenter. Participants obtained 1 point when they gave the experimenter a due turn. If they gave all due turns to the experimenter, they would get 10 points. Following Kochanska et al. (1996), children gained 1 point for carefully arranging the tower in good order to prevent it from collapsing and lost a point when knocking it down. If children waited for 10 s after placing their due block, they were credited with 1 point even when they did not show any explicit signals of giving a turn to the experimenter.

Following Espy et al. (2004), the self-control task was administered at the end of testing because a present served as "a thank-you gift" for each child's participation in the experiment. The experimenter put an attractively wrapped gift in front of each child and instructed him/her not to touch the gift until she finished her work. The experimenter went through her test sheets while secretly observing the child. The dependent measure was the latency to touch the present (with a maximum of 150 s).

Working memory measures. Following Gordon and Olson's (1998) procedure, in the finger tapping and labeling trials, children were asked to continue to tap a finger on the table and at the same time to label three objects when the experimenter lifted them. The three objects used were a toy rabbit, a pencil, and a teaspoon. If children performed incorrectly (i.e., by ceasing to tap as they labeled an object), the experimenter repeated the demonstration and asked them to do the dual task. A maximum of two demonstration trials and two test trials were administered. Children were regarded as having passed if they performed correctly on at least one of two test trials.

The eight boxes scrambled test was adapted from the six boxes task by Diamond, Prevor, Callender, and Druin (1997). Children were required to find eight stickers one by one hidden in eight boxes that differed from one another by color and patterns decorated on their lids. Whenever each child opened a box, the experimenter re-covered it, placed a big screen between him/her and the boxes, and scrambled them while they counted to 10 together to fill the delay. The test continued until children found all the stickers or reached the wrong boxes seven times in a row. The dependent measure was the total number of reaches.

The backward word span task was employed following Carlson et al.'s (2002) procedure. The experimenter asked children to repeat a list of words in reverse order. Practice trials were administered with two-word lists. If they performed correctly in the practice trial, the experimenter proceeded to the test phase. In the test trials, the size of the word list increased from two to four with successful performance, with a minor difference from previous studies of the new word adding to the old list instead of replacing it. Span was the maximum number of words that children recalled in reverse order without making a mistake. The score ranged between 1 and 4. Following Carlson et al., children who failed the two-word list obtained 1 point and those who achieved the four-word list obtained 4 points.

Switching measures. The DCCS test was based on the procedure of the standard version of the DCCS used by Frye et al. (1995). Children were shown two model cards that were each attached to a back wall of two sorting trays. One model card depicted a red car and the other depicted a blue star. Then the experimenter presented two types of sorting cards to children. The sorting cards depicted either a blue car or a red star, and thus, they did not match a model card on both dimensions (color and shape). The experimenter asked children to sort six cards according to one dimension (either color or shape and this dimension was counterbalanced). During these preswitch trials, feedback was given. If children sorted six cards correctly according to the first sorting dimension, the sorting dimension was altered. Unlike in the preswitch phase, the experimenter did not demonstrate sorting the cards in the postswitch phase. Following Kirkham et al. (2003), on each trial the experimenter randomly selected a card and labeled only the relevant dimension by saying, "Here's a blue one (a car)," "Here's a red one (a star)," and asked children, "Where does this go in the color (shape) game?" No feedback was given during the postswitch trials. Before the postswitch phase began, the experimenter said, "If you sort the cards as before, you will lose this new game." If children correctly sorted five of six cards, they were considered to have successfully switched the sorting dimension.

As described in the Introduction to this experiment, the fruit animal alternation test was a newly developed measure. Ten laminated cards included a pair of pictures, each with a different item of fruit on the left and a different animal on the right. The cards were shown to participants one by one. On each trial,

children were required to label either the fruit or animal item in alternation. Five demonstration and five practice trials were administered with the cards faceup. Then the test phase began. During the test trials, the experimenter turned the card over so that children could not see the previously presented card. As 10 cards were presented and the alternation trial began from the second card, there were nine alternation trials. The score was the number of perseverative errors (whenever children persisted in naming items of the same category, the experimenter corrected it and resumed testing).

False Belief Measure

The procedure of the deceptive box test was based upon Gopnik and Astington (1988). Children were shown a closed chocolate box and asked a question "What do you think is inside?" Then the experimenter opened the box and took a pen out of it. After participants identified the true object inside the box, the experimenter put the pen back in the box and closed it. Then children were asked about their own prior false belief and the other's current false belief about the chocolate box. Finally, they were asked the reality control question, "What is actually inside the box?" Credits for both self and other questions were given to children only in the case of giving a correct answer on this reality control question.

Procedure

Children were tested individually in an empty classroom of the nursery school. The vocabulary scale

was always presented first and self-control was always administered last. The rest of the tasks were administered in a fixed order, but a Latin square determined a different starting point for successive children: day/night, eight boxes scrambled, DCCS, deceptive box, tower building, backward word span, fruit animal alternation, Luria's hand game, and finger tapping and labeling. The first author tested all the children. It took 40–50 min to administer all the tests to one child.

Results and Discussion

Verbal Ability

The relatively young 4-year-old children ($M = 15.00$, $SD = 5.65$) performed significantly better than the relatively old 3-year-old children ($M = 11.75$, $SD = 2.55$) on the vocabulary scale of K–WPPSI, according to the Mann–Whitney U test, $z = 2.03$, $p < .05$. The levels of performance were slightly above those of the standardization sample, which is to be expected given the social mix of the sample.

Executive Functioning Performance

A summary of children's performance on the nine executive measures is reported in Table 1. The preliminary analyses did not reveal age effects in any of the nine measures.

Conflict inhibition performance. As Table 1 shows, on the conflict inhibition measures, children showed ceiling or near-ceiling performance. This pattern is very different from the findings on Western

Table 1
Descriptive Statistics for Executive Performance as a Function of Age

	Older 3 years	Younger 4 years	Range
Conflict Inhibition Measures			
Day/night (no. correct trials)	14.65 (2.11)	14.65 (1.95)	9–16 [16]
Hand game (no. correct trials)	14 (0.00)	13.7 (0.66)	12–14 [14]
Delay Inhibition Measures			
Tower building (point score)	9.65 (1.90)	8.90 (1.94)	3–14
Self-control (latency)	135.90 (38.41)	123.40 (50.66)	6–150 [150]
Working Memory Measures			
Finger T&L (percentage passing)	100	100	
Eight boxes scrambled (no. reaches)	11.90 (2.90)	10.85 (2.62)	8–20 [8]
Backward word span (span score)	1.65 (0.93)	2.25 (1.07)	1–4 [4]
Switching Measures			
DCCS (percentage passing)	95	100	
Fruit animal alternation (no. errors)	3.3 (1.89)	2.15 (1.98)	0–8 [0]

Note. Standard deviations are shown in parentheses. The scores that reflect the possible highest performance on each measure are shown in square brackets in the far right column. In the tower-building task, the highest score for giving all due blocks is 10 excluding additional credits. DCCS = dimensional change card sort; T&L = tapping and labeling.

preschoolers, who invariably face difficulties in overcoming the habitual tendency to provide novel responses on these measures. On the day/night task, the mean percentages of correct trials for the two age groups were exactly the same at 91.6%. These data contrast with those from the study done by Gerstadt et al. (1994) in which the mean percentages of correct trials for the older 3- and younger 4-year-olds were 71.5% and 68.8%, respectively.

On Luria's hand game most children performed at ceiling. Of 40 children, 36 (90%) did not make any mistake over the 14 conflict trials.

Delay inhibition performance. The majority of children in the tower-building task gave most due turns (10 blocks) to the experimenter.

