

## **Relationship between theory of mind and executive function in young children assessed by class teachers**

Tomoko SUGIMURA

Ayako ANDO

Fukuoka University of Education

(Accepted August 20, 2009)

Three- to 6-year-olds were given first-order and second-order false-belief tasks, and their class teachers assessed their everyday behavior in the nursery school using the Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P) and T-K Developmental Scale for Children (T-K Scale). Main results are as follows: a) The subscale ratings of the BRIEF-P, i.e., Inhibit, Working Memory and Plan/Organize, were related to the performances of both first- and second-order false-belief tasks. b) Four of the 5 BRIEF-P subscales were related to the second-order false-belief task; however, only one significant correlation among the four subscales of the T-K Scale was observed. These results are discussed from the perspective of cognitive skills to better understand higher-level mental states.

Key words: young children, theory of mind, executive function

Numerous recent studies of preschoolers have reported robust associations between the acquisition of a theory of mind (ToM) and performance on tests of executive function (EF). ToM is the ability to ascribe mental states, such as desires, beliefs, feelings and intentions, to oneself and to others. Knowing what people want, think, feel and intend enables one to make behavioral predictions about how people will act (Perner, 1991; Perner & Lang, 1999). The term EF refers to higher-order, self-regulatory cognitive processes that aid in the monitoring and control of thought and action (Carlson, 2005). These processes, responsible for higher-level action control, include working memory, inhibitory control, attentional flexibility, and planning. The frontal lobes of the brain are believed to be heavily associated with EF.

The most frequently used measure of ToM is the false-belief paradigm. In the standard task (e.g., Wimmer & Perner, 1983), children observe a protagonist put an object into a location X and then witness, in the absence of the protagonist, the object being transferred from location X to location Y. To assess the children's understanding of the protagonist's false belief of the location of the object, they are asked to predict where the protagonist will look for the object upon their return. The typical developmental trend is that, at 3 years, almost all children answer incorrectly with the actual location of the object (i.e., location Y), whereas most children of 4 years and older answer correctly (i.e., location X).

Executive function is often assessed by means of neuropsychological tasks suitable for young children. These tasks include Backward Digit Span (Carlson, Moses, & Breton, 2002) as a measure of working memory, Luria's Hand Game (Hughes, 1998) measuring inhibitory control, Dimensional Change Card Sorting (DCCS) (Carlson & Moses, 2001) measuring attentional flexibility, and Tower of Hanoi (Carlson, Moses, & Claxton, 2004) measuring planning.

Recent studies investigating ToM and EF performance among young children have demonstrated that the understanding of false-beliefs is strongly correlated with several components of EF tasks. For example, in Gordon and Olson (1998), a partial correlation controlling the effect of age between the performance of a counting and labeling task (as a measure of working memory) and the false-belief task was significant for 3- to 5-year-olds. It has also been shown that inhibitory control tasks (Carlson, et al., 2004) and attentional flexibility tasks (Frye, Zelazo, & Palfai, 1995) are significantly related to the understanding of false-beliefs.

In contrast to such previous studies that assessed children's EF by means of EF tasks in which the children themselves performed, recently Gioia, Espy, and Isquith (2002) developed an EF scale for teacher and parent rating of preschool-aged children, called the Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P). The BRIEF-P is a questionnaire to assess EF behaviors in the home and preschool environments, including behavior-rating items such as "Does not stop laughing at funny things or events when others stop." It contains 63 items comprising 5 subscales that measure different aspects of EF: Inhibit, Shift, Emotional Control, Working Memory, and Plan/Organize.

The aim of the present study is to examine the relationship between performance in the false-belief task and EF behaviors as assessed by the teachers of a nursery school. Little is known about whether the correct understanding of a false-belief is related to EF behaviors observed by teachers and parents in everyday contexts. Thus, in this study the relationship between the ToM performance of children and EF behavior ratings by their teachers was examined.