On the self-control measure, the majority of children (72.5%) were at ceiling (delaying for 150 s). Meanwhile, one child touched the present when only 6 s passed and 15% of children touched it within 50 s.

These patterns of performance on both delay tasks are hard to compare directly with those of previous studies because Kochanska et al. (1996) used a total score based upon a larger number of trials on the tower-building task and the version of self-control in the study done by Espy et al. (2004) was used with 2- to 5-year-old children.

Working memory performance. On finger tapping and labeling, all the children successfully performed the dual task apart from one 3.5-year-old child who refused to tap a finger on the table due to apparent shyness. Indeed, most participants needed only one trial to engage successfully in the dual activities. Although the age ranges are not exactly corresponding, the Canadian profiles (Gordon & Olson, 1998) in which 46% of 3-year-olds and 71% of 4-year-olds passed at least one of two trials are in clear contrast with the current data. Having administered this test, we doubt whether this measure taps working memory skills. Children are required to resist the automatic tendency to stop tapping a finger when labeling and to produce two conflicting actions simultaneously. Thus, this task may well measure conflict inhibition skills and Korean children's ceiling performance makes sense considering their equivalent performance on day/night and Luria's hand game.

On the eight boxes scrambled task, the children made on average three errors in finding the eight stickers. The number of boxes was increased in this study because even 3-year-old children have been found to perform at ceiling on the six boxes scrambled test (e.g., Espy et al., 2004). In the current study, it was observed that most children did not make an error until they retrieved the fifth or sixth sticker, suggest-

ing that Korean preschoolers would also perform at ceiling on the six boxes scrambled task.

On the backward word span task, it appeared that the processing capacities of these Korean preschoolers were comparable with those of similarly aged American children. American preschoolers' backward word span (1.58 for 3-year-olds and 2.21 words for 4-year-olds) in the study by Carlson et al. (2002) was very close to that in these Korean preschoolers.

Switching performance. The DCCS data show that except for only one 3.5-year-old child, all the children correctly sorted at least five of six cards according to the postswitch sorting criterion. This ceiling performance, particularly by 3-year-olds, is in sharp contrast with earlier findings that 3-year-olds usually have difficulty in switching the sorting criterion on the standard version of this test (e.g., Frye et al., 1995; Kirkham et al., 2003).

On the fruit animal alternation task, some children committed seven or eight errors across nine alternation trials. It might be the case that they did not understand the rule of this task properly. This consideration led to a slight modification of the procedure in the second experiment to present a more thorough way for children to understand the alternation rule during the practice trials.

False Belief Performance

The two age groups showed similar performance on the deceptive box measure, with fewer than 40% correct in each group. One 4-year-old child did not answer the question "What do you think is inside the box?" and, therefore, was excluded. On the "self" false belief, 30.0% of the younger children and 31.6% of the older ones passed; on the "other" false belief, 35.0% of the younger children and 36.8% of the older ones were correct. This pattern deviates from the norm in Western studies where a shift between 3- and 4-year-olds' performance on the false belief tasks is commonly found. The lack of an age group difference in this study might be due to the narrow age range. Thus, the second experiment looked at the false belief performance by Korean children with a broader age range.

Correlations Among Executive Functioning, False Belief, and Language Measures

In order to explore further the relationships between the measures, a series of Spearman's correlations was conducted, given the negative skew in the performance on some executive tests. Indeed, two

measures (DCCS and finger tapping and labeling) were omitted because of ceiling performance. For convenience, the scores on the eight boxes scrambled and fruit animal alternation tasks were reversed so that a higher score indicated better performance. An overall false belief scale was created (range 0–2) by summing each binary score for the self and other false belief items of the deceptive box task.

Table 2 presents these correlations. It shows that two groups of variables appeared to be related to one another. First, two measures—backward word span and fruit animal alternation—were significantly correlated with each other and these two were significantly related to performance on the vocabulary subtest of K–WPPSI (see Table 2). This set of correlations may confirm our suspicions that the fruit animal alternation task might be more of a working memory test than a measure of switching. Second, some measures of inhibitory control—day/night, Luria's hand game, and self-control—were correlated with tower building but not with each other. Only one measure, backward word span, significantly correlated with the total false belief scores, but given the relationship between the working memory measure and vocabulary, it might be the case that even this relationship is mediated by another factor. Indeed, an exploratory partial (Pearson) correlation (which must be tentative as both measures were positively skewed) showed that the false belief/word span correlation became nonsignificant if vocabulary was taken into account, although false belief did not correlate significantly with vocabulary.

The correlations reveal two apparent differences between these Korean data and those collected in the West. First, the relationships found in the United States (e.g., Carlson et al., 2002) between conflict inhibition and working memory as predictors of false belief were not clearly in evidence in this study. Second, Korean children showed more advanced performance on the inhibitory measures, particularly on the conflict inhibition measures.

Experiment 2

While Experiment 1 appeared to show patterns of executive control that are different from those in the West, a close comparison between a Korean and a Western sample is required. One finding that needs particular attention is the excellent performance on conflict inhibition tasks, given their centrality in recent debate (Moses & Carlson, 2004). To explore this question, we decided to focus in Experiment 2 on these tasks and defer research on delay tasks for future studies. We kept Luria's hand game and the day/night task and included two new procedures that we assumed would not seem as easy for 3- and 4-year-olds. First, we incorporated Luria's tapping test, which appears to provide a slightly greater inhibitory demand (Diamond & Taylor, 1996). We also modified the grass/snow test (Carlson & Moses, 2001) in which children have to point to a sheet of white paper when the experimenter says "grass" and a green paper when the experimenter says "snow." A variant of this task, the blue/red task, provides preschoolers with the more directly contrasting cue of pointing to a blue (red) square of paper when the experimenter says "red" ("blue"). While we were conducting this procedure, Simpson and Riggs (2005b) were administering a similar task. To accommodate these measures, we dropped finger tapping and labeling, in part because children were very successful at this task and

Table 2
Spearman Correlations Among the Vocabulary, Executive, and False Belief Measures

	Conflict IC		Delay IC		Working memory		Switch	False belief
	Day/night	Hand game	Tower building	Self-control	Eight boxes	Word span	Fruit animal alternation	Deceptive box
Vocabulary	.12	−.21	−.08	.18	.24	.58***	.46**	.19
Day/night		.06	.42**	.30	.01	.29	.17	.17
Hand game			.40**	.24	.16	.01	−.12	−.29
Tower building				.35*	.15	.05	.09	.24
Self-control					.28	.16	.17	.14
Eight boxes						.17	−.16	−.01
Word span							.54***	.33*
Fruit animal alternation								.15

Note. IC = inhibitory control.
*$p < .05$. **$p < .01$. ***$p < .001$.

also because it was unclear to us whether it assessed working memory, as suggested by Gordon and Olson (1998), or conflict inhibition as we suspected.

Three further modifications were made to the tests in Experiment 2. First, we extended the mental state test to include both the deceptive box task and the unexpected transfer formats. Second, we changed the procedures in the tasks that were made in Experiment 1 to make testing Korean children more participant friendly, as we sought to rule out the possibility that child-friendly protocols in some procedures might have enhanced children's performance to a ceiling level in Experiment 1. So, in keeping with most published data on executive skills in preschoolers (the exception is the DCCS—to be discussed later), the child-friendly prompts, like ''good'' between trials, were omitted in Experiment 2. Similarly, we dropped the prompt in the DCCS and Luria's hand game (e.g., in the latter ''If you copy me, you will lose this new game'') as the switch to the new rule was made. Third, we were concerned that two tests might not have been completely understood by the children. In the fruit animal alternation task, we extended the pretest phase because we felt that children might have erred because they had not learned the rules. We wanted to examine further whether this task assesses switching, working memory, or a combination of both. In the backward word span, we were concerned that the procedure is not child friendly enough. We thus included a visual–spatial analogue to help children to understand what ''reverse order'' might mean.

The two aims of Experiment 2 were, first, to see whether the patterns of data in Experiment 1 would be replicated in a larger sample of Korean preschoolers with an age range extending across the two-year span of 3–4 years. Second, we included a sample from the United Kingdom, so that we could make comparisons between an Eastern and a Western culture.

Method

Participants

Seventy-six (37 girls) 36- to 59-month-old ($M = 47.2$ months, $SD = 6.5$) Korean preschool children were recruited from two nursery schools and a kindergarten located in an area of Seoul with predominantly professional families. A comparable sample of 64 (35 girls), 36- to 62-month-old ($M = 47.4$ months, $SD = 6.8$) English children, all White, were drawn from a comparable sample in Lancaster, UK. Both samples were predominantly upper middle class but included families from across the social spectrum.

In the analyses, participants were divided into four age groups. They were 3-year-olds (Korean: $M = 39.1$ months, $SD = 1.83$, range = 36–41; English: $M = 38.0$ months, $SD = 1.41$, range = 36–41), 3.5-year-olds (Korean: $M = 44.29$ months, $SD = 1.99$, range = 42–47; English: $M = 44.32$ months, $SD = 1.67$, range = 42–47), 4-year-olds (Korean: $M = 50.23$ months, $SD = 1.63$, range = 48–53; English: $M = 50.67$ months, $SD = 1.80$, range = 48–53), and 4.5-year-olds (Korean: $M = 55.88$ months, $SD = 1.73$, range = 54–59; English: $M = 56.06$ months, $SD = 2.05$, range = 54–62) groups. Each age group consisted of 20, 17, 22, and 17 Korean and 14, 19, 15, and 16 English children, respectively.

Measures

Executive Functioning Battery

Conflict inhibition measures. The day/night task and Luria's hand game procedures were the same as those in Experiment 1 except that the experimenter did not make the comments such as ''good'' or ''well done'' during the former, and children were not warned about losing the game before the conflict trials in the latter.

Two new tasks were included in the battery. Adapting the blue/red test from the grass/snow test devised by Carlson and Moses (2001), children were shown a white sheet of paper on which a piece of blue paper and a piece of red paper were attached side by side at the top. Children were asked to place their one hand on the middle of the white paper. Then they were instructed to point to the blue paper when the experimenter said ''red'' and to point to the red paper when she said ''blue.'' Sixteen test trials (eight blue and eight red trials) were presented in a pseudorandom order. The dependent measure was the number of correct trials.

In Luria's tapping test (following Diamond & Taylor, 1996), children were instructed to tap a wooden dowel on the table once when the experimenter tapped twice and to tap twice when she tapped once. Sixteen test trials were administered (eight trials of each tapping action), presented in a pseudorandom random order. The dependent measure was the number of correct trials.

Working memory measures. In the eight boxes scrambled and backward word span tasks, the procedures were the same as those in Experiment 1, with only a slight modification to the instructions in the latter. As younger children appeared to find the term ''in reverse order'' difficult in the first experiment, the experimenter constructed a horizontal line of small squares of paper in front of participants. She pointed to each square in turn while saying a word on the word list and then asked children to repeat what she

had just said while pointing to the pieces of paper in reverse order. The number of the pieces of paper presented corresponded to the size of the word lists. Unlike in Experiment 1, practice trials were not required because all but two children passed a two-word trial at the start. The span was increased to a seven-word list. The scores ranged between 2 and 7.

Switching measures. The procedure of the DCCS in this experiment was the same as that in Experiment 1 except that the experimenter did not say "If you sort the cards as before, you will lose this new game" right before the postswitch trials.

The fruit animal alternation procedure was different from that in Experiment 1 only in the administration of practice trials. In Experiment 1, five practice trials were administered to all children regardless of their performance. In this experiment, the experimenter proceeded to the test phase only when children correctly performed on four consecutive practice trials in order to ensure that they understood the rule of alternation thoroughly before the test trials began.

False Belief Battery

The procedure in the deceptive box test was the same as that in Experiment 1 except that different materials were used. Instead of a chocolate box, a plaster box was used and the content inside the box was a hair clip instead of a pen. In addition, children were administered the unexpected transfer test. Following the procedure of Baron-Cohen, Leslie, and Frith (1985), the experimenter introduced two doll protagonists—Sally and Anne—and told a story to the children. Sally had a marble. She put her marble in a pouch and left the place. In Sally's absence, her friend Anne took the marble out of the pouch and moved it to a box. Then Sally came back. The experimenter asked the false belief question, "Where will Sally look for her marble?" Three control questions followed: "Where is the marble now?" "Where did Sally put her marble before she left the room?" and "Did Sally see Anne move the marble?" If children pointed to the pouch, they were scored as having passed the false belief question.

Procedure

Given the constraints of working within Korean nurseries, we tested the children in both cultures in a single session that usually took about 35–45 min. Participants were tested individually in empty classrooms. The vocabulary test was administered first. The remaining tasks were administered in the following order, using a Latin square to determine the

child's first test to rule out order effects: Luria's hand game, backward word span, unexpected transfer, DCCS, eight boxes scrambled, blue/red, fruit animal alternation, Luria's tapping test, deceptive box, and day/night.

Results and Discussion

Verbal Ability

Although the vocabulary measure was derived from the same test, the items were different. Thus, the two cultures were explored separately. In the Korean sample, the mean scores for 3-, 3.5-, 4-, and 4.5-year-old groups were 11.35 ($SD = 2.13$), 12.29 ($SD = 2.44$), 14.09 ($SD = 2.07$), and 16.06 ($SD = 2.16$), respectively, and significant age-related differences were found, $F(3, 72) = 16.30, p < .001, \eta_p^2 = .40$. Tukey's tests revealed significant differences between 3- and 4-year-olds and the three younger age groups (3-, 3.5-, and 4-year olds) and the 4.5-year-olds. In the English sample, the means in the four groups were 8.07 ($SD = 1.82$), 9.56 ($SD = 3.29$), 11.73 ($SD = 3.56$) and 13.31 ($SD = 3.72$). There were significant differences, $F(3, 59) = 7.89, p < .001, \eta_p^2 = .29$, and Tukey's tests revealed differences between the 4.5 and two youngest groups and also between the 3- and 4-year-olds.

Executive Functioning Performance

The reliability of the executive tasks was first examined by dividing those with clear binary trials into two sets—the sums of odd numbered and even numbered trials. Correlations between each set were all acceptably high (range: $r = .75 – .94$); therefore, we may assume that these alternate forms show sufficient levels of consistency. Table 3 displays children's performance on the eight executive function measures. It shows that on some tasks, the Korean children matched those in Experiment 1, whereas the performance on others was more variable, perhaps because a wider age range was sampled.