We first prepared two kinds of false-belief tasks different with respect to the mental state level they involved, that is, first-order and second-order false-belief tasks (e.g., Perner & Wimmer, 1985). Although first-order tasks refer to the standard, abovementioned paradigm and assess the understanding of another person's perception of a situation (e.g., "She thinks..."), second-order tasks are used to examine the understanding of different individuals' perceptions regarding the mental states of others (e.g., "She thinks that he thinks..."). We used two kinds of first-order false belief tests that differed in the level of abstraction used: a concrete-object version in which an experimenter gave instructions with the assistance of concrete objects, including a dollhouse and cupboards, and a picture-story version in which instructions were given with the assistance of a series of picture-story cards. It was assumed that the performance of these three tasks differing in the levels of mental state and abstraction they involve would relate to different aspects of EF.

Second, we separately analyzed the understanding of false-belief, reality, and the situation prior to any action being performed. Previous studies have defined performance on the false-belief task as the sum score of the false-belief question asking where the protagonist will look for the object, the reality question asking about the present location of the object, and the past question asking about the past location of the object. However, it seems that these questions measure different aspects of cognitive skills. To provide a correct response to the reality question, the present situation must be accurately explained. For the past question, the child must remember the situation prior to any action being performed and describe it. Finally, for the false-belief question, the child must infer the protagonist's belief based on both present and past information. In this way, we hypothesized that performance on these questions was related to different aspects of EF.

Finally, we examined the relationship between ToM tasks and children's behavior in a nursery school as assessed by a traditional developmental scale in addition to the BRIEF-P. The T-K Developmental Scale for Young Children (1979), including the Verbal Comprehension, Peer Relation and Self Control assessment items, were used. We assumed that Peer Relation was related to the understanding of false-beliefs because it comprises items assessing cooperation with friends and leadership roles. We also assumed that Verbal Comprehension might be associated with an

understanding of reality that requires a verbal description of the present situation. We hypothesized that different patterns of relationship between the T-K Scale and ToM, and the BRIEF-P and ToM, would be observed.

## Method

### Participants

Thirty-six young children (ages 3:6-6:4, mean age = 4:11) participated in the experiment.

### Materials

#### *Theory of mind scale*

First-order false-belief test (concrete-object version): Three kinds of first-order false belief tasks, each selected from the Theory of Mind Development Test (ToM test) produced by Morinaga, Azuma, Mayuzumi, Kakinuma, & Konno (2002), were used. An experimenter gave instructions with the assistance of experimental tools, including a dollhouse, girl and boy dolls, and shoe cupboards. An example of the main story and instructions for the false-belief tasks (i.e., shoe-cupboard task) is as follows: 1) Akiko comes back to her house, puts her shoes in one of two shoe cupboards (i.e., cupboard A), and enters the house. The outside is invisible from the inside of the house. 2) Taro then approaches Akiko's house, moves Akiko's shoes from cupboard A to the other cupboard (i.e., cupboard B), and then runs away. 3) Akiko prepares to leave the house and is about to pick up her shoes. 4) "Which cupboard is Akiko going to open to pick up her shoes, cupboard A or cupboard B?" (i.e., false-belief question). 5) "Now, in which cupboard are Akiko's shoes, cupboard A or cupboard B?" (i.e., reality question). 6) "In which cupboard did Akiko put her shoes at the beginning?" (i.e., past question). The other two tasks had the same structure as this shoe-cupboard task, although their contexts were different.

First-order false-belief test (picture-story version): Three kinds of first-order false belief tasks in which instructions were given with the assistance of a series of picture-story cards were prepared. The structure of each basic story and set of instructions was very similar to those of the concrete-object version.

Second-order false-belief test: Two kinds of second-order false belief tasks, in which instructions were given with the assistance of a series of picture-story cards, were prepared. An example of the main story and task instructions (i.e., a box task) is as follows: 1) Ayako comes to her room, places her doll inside one of two boxes (i.e., box A), and leaves the room. 2) Tomoko comes into the room and moves the doll from box A to the other box (i.e., box B). Ayako sees what Tomoko does from outside the house, through a window, but Ayako is not aware of this. 3) Ayako comes back to her room and is about to pick up her doll. 4) "Which box does Tomoko think Ayako is going to open to pick up her doll, box A or box B?" (i.e., second false-belief question). 5) "Now, which box is Ayako actually going to open to pick up her doll, box A or box B?" (i.e., reality question). 6) "In which box did Ayako put her doll at the beginning?" (i.e., past question).