Conflict inhibition performance. As in Experiment 1, a near ceiling effect appeared in the Korean sample for the day/night task. Forty children were at ceiling and a further 16 made only one error across 16 trials. Thus, only 18 children made more than one error (2 refused to participate) and these were spread across the four age groups, Kruskal–Wallis, $\chi^2(df = 3, N = 74) = .69$, *ns*. These data contrast with those from the English children. Only 11 of the 64 were at ceiling, and three made only one error (and additional seven children [six 3-year-olds and one 3.5-year-old] failed to complete the task). As Table 3 shows, there was an

Table 3

Descriptive Statistics for Executive Performance as a Function of Age in the Two Samples

	3-year-olds	3.5-year-olds	4-year-olds	4.5-year-olds	Range
Conflict Inhibition Measures (no. correct trials)					
Day/night, K	14.68 (1.97)	14.94 (1.60)	14.90 (1.67)	15.35 (0.93)	9–16 [16]
Day/night, E	5.63 (4.75)	8.61 (5.19)	12.00 (3.51)	13.19 (3.21)	2–16
Hand game, K	13.35 (1.35)	13.94 (0.24)	13.91 (0.43)	14 (0.00)	10–14 [14]
Hand game, E	6.63 (3.62)	11.63 (2.92)	12.80 (0.94)	13.06 (1.48)	2–14
Blue/red, K	11.65 (3.31)	11.88 (2.06)	13.09 (2.29)	14.82 (1.81)	4–16 [16]
Blue/red, E	6.73 (3.50)	10.00 (5.16)	12.27 (4.79)	15.38 (0.96)	2–16
Tapping test, K	12.55 (1.86)	12.56 (2.34)	14.05 (1.91)	14.59 (2.18)	9–16 [16]
Tapping test, E	5.33 (1.51)	10.72 (3.53)	11.67 (3.92)	12.00 (4.49)	2–16
Working Memory Measures (no. reaches [eight boxes] and span scores [word span])					
Eight boxes, K	10.94 (2.31)	10.65 (2.91)	10.09 (2.27)	10.12 (2.32)	8–17 [8]
Eight boxes, E	15.29 (3.07)	12.89 (2.31)	11.80 (2.24)	12.38 (3.05)	8–23
Word span, K	3.35 (0.99)	3.65 (0.86)	4.32 (1.52)	4.82 (1.01)	2–7 [7]
Word span, E	2.21 (0.58)	3.21 (1.08)	3.53 (0.92)	3.81 (0.91)	1–5
Switching Measures (percentage of passing [DCCS] and no. errors [fruit animal alternation])					
DCCS, K	40	70.6	68.2	100	
DCCS, E	14.3	57.9	86.7	81.3	
Fruit animal, K	3.29 (1.83)	2.71 (1.65)	1.62 (1.53)	0.71 (0.85)	0–6 [0]
Fruit animal, E	7.86 (1.88)	3.05 (2.46)	2.47 (1.41)	1.63 (1.54)	0–9

Note. For each measure, Korean children's performance is indicated in the upper row (identified by a ''K'') and English children's performance is presented in the lower row (identified by an ''E''). The possible maximum performance for each task is indicated in the brackets in the far right column. Standard deviations are presented in parentheses. DCCS = dimensional change card sort.

improvement in performance with age, Kruskal–Wallis, χ^2(df = 3, N = 57) = 15.32, p < .01, but even the 4.5-year-olds were not performing as well as the younger 3-year-old Korean children. Comparisons between each age group showed that the Korean children were performing better than their English peers, Mann–Whitney, z = 3.74, p < .001 at 3; z = 3.65, p < .001 at 3.5; z = 2.75, p < .01 at 4; z = 2.36, p < .05 at 4.5—all two tailed.

As Table 3 shows, the same patterns were in evidence for Luria's hand game. The Korean children performed at ceiling, with 69 children of the 76 (90.8%) successful across 14 trials. The age groups were significantly different because 5 of the 20 children in the youngest age group made one or more errors and all the children older than 4.5 years of age were at ceiling, Kruskal–Wallis, χ^2(df = 3, N = 76) = 8.59, p < .05. This compares with the English sample who showed a similar age progression, Kruskal–Wallis, χ^2(df = 3, N = 55) = 14.46, p < .01, but even the oldest English children were performing below the level of the Korean 3-year-olds. Again, each Korean age group was more successful than its English counterpart, Mann–Whitney, z = 3.81, p < .001 at 3; z = 3.46, p < .01 at 3.5; z = 4.17, p < .001 at 4; z = 3.00, p < .01 at 4.5—all two tailed. Therefore, our concerns that differences between Korean and Western children on

these tasks might be caused inadvertently by subtle differences in their administration can be dispelled.

On the blue/red task, the Korean children were, as expected, not at ceiling and the English children were not at floor. Checks showed that the data for each sample were within normal ranges. A 2 (culture) × 4 (age group) factorial analysis of variance (ANOVA) produced significant main effects for culture, F(1, 128) = 9.74, p < .01, η_p^2 = .07; age group, F(3, 128) = 18.87, p < .001, η_p^2 = .31; and an interaction, F(3, 128) = 3.88, p < .05, η_p^2 = .08. Simple effects analysis comparing cultures in each age group showed that it was in the youngest group that the cultural difference was significantly different, t(29) = 3.88, p < .01, whereas for the other ages, the results were not significant, t(22.54) = 1.43 at 3.5; t(18.40) = .62 at 4; t(24.60) = 1.10 at 4.5—controlling for unequal variances. As Table 3 shows, there were clear developments with age in both cultures, as seen by the larger effect size for age than for culture and the interaction.

On Luria's tapping test, 11 Korean and 9 English children were unable to follow the test procedure and were excluded. Checks revealed that the data for the each subgroup were in the normal range on this measure. A two-way factorial ANOVA revealed main effects for culture, F(1, 112) = 35.81, p < .001, η_p^2 = .24;

age group, $F(3, 112) = 8.51$, $p < .001$, $\eta_p^2 = .19$; and an interaction, $F(3, 112) = 3.15$, $p < .05$, $\eta_p^2 = .08$. The effect of culture was replicated in the comparisons of three of the four age groups, $t(15) = 8.11$, $p < .001$ at 3; $t(18.77) = 2.17$, $p < .05$ at 4; $t(21.42) = 2.09$, $p < .05$ at 4.5—in the latter two unequal variances required adjustment. Only in the 3.5 age group was the difference between the cultures nonsignificant, $t(32) = 1.77$, and this most likely explains the interaction. Tukey's tests following up the main effect for age showed that the 4- and 4.5-year-olds scored more highly than the younger 3-year-olds.

How do these data compare with previous studies? In the study by Diamond and Taylor (1996), slightly older children were studied. At 3.5 years, the mean percentage of correct trials was 64% and that of 4-year-olds was 81% and of 4.5-year-olds was 77%. In the present study, the mean percentages of correct responses of the English 3.5-, 4-, and 4.5-year-old groups were comparable (67%, 72.9%, and 75%, respectively), whereas those in the Korean sample were higher (3.5-, 4-, and 4.5-year-old groups' mean percentages of correct trials were 78.5%, 87.7%, and 91.2%, respectively).

Working Memory Performance. In the eight boxes scrambled task, 2 children (both Korean) refused to continue with the test. Almost half the Korean sample either made no error (19.7%) or opened only one empty box (28.6%) while retrieving the stickers. Thus, the average number of "false reaches" was just over 2 (see Table 3). The English children showed a developmental progression but even the oldest children made on average five false reaches. In a two-way factorial ANOVA, there were main effects for culture, $F(1, 130) = 36.15$, $p < .001$, $\eta_p^2 = .22$, and age group, $F(3, 130) = 4.60$, $p < .01$, $\eta_p^2 = .10$. The interaction, $F(3, 130) = 1.69$, was not significant. Tukey's tests revealed that the younger 4-year-olds made significantly fewer moves than the younger 3-year-olds.

In the backward word span task, the children's span increased in both cultures and the Korean children performed at a higher level than the English children. There were main effects for culture, $F(1, 132) = 22.09$, $p < .001$, $\eta_p^2 = .14$, and age group, $F(3, 132) = 13.27$, $p < .001$, $\eta_p^2 = .23$, but no interaction, $F(3, 132) = .74$. Tukey's tests showed that the 4.5-year-olds' span was significantly longer than the 3-year-olds' and the 3.5-year-olds', and the 4-year-olds' span was significantly longer than the 3-year-olds'.

The contrast between the Korean children's span scores for Experiments 1 and 2 is noteworthy (compare Tables 1 and 3). Although this is most likely attributable to the use of a visual–spatial analogue (pointing to a line of squares in the same or reverse order) in Experiment 2, we have no direct comparison of this procedure with the verbal instructions given in Experiment 1. Further work could directly compare the two means of administering working memory tests with young children, but for the purposes of this analysis, the span task clearly differentiated children, and the age differences suggest that such differences indicate developmental trends.

Switching Performance. There were clear age differences in the performance on the DCCS. The Korean children older than 4.5 years of age were at ceiling, whereas only 40% of those younger than 3.5 years of age showed successful performance. This compared with the figures of 81% and 14% in the comparable English groups. A logistic regression was conducted on success and failure on the DCCS, with age and culture loaded as categorical variables. The overall model was significant, $\chi^2(df = 4) = 33.48$, $p < .001$; Nagelkerke $R^2 = .29$. Culture made no contribution to this result, Wald $\chi^2(df = 1) = 1.61$. Age had an overall effect, Wald $\chi^2(df = 3) = 25.35$, $p < .001$. The SPSS output also compares one level (the younger 3-year-olds) with the other levels on the same factor. The two groups of children older than 4 years of age were significantly different from the youngest group ($p < .05$) and nearly different from the 3.5-year-olds ($p = .09$).

The difference in performance between the two experiments might reflect the fact that children in the first were given a verbal prompt ("If you sort the cards as before, you will lose this new game"). It is not certain whether the slight difference in protocols had an effect, but it is conceivable that it helped some children to shift their attention to the other dimension during the postswitch trials. If that were the case, it might lend support to the argument of Kirkham et al. (2003; see also Towse et al., 2000), discussed in the Introduction, that younger children's difficulty with the DCCS lies in their failure to inhibit the tendency of "attentional inertia." Alternatively, Brooks, Hanauer, Padowska, and Rosman (2003) have shown that preschoolers' performance on sorting tasks may simply vary because of small contextual manipulations. Therefore, casting the card sort into a game (children have to sort into "same" categories or "silly" ones) may enhance performance, whereas adding irrelevant color to the cards leads to more errors. Thus, more systematic manipulations would be required to clarify whether the previous instruction plays a critical role in children's card-sorting performance.

The data from the fruit animal alternation task match those of Experiment 1, although 4 Korean children were dropped from this condition because they failed to grasp the rules of the activity in

the slightly longer warm-up phase. In a two-way factorial ANOVA, there were main effects for culture, $F(1, 128) = 31.99$, $p < .001$, $\eta_p^2 = .20$, and age group, $F(3, 128) = 39.34$, $p < .001$, $\eta_p^2 = .48$, and the interaction, $F(3, 128) = 10.31$, $p < .001$, $\eta_p^2 = .20$, was also significant. To unpack the interaction, the two cultures were compared across each age group. T-test comparisons showed that the two middle groups were not significantly different from one another, $t(34) = .49$ at 3.5 years; $t(34) = 1.69$ at 4 years. The Korean children made fewer errors at 3 years, $t(29) = 6.84$, $p < .001$, and at 4.5 years, $t(23.01) = 2.10$, $p < .05$, after adjustments were made for unequal variances. Thus, as Table 3 shows, the overall cultural difference was accounted for mainly by the differences in the oldest and youngest age groups. In both cultures, there were clear improvements in performance with age.

False Belief Performance

Table 4 summarizes children's performance on the false belief battery. It shows the expected patterns of change in both samples. A series of logistic regressions was conducted on success and failure on each task. As before, age and culture were entered as categorical variables. Each overall model was significant: unexpected transfer, $\chi^2(df = 4) = 29.29$, $p < .001$: Nagelkerke $R^2 = .26$; deceptive box "self," $\chi^2(df = 4) = 28.08$, $p < .001$: Nagelkerke $R^2 = .28$; deceptive box "other," $\chi^2(df = 4) = 22.80$, $p < .001$: Nagelkerke $R^2 = .23$. Culture had no effect in any: unexpected transfer, Wald $\chi^2(df = 1) = .05$; deceptive box "self," Wald $\chi^2(df = 1) = .01$; deceptive box "other," Wald $\chi^2(df = 1) = .24$. Age was significant in all three: unexpected transfer, Wald $\chi^2(df = 3) = 20.70$, $p < .001$; deceptive box "self," Wald $\chi^2(df = 3) = 20.65$, $p < .001$;

deceptive box "other," Wald $\chi^2(df = 3) = 16.41$, $p < .01$. In the unexpected transfer test, all three older groups performed significantly better than the 3-year-olds ($p < .05$), whereas in both items of the deceptive box test, the two groups of 4-year-olds performed better than the younger 3-year-olds ($p < .05$). To test children's competence in each age group in the two cultures, performance in each test was assessed against chance (assuming 50% likelihood of guessing the correct answer in each test and using two-tailed binomial tests). The 3-year-olds were below chance on all three tests ($p < .001$), the 3.5-year-olds were below chance on the unexpected transfer and "self" tests ($p < .02$), whereas only the 4.5-year-olds were above chance on the two deceptive box tests ($p < .05$). Thus, the patterns of performance are comparable with those in most studies (Wellman, Cross, & Watson, 2001).

Relationships Among Executive Functioning, False Belief, and Language Measures

A false belief scale was created by summing the binary scores for the three false belief questions. Table 5 presents the correlation matrix for all the variables in the study, with the exception of Luria's hand game and the day/night task in which there was too little variation in the Korean children's performance for them to be included. Pearson's correlation was used because checks showed that these measures were within normal ranges. As in Experiment 1, the scores on the fruit animal alternation and eight boxes scrambled tasks were reversed so that a higher score indicated better performance. Table 5 shows first that the correlations between the executive tests in the two cultures were slightly different. For a start in the Korean sample, two of the executive measures, the eight boxes scrambled and tapping tasks, showed

Table 4

Percentage of Children in Each Age Group Passing Both False Belief Questions

	3-year-olds	3.5-year-olds	4-year-olds	4.5-year-olds
Korean Sample				
Unexpected transfer	10.0 ($n = 20$)	29.4 ($n = 17$)	31.8 ($n = 22$)	58.8 ($n = 17$)
Deceptive box	($n = 17$)	($n = 15$)	($n = 20$)	($n = 17$)
Self false belief	11.8	26.7	45.0	70.6
Other false belief	17.6	33.3	45.0	76.5
English Sample				
Unexpected transfer	0 ($n = 14$)	21.1 ($n = 19$)	46.7 ($n = 15$)	68.8 ($n = 16$)
Deceptive box	($n = 7$)	($n = 13$)	($n = 15$)	($n = 16$)
Self false belief	0	23.1	53.3	68.8
Other false belief	0	53.8	60.0	68.8

Note. The data for 7 Korean and 13 English children who could not answer the question "What do you think is inside the box?" in the deceptive box task were rated as missing.