#### *Executive function scale*

To measure executive function, the Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P) by Gioia, Espy, and Isquith (2002) was used. The BRIEF-P consists of 16 Inhibit items (e.g., "Does not stop laughing at funny things or events when others stop,"

"Is unaware of how his/her behavior affects or bothers others"), 10 Shift items (e.g., "Become upset with new situations," "Has trouble changing activities"), 10 Emotional Control items (e.g., "Mood changes frequently," "Becomes upset too easily"), 17 Working Memory items (e.g., "When given two things to do, remembers only the first or last," "Cannot stay on the same topic when talking"), and 10 Plan/Organize items (e.g., "Needs to be told to begin a task even when willing to do it," "Does not complete tasks even after given directions"). Each item was assessed by a 3-

point scale with anchored ratings: 2 = Never, 1 = Sometimes, and 0 = Often.

#### *Developmental scale*

Ninety items selected from the T-K Developmental Scale for Young Children (1979) were used: 30 Verbal Comprehension items (e.g., “Addresses the names of his/her classmates and class teacher,” “Reads a picture story book”), 20 Peer Relation items (e.g., “Makes sand mound in cooperation with classmates,” “Keeps rules and takes turns when playing with a ball”), 20 Self Control items (e.g., “Acts up to a promise,” “Does not get angry or tearful when losing a game”), and 20 Autonomy items (e.g., “Gives willing assistance to a classmate who is in trouble,” “Puts on a hat in the sun unless he/she has been given specific instructions”). Each item was assessed by a 4-point-scale with anchored ratings: 3 = Very good, 2 = Good, 1 = Poor, and 0 = Very poor.

#### *Procedure*

A female experimenter individually gave the ToM tests (i.e., both kinds of first-order test and the second-order test) to the children in a quiet room of the nursery school. The three tests were conducted over a three-day period, one test per day, to ensure the children could maintain their concentration. The order of the three tests was randomized for each participant. The BRIEF-P and T-K Developmental Scale items of each participant were assessed by his or her respective class teacher.

## Results

### *Scoring*

For the ToM tests, one point was given for each correct response to a question (i.e., false-belief question, reality question, and past question). For the first-order false-belief tests, the summed scores of the three tasks for each question were defined as the false-belief score, reality score, and past score. For the second-order false belief test, the summed scores of the two tasks for each question were defined in the same way as described above. For the BRIEF-P, the rating points (i.e., 2 = Never, 1 = Sometimes, and 0 = Often) of the items were summed for each of the 5 subscales. For the T-K Developmental Scale, the rating points (i.e., 3 = Very good, 2 = Good, 1 = Poor, and 0 = Very poor) of the items were summed for each of the 4 subscales. Table 1 shows the mean score and SD of the ToM tests, BRIEF-P, and T-K Developmental Scale.

### *Correlation*

Table 1 Mean score and SD of Theory of mind tests, T-K Developmental Scale and BRIEF-P

Test or Scale	Category	Max.	Mean	SD
First-order false-belief test (concrete-object version)	False-belief	3.00	1.83	1.11
	Reality	3.00	2.94	0.23
	Past	3.00	2.67	0.83
First-order false-belief test (picture-story version)	False-belief	3.00	1.83	1.21
	Reality	3.00	2.92	0.28
	Past	3.00	2.53	0.88
Second-order false-belief test	False-belief	2.00	1.28	0.91
	Reality	2.00	0.92	0.97
	Past	2.00	1.78	0.59
T-K Developmental Scale	Verbal Comprehension	90.00	54.36	12.29
	Peer Relation	60.00	41.64	10.16
	Self Control	60.00	34.75	11.24
	Autonomy	60.00	35.06	9.59
	Total	270.00	165.81	38.20
BRIEF-P	Inhibit	32.00	21.78	8.30
	Shift	20.00	16.89	2.73
	Emotional Control	20.00	13.06	4.07
	Working Memory	34.00	25.69	7.28
	Plan/Organize	20.00	16.64	4.10
	Total	126.00	94.06	22.02