Table 5

Pearson Correlations Among the Vocabulary, Executive, and False Belief Measures in the Korean (Above the Diagonal) and English (Below the Diagonal) Samples

	Vocabulary	Blue/red	Tapping test	Eight boxes	Word span	DCCS	Fruit animal	False belief
Vocabulary		.30**	.26*	.08	.31**	.36**	.33**	.43***
Blue/red	.38**		.30* (.22)	.06 (.04)	.42*** (.32*)	.26* (.22†)	.36** (.31*)	.32** (.09)
Tapping test	.19	.42** (.27†)		.10 (.01)	.15 (.05)	.22 (.14)	.09 (.11)	.24† (.05)
Eight boxes	.08	.04 (−.19)	.26† (.15)		.29* (.20)	−.01 (−.25†)	.15 (.04)	.05 (−.17)
Word span	.29*	.23† (−.06)	.29* (.15)	.23† (.08)		.36** (.13)	.54*** (.42**)	.06 (−.21)
DCCS	.31*	.40** (.12)	.36** (.29†)	.26* (.17)	.40** (.14)		.39** (.17)	.40** (.15)
Fruit animal	.43***	.37** (−.15)	.55*** (.38*)	.36** (.29†)	.49*** (.34*)	.52*** (.35*)		.16 (−.17)
False belief	.43**	.43** (.20)	.26† (.14)	.06 (−.19)	.36** (.18)	.46** (.33*)	.47*** (.20)	

Note. Partial correlations controlling for age in months and vocabulary are shown in parentheses. DCCS = dimensional change card sort.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

almost no correlations with the other executive tests. The exception was that the boxes task correlated with the other working memory measure, word span. In the English sample, most of the executive tests correlated with one another. Second, a difference was apparent in the link with false belief. Whereas in the Korean sample, blue/red and the DCCS were associated with false belief, in the English sample, all the executive measures correlated with this social understanding scale, except for a single working memory measure, the eight boxes scrambled task.

Third, in both samples, most of the executive measures were correlated with vocabulary. Given the links with this general language measure, Table 5 also shows the correlations between the tasks with vocabulary scores and age in months partialled out, in parentheses. In the Korean sample, a few of the correlations between the executive measures hold when vocabulary and age are taken into account, showing three-way links between the blue/red, word span, and fruit animal alternation tasks. However, no individual executive measure continued to correlate with false belief once the language scores and age were taken into account. In the English sample, three partial correlations between the executive measures were maintained and all involved correlations with fruit animal alternation: the tapping, word span, and DCCS tasks. In addition, the DCCS task remained a significant correlate with false belief.

Previous studies have found that the relationships between the executive skills and false belief are more discernible when the results of the former are constructed in composite scales (e.g., Carlson & Moses, 2001; Sabbagh, Xu, Carlson, Moses, & Lee, 2006; Tardif, So, & Kaciroti, 2007). We followed this precedent by carrying out two sets of correlations. We performed two because the new measure, fruit animal alternation, was originally constructed as a switching

task, but in both studies, it was highly correlated with word span (see Tables 2 and 5). We standardized each of our executive measures and constructed composite scores of each of the three executive constructs. In the two analyses, we compared the effect of including the fruit animal alternation task within the composite switching versus including it within working memory. In both analyses, the raw correlations were significant, so we can rule out possible Type 1 errors. However, the correlations were stronger when fruit animal alternation was included as a working memory task, so we report this analysis here. This leaves the DCCS as the only switching measure. Table 6 presents the correlations between these composite scores, including total false belief, separately for each culture. It shows that in each culture, the raw correlations between the composite measures were significant, but two differences are noteworthy. First, most of the correlations in the Korean data were less strong than those in the English sample. Second, when language and age were partialled out, these associations were almost completely diluted in the Korean sample. The only significant correlation, between working memory and false belief, was negative. In the English sample, there was greater evidence of links between these measures, with clearly significant partial correlations between the DCCS and both inhibition and false belief.

General Discussion

The Korean children in these experiments appear to show precocious performance in some executive tests but also an interesting variability in performance when compared with the English children. On the boxes task and three of the four measures of inhibitory control (day/night, Luria's hand game, and tapping

test), not only was the difference between the two cultures significant, but on each measure, the younger Korean 3-year-olds were performing above the level of their English counterparts who were almost 5 years old. It is hard to interpret this difference in the working memory test. It could be that the boxes task promotes the same impetuous reaching associated with low inhibitory control. However, the low correlations between the boxes task and the measures of inhibition shown in Tables 2 and 5 militate against such an inference. The clearest finding is that the performance of the Korean children in the inhibition tasks across both experiments shows proficiency that demands attention in this section.

The other results show that Korean children are not 1 year ahead in all executive tasks. Scores on one working memory test, word span; one test of inhibitory control, blue/red; and one measure originally designed as a switching task, fruit animal alternation, showed significant differences between the two cultures but only a slight advantage in the Korean children (see Table 3). Despite ceiling performance in Experiment 1 and an apparent difference in the DCCS between the cultures in Experiment 2, the trend in the latter was not significant, so we cannot conclude that the patterns of success on this measure were different in Korea.

These results raise three issues that will each be addressed in order. First, the data question our assumptions about the nature of the executive system in preschoolers and its relation to the means by which we assess its component skills. Second, they allow us to reflect upon the links between executive function tasks and ''theory of mind'' skills over the 4th and 5th years of life. Third, and most importantly, the precocious inhibitory control skills in Korean children need to be explained.

Nature of the Executive System

We first consider the data on all the executive measures, in light of the issue of whether there are separate components of executive functioning. The raw correlations in Tables 2 and 5 suggest that there were significant associations between a mixture of individual measures that cross familiar boundaries of inhibition, working memory, and switching. Such a pattern would be expected in a model of the executive system in which component skills are distinct yet interdependent, as demonstrated in adults (Miyake et al., 2000) and even older children (Lehto et al., 2003). Yet, there is much in the analyses presented here, which questions such an assumption. First, it is not clear why in both cultures there were no clear associations between the individual tests assessing each construct—the relationships were more piecemeal and straddled the three categories, particularly in the Korean sample (see Table 5). Second, once language and age were taken into account, most of the correlations in Table 5 were no longer significant. Fruit animal alternation seemed to be the only variable in each sample that retained its links with others once the effects of language and age were partialed out. Given that this is a measure designed for this study, further work should explore why this is the case. Third, the patterns of the correlations among the executive components in the two cultures were different. Table 6 shows clear relationships between inhibitory control, working memory, and switching in the English sample, of which only the inhibition-switching link remained after controlling for language and age. Meanwhile, the relationships between these constructs in the Korean children were weaker than those in the English sample and no relationship remained significant after the control factors were held constant. These factors raise questions about whether the executive system is as unified in preschoolers as has been suggested and whether the structure of the executive system is consistent across cultures. The data add to recent critical analyses of a model of executive function with three distinct but unified components (Huizinga, Dolan, & van der Molen, 2006). The conclusion we draw is that we need more thorough analyses of tests of executive functioning in preschoolers. Only close analyses of

Table 6

Correlations Between the Composite Executive Measures and False Belief

	Korean sample			English sample		
	Inhibition	Working memory	DCCS	Inhibition	Working memory	DCCS
Working memory	.36** (.13)			.58*** (.25[†])		
DCCS	.26* (.01)	.33** (.13)		.52*** (.33*)	.52*** (.25[†])	
False belief	.28* (.04)	.07 (−.25*)	.40** (.21[†])	.49*** (.24[†])	.37** (.09)	.46** (.32*)

Note. Partial correlations controlling for age and vocabulary are presented in parentheses. DCCS = dimensional change card sort.
[†]$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

individual procedures and more complex multivariate designs, like hierarchical modeling or structural equation models, will help us to understand the origins of patterns found in older children.

Links Between Executive Function and Mental State Understanding

The second question raised in these data concerns the relationship between executive skills and mental state understanding. They allow us to discuss each of the three perspectives described in the Introduction section. The claim that working memory accounts for developments in false belief understanding is not wholly supported in this analysis of children from either culture. In Experiment 1, word span scores were correlated with false belief performance, but this was not replicated in the Korean sample or with the English children in Experiment 2 once language and age were taken into account. The same patterns hold when the composite measures are explored (Table 6), which is not convergent with previous research in the West (Davis & Pratt, 1995) and in China (Tardif et al., 2007), showing that working memory contributes to false belief understanding. The data also suggest that the relationship between conflict inhibition and mental state understanding is less obvious than that in previous studies. The correlation between the blue/red and false belief tests in Experiment 2 provides some support for the account that inhibitory control is necessary for the emergence of mental state reasoning (Carlson & Moses, 2001), but again such correlation did not hold when language and age were taken into account.