Note. N = 36



Table 2 shows the coefficients of partial correlation (i.e., age was controlled) between the three kinds of score for each ToM task and the score for each subscale of the BRIEF-P and T-K Developmental Scale. For the relationship between the ToM task and BRIEF-P, the Inhibit, Working Memory and Plan/Organize items were significantly associated with all of the false-belief scores, whereas the Shift items were only significantly associated with the second-order false-belief scores. In addition, the Shift and Working memory items were significantly associated with the past score in the concrete-object version of the first-order task.

As for the relationship between the ToM task and T-K Developmental Scale, the Peer Relation, Self Control and Autonomy items were significantly associated with the false-belief scores in the concrete-object version of the first-order task, while the Peer Relation and Self Control items were significantly associated with the false-belief scores in the picture-story version. Only the Peer Relation items were found to be significantly associated with the false-belief scores in the second-order task. Furthermore, the Verbal Comprehension, Peer Relation, and Autonomy items were significantly associated with the reality score, while the Peer Relation and Self Control items were significantly associated with the past score in the concrete-object version of the first-order task. Only the Peer Relation items were significantly associated with the Past scores in the picture-story version of the first-order task.

### Discussion

According to the relationship observed between the BRIEF-P and first-order tasks, Inhibit and Working Memory were related to both of the false-belief scores. These results correspond with the results of previous studies that examined the relationship between children's performance on false-belief and EF tasks, which involve working memory as assessed by the digit span task (Carlson, et al., 2002; Gordon & Olson, 1998) and inhibitory control as assessed by the whisper task (Carlson, et al., 2004). These results indicate that not only EF tasks but also EF behavior as rated by teachers was related to the understanding of false beliefs among the participants, suggesting that basic cognitive skills such as working memory reflect children's adaptive behaviors, such as "Can stay on the same topic when talking."

Although a previous study (Carlson, et al., 2004) demonstrated that planning assessed by EF tasks such as Tower of Hanoi was not associated with ToM performance, in this study the Plan/Organize items were found to be related to the false-belief scores. This contradiction may be because of the different nature of "planning" between EF tasks and EF behavior ratings. In EF tasks such as Tower of Hanoi, children are given new rules that are never used in everyday

**Table 2** Coefficients of partial correlation between Theory of mind tests, T-K Developmental Scale and BRIEF-P

Test or Scale	Subscale	First-order (concrete-object version)			First-order (picture-story version)			Second-order		
		False-belief	Reality	Past	False-belief	Reality	Past	False-belief	Reality	Past
BRIEF-P	Inhibit	.431**	.187	.185	.438***	-.093	.097	.357**	-.105	-.060
	Shift	.241	.226	.312*	.170	.091	.243	.375**	-.122	.117
	Emotional Control	.195	-.030	.173	.120	-.102	.020	.180	-.130	.034
	Working Memory	.481**	.257	.337**	.403**	-.023	.236	.587***	-.178	.031
	Plan/Organize	.366*	.232	.218	.427**	.022	.221	.336**	-.047	.074
	Total	.461**	.221	.293*	.427***	-.049	.181	.473***	-.147	.021
T-K Developmental Scale	Verbal Comprehension	.275	.367**	.210	.164	.060	.059	.272	-.105	-.058
	Peer Relation	.404**	.292*	.433***	.311*	.101	.292*	.441**	-.225	.165
	Self Control	.450***	.275	.304*	.367**	-.062	.203	.245	-.247	.053
	Autonomy	.307*	.322*	.189	.278	-.205	.051	.224	-.122	-.018
	Total	.412**	.355**	.321*	.322*	-.034	.172	.330*	-.202	.039

Note: \* p < .10, \*\* p < .05, \*\*\* p < .01

situations and required to practice goal-oriented manipulation while following the rules. However, EF “planning” behavior in the BRIEF-P assesses the ability to control goal-oriented behavior learned from everyday experience. As a result, the EF behavior observed related to the Plan/Organize items might be associated with the false-belief tasks performed in everyday contexts.