The bulk of inhibitory control data showed a relatively high level of competence in Korean 3-year-olds across the two experiments, whereas even the older 4-year-old Korean children were at chance on two false belief measures in Experiment 2. This suggests that if there is a necessary relationship between these skills, the ability to inhibit a prepotent response is unlikely to be a strong grounding for "theory of mind." However, this inference needs to be tested within this cultural context using a wider range of "theory-of-mind" measures (following, e.g., Wellman & Liu, 2004) and within a longitudinal design. At the same time, the very discrepancy between the inhibitory control and false belief data presented here militate against the claim that mental state understanding provides a means of gaining self-control (Perner & Lang, 2000).

The only executive perspective to receive some support was the account arguing that switching plays a critical role in false belief understanding, as demonstrated by the robust relationship between the DCCS and false belief in the English sample. This might indicate that an acquisition of mental state understanding might require an ability to understand embedded rules, but equally could suggest the opposite. Again, only longitudinal research could tease apart causal factors here, although the evidence suggests that this would be more likely to lead to a demonstration of the interdependence of these factors (Kloo & Perner, 2003). The lack of consistent relations between executive skills and mental state understanding, particularly in the Korean sample, may well be explained by the high levels of performance on the former, particularly the measures of inhibitory control, in the Korean children. The lack of transfer of these precocious executive skills into false belief understanding in the Korean preschoolers casts doubts on the necessity of a functional relationship between these two areas of cognitive development that has been assumed with reference to Western data in accounts described in the Introduction. We turn to explore the third issue—Korean children's inhibitory skills—in the remainder of this discussion.

Why Are Korean Children So Good at Inhibitory Control?

As some of these findings are somewhat different from developmental patterns found in Western preschoolers and some emerging data in Chinese (e.g., Sabbagh et al., 2006; Tardif et al., 2007) and even in Japanese children in the study done by Ogawa and Koyasu (2006), which failed to replicate the present findings, it is important to consider just why these Korean children were so adept at tasks involving self-control. There is a long-standing debate about whether existing developmental tests fully tap the executive system (Hughes & Graham, 2002), particularly when methods are transferred from one culture to another (Vinden, 2005). It might be claimed, for example, that Korean children's daily experiences allow them to perform tasks like tower building and Luria's hand game without employing self-control, as is required in Western cultures. In the case of delay inhibition, it is possible to envisage cultural practices that require waiting to take a turn or open a present. The need to delay may not even occur to children and therefore may allow them to restrain themselves without inhibiting a prepotent response—either because such patience is highly practiced or because the response is punished. This is one reason why we did not include these tasks in Experiment 2. However, in the case of conflict inhibition, it is difficult to envisage performing these tasks without employing some inhibitory control. The impetus to copy a gesture

or say "day" when the picture of a sun is presented is difficult to understand in terms of a cultural practice that bypasses the executive system.

Although the translation of tests and constructs from one culture to another should always be considered, we propose that the performance of these Korean children raises intriguing theoretical questions about the ideas and data presented on the development of executive skills in the West. The variations between individual tests, like the day/night versus blue/red contrast and indeed the general levels of performance in the inhibition tests may provide key information in our quest to understand individual skills and the executive system as a whole. In addition, the advantage in inhibitory control generalizes to another Confucian culture, China, where preschoolers show improved performance across a battery of executive tasks (Sabbagh et al., 2006) and in delay gratification (Zelazo & Qu, 2005).

Why do Confucian cultures appear to promote early executive skills? There could be several explanations for the results obtained here and elsewhere. First, Sabbagh et al. (2006) suggest that genetic factors might be involved. Studies show that problems in inhibitory control are linked with attention deficit hyperactivity disorder (ADHD; Barkley, 1997), that patterns of ADHD are more infrequent in East Asian cultures, and that a particular gene linked with ADHD is found to be present in up to 40% of children from the Americas, but only 1.7% of East Asian children (Chang, Kidd, Livak, Pakastis, & Kidd, 1996). By this chain of association, there might be a genetic influence on Korean children's performance. However, individual data would be needed to test this possible explanation. At present, there are several problems in making claims about the genetic links with developmental disorders (Buitelaar, 2005; Hutchison, Stallings, McGeary, & Bryan, 2004); therefore, we must exercise caution about generalizing from this literature to the speculation that there is a genetic basis for advanced inhibitory control in East Asian, and particularly Korean, children.

We feel that a cultural explanation is more likely. The literature summarized in the Introduction section (e.g., French & Song, 1998) suggests that the Korean national preschool curriculum draws from Dewey's and Piaget's emphasis on the individual child's experiences with educational materials, but it is clear that traditional Confucian beliefs in collective responsibilities persist. As Kwon's (2004) interview study suggests, teachers continue to prefer to adopt "whole class teaching, teachers' authority, extrinsic motivation, worksheets, and separation of play time and work time, which are considered inappropriate in Western early years education" (p. 311). It could be that the emphasis placed by Korean early education teachers on self-control accounts for the findings we obtained. We note that researchers might be able to capitalize upon differences between East Asian cultures in their interpretation of Confucian values that have long been discussed (e.g., Tobin, Wu, & Davidson, 1989) and the rapid cultural changes that accompany shifts in the nature of educational practices in preschools in different East Asian cultures (Tobin, Karasawa, & Hsueh, 2004).

We should not simply dwell upon teachers. The evidence suggests that cultural influences permeate every relationship, and we could easily have identified many aspects of parenting that reflect the Confucian ideal and that contrast with parenting in the West (for a review, see Chao & Tseng, 2002). For example, analysis of parental values by Park and Cheah (2005) shows that Korean mothers emphasized the importance of sharing and helping over controlling their children's emotions. Almost 90% of children younger than 7 years of age sleep with their parents, who explain this practice as emphasizing familial bonds and interpersonal relationships (Yang & Moon, 2002).

Although there is agreement between parents and teachers on the importance of socializing Korean children into a society emphasizing self-control within harmonized social interactions, there might be other possible reasons why Korean children appear to show higher levels of performance on tests involving inhibition. For example, child-directed speech may be a cradle for learning the elements of control. Such speech directed towards babies in Korea is characterized by a greater use of verbs than that in other languages (Choi & Gopnik, 1995). The emphasis is on action and, by implication, its control (Kim, McGregor, & Thompson, 2000). It seems likely that this input and a related early child output of action terms are geared towards stressing the importance of self-control: "We might think of the input to Korean children as a medium that naturally teaches a great deal about actions and relations, whereas English input focuses more intently upon object classes" (Gopnik, Choi, & Baumberger, 1996, p. 201). Gopnik et al. (1996) drew upon analyses of the role of language in cognition (Choi & Bowerman, 1991) that are resonant with the cultural approach to the executive function expressed by Luria (1961), in which cultural and linguistic practices exert a strong influence upon cognitive development. Certainly, there is a need to examine the role of early language in the development of executive control and false belief within a Korean sample.

This study points to the need to examine cultural influences on the executive system. One way to disentangle cultural and possible hereditary influences would be to study Korean children who are assimilating into Western cultures. For example, Korean American parents who identify more with American cultural values have children who are rated as more disruptive in preschool settings (Farver & Lee-Shin, 2000). Such designs could clarify the specific social components that influence the development of the executive system.

References

Barkley, R. A. (1997). *ADHD and the nature of self-control*. New York: Guilford Press.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Brooks, P. J., Hanauer, J. B., Padowska, B., & Rosman, H. (2003). The role of selective attention in preschoolers' rule use in a novel dimensional card sort. *Cognitive Development, 18*, 195–215.

Buitelaar, J. K. (2005). ADHD: Strategies to unravel its genetic architecture. *Journal of Neural Transmission, 69*, 1–17.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contribution of inhibitory control and working memory. *Infant and Child Development, 11*, 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*, 299–319.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory control in young children's difficulties with deception and false belief. *Child Development, 69*, 672–691.

Chang, F.-M., Kidd, J. R., Livak, K. J., Pakastis, A. J., & Kidd, K. K. (1996). The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. *Human Genetics, 98*, 91–101.

Chao, R., & Tseng, V. (2002). Parenting of Asians. In M. Bornstein (Ed.), *Handbook of parenting. Vol. 4: Social conditions and applied parenting* (2nd ed., pp. 59–93). Mahwah, NJ: Erlbaum.

Chasiotis, A., Kiessling, F., Winter, V., & Hofer, J. (2006). Sensory motor inhibition as a prerequisite for theory of mind: A comparison of clinical and normal preschoolers differing in sensory motor abilities. *International Journal of Behavioral Development, 30*, 178–190.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*, 83–121.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 27*, 497–530.

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

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62*(4).

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "do as I say, not as I do." *Developmental Psychobiology, 29*, 315–334.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contributions of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26*, 465–486.

Farver, J. A. M., & Lee-Shin, Y. (2000). Acculturation and Korean-American children's social and play behavior. *Social Development, 9*, 316–336.

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7*, 103–115.

French, L., & Song, M. (1998). Developmentally appropriate teacher-directed approaches: Images from Korean kindergartens. *Journal of Curriculum Studies, 30*, 409–430.

Frye, D. (1999). The development of intention: The relation of executive function to theory of mind. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 119–132). Mahwah, NJ: Erlbaum.

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

Gerstadt, C. L., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½-7 years old on a Stroop-like day-night test. *Cognition, 53*, 129–153.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*, 26–37.

Gopnik, A., Choi, S., & Baumberger, T. (1996). Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development, 11*, 197–227.

Gordon, A. C. L., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70–83.

Hala, S., Hug, S., & Henderson, A. (2003). Executive function and false-belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognition and Development, 4*, 275–298.

Happaney, K., & Zelazo, P. D. (2003). Commentaries: Inhibition as a problem in the psychology of behavior. *Developmental Science, 6*, 468–470.

Hughes, C. (1996). Control of action and thought: Normal development and dysfunction in autism: A research note. *Journal of Child Psychology and Psychiatry, 37*, 229–236.

Hughes, C. (1998a). Executive function in preschoolers: Links with theory of mind and emotion and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.

Hughes, C. (1998b). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology, 34*, 1326–1339.

Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions. *Child and Adolescent Mental Health, 7*, 131–142.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017–2036.

Hutchison, K. E., Stallings, M., McGeary, J., & Bryan, A. (2004). Population stratification in the candidate gene study: Fatal threat or red herring? *Psychological Bulletin, 130*, 66–79.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English- and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language, 27*, 225–254.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–467.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*, 1823–1839.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490–507.

Kwon, Y.-I. (2002). Western influences in Korean preschool education. *International Education Journal, 3*, 153–164.

Kwon, Y.-I. (2004). Early childhood education in Korea: Discrepancy between national kindergarten curriculum and practices. *Education Review, 56*, 297–312.

Lang, B., & Perner, J. (2002). Understanding of intention and false belief and the development of self-control. *British Journal of Developmental Psychology, 20*, 67–76.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59–80.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behaviour*. London: Pergamon Press.

McMullen, M., Eliker, J., Wang, J., Erdiller, Z., Lee, S.-M., Lin, C.-H., et al. (2005). Comparing beliefs about appropriate practice among early childhood education and care professionals in the U.S., China, Korea and Turkey. *Early Childhood Research Quarterly, 20*, 451–464.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex ''frontal lobe'' tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Moses, L. J., & Carlson, S. M. (2004). Self-regulation and children's theories of mind. In C. Lightfoot, C. Lalonde, & M. Chandler (Eds.), *Changing conceptions of psychological life* (pp. 127–146). Mahwah, NJ: Erlbaum.

Ogawa, A., & Koyasu, M. (2006, January). *The relationship between theory of mind and executive function in Japanese young children*. Poster presented at the International Symposium on Inhibitory Processes in the Mind, Kyoto, Japan.

Park, H., Kwak, K., & Park, K. (1995). *Korean–Wechsler Preschool and Primary Scale of Intelligence*. Seoul, South Korea: Psychological Corporation.

Park, S.-Y., & Cheah, C. S. L. (2005). Korean mothers' proactive socialization beliefs regarding preschoolers' social skills. *International Journal of Behavioral Development, 29*, 24–34.

Parmar, P., Harkness, S., & Super, C. M. (2004). Asian and Euro-American parents' ethnotheories of play and learning: Effects on preschool children's home routines and school behavior. *International Journal of Behavioral Development, 28*, 97–104.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship? In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience* (pp. 150–181). Oxford, UK: Oxford University Press.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child Development, 73*, 752–767.

Perner, J., Stummer, S., & Lang, B. (1999). Executive functions and theory of mind: Cognitive complexity or functional dependence? In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 133–152). Mahwah, NJ: Erlbaum.

Russell, J. (1996). *Agency: Its role in mental development*. Hove, UK: Erlbaum Taylor & Francis.

Sabbagh, M., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*, 74–81.

Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (Eds.). (2005). *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind*. Mahwah, NJ: Erlbaum.

Simpson, A., & Riggs, K. J. (2005a). Factors responsible for performance on the day-night task: Response set or semantics? *Developmental Science*, *8*, 360–371.

Simpson, A., & Riggs, K. J. (2005b). Inhibitory and working memory demands of the day-night task in children. *British Journal of Developmental Psychology*, *23*, 471–486.

Tardif, T., So, C. W.-C., & Kaciroti, N. (2007). Language and false belief: Evidence for general, not specific effects in Cantonese-speaking preschoolers. *Developmental Psychology*, *43*, 318–340.

Tobin, J., Karasawa, M., & Hsueh, Y. (2004). Komatsudani then and now: Continuity and change in a Japanese preschool. *Contemporary Issues in Early Childhood*, *5*, 128–144.

Tobin, J., Wu, D., & Davidson, D. (1989). *Preschool in three cultures: Japan, China and the United States*. New Haven, CT: Yale University Press.

Towse, J. N., Redbond, J., Houston-Price, C. M. T., & Cook, S. (2000). Understanding the dimensional change card sort perspectives from task success and failure. *Cognitive Development*, *15*, 347–365.

Vinden, P. (2005, April). *Discussion in symposium on the role of culture in the development of executive function and mental state understanding in preschoolers*. Presented at the bien-nial conference of the Society for Research in Child Development, Atlanta, GA.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

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.

Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development*, *75*, 523–541.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, *7*, 131–149.

Yang, C. K., & Moon, H. M. (2002). Cosleeping in young Korean children. *Journal of Developmental and Behavioral Pediatrics*, *23*, 151–157.

Zelazo, P. D., & Jacques, S. (1997). Children's rule use: Representation, reflection, and cognitive control. In R. Vasta (Ed.), *Annals of child development* (Vol. 12, pp. 119–176). London: Jessica Kingsley Press.

Zelazo, P. D., & Qu, L. (2005, April). *Early development of executive function in Chinese and Canadian children*. Paper presented at the biennial Conference of the Society for Research in Child Development, Atlanta, GA.