As for the relation between the BRIEF-P and second-order tasks, the Inhibit, Working Memory and Plan/Organize items were related to the false-belief score, similar to the results for the first-order task. These results suggest that these three aspects of EF have a significant role in the cognitive process of understanding first- and second-order false beliefs. The Shift items, on the other hand, which were not related to the first-order tasks, were significantly associated with the second-order false-belief scores. This result may be because of the different processes involved in the understanding of first- and second-order false beliefs. In first-order understanding, children need to keep both the past and present location of an object in their working memory while inhibiting primary information of the present location. In second-order understanding, on the other hand, in addition to these processes children need to understand multiple perspectives of two other people and shift perspectives from one person to the other. However, thus far no study has directly examined the relationship between the EF performance of children and their understandings of second-order false beliefs. Further studies are needed to elucidate these issues.

It was interesting to find that Emotional Control was not correlated with any aspect of ToM understanding, whereas the other four subscales of the BRIEF-P were associated with some aspects of ToM. These results may have been caused by the nature of the Emotional Control items, which ask about individual traits such as “Becomes upset too easily.” In contrast, the other subscales more likely assess adaptive behavior to other people and to environmental changes, as the Inhibit items (e.g., “Is unaware of how his/her behavior affects or bothers others.”) and Shift items (e.g., “Becomes upset with new situation.”) do.

According to the relationship between the T-K scale and ToM tasks, the Peer Relation items were related to all of the false-belief scores. The Peer Relation items refer to the skills of cooperating with peers and maintaining good relationships with them. It is assumed that such social skills observed in everyday activities are based partly on the understandings of first- and second-order false beliefs. The Self Control items were also related to both types of first-order false-belief scores. This is probably because the Self Control items included items concerning group activities, such as “Can control his/her behavior when acting in a group.” As for the Verbal Comprehension items, however, no relationship was found with the false-belief scores; these items were associated with the Reality scores of the concrete-object first-order test only. These results may reflect the fact that the correct responses for the reality question require only a verbal description of the present situation, whereas the false-belief question requires the inferring of the protagonist’s belief based on both present and past information.

A number of T-K Scale subscales were related to many aspects of the concrete-object version of the first-order test. Specifically, the Verbal Comprehension, Peer Relation, and Autonomy subscales were related to the understanding of reality, while the Peer relation and Self Control items were related to the understanding of the past. These results suggest that children’s behavior as assessed by traditional developmental scales reflects the basic abilities of understanding of the present situation and recalling a past event from a short time before the present.

However, there was only one significant correlation found among the four T-K subscales and the three types of second-order scores. This result contrasts with the results of the BRIEF-P, in which four subscales were found to be related to the second-order false-belief scores. These results indicate that the understanding of higher level mental states is more strongly related to EF behavior than to social behavior as assessed by traditional developmental scales. It can therefore

be concluded that EF behavior better reflects the understanding of complicated peer relationships and an awareness of changes that can occur in the minds of one's peers.

The present study demonstrated that children's EF behaviors assessed by their teachers are significantly related to the understandings of not only first-order but also second-order false beliefs. EF behaviors in everyday contexts are likely to provide more reliable indicators of the social development of young children. However, further studies are needed to reveal which factors among the EF behaviors contribute most to a better understanding of ToM and social development.

#### References

- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28, 595-616.
- 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? Contributions 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.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10, 483-527.
- Gioia, G. A., Espy, K. A., & Isquith, P. K. (2003). Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P). Lutz, FL: Psychological Assessment Resources.
- 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.
- Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychiatry*, 16, 233-235.
- Morinaga, R., Azuma, H., Mayuzumi, M., Kakinuma, M., & Konno, M. (2002). The Theory of Mind Development Test. Tokyo: Bunkyo-Siryou Institute.
- Perner, J. (1991). Understanding the representational mind. Cambridge, MA: MIT Press.
- Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, 3, 337-344.
- Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437-471.
- Tanaka Institute for Educational Research (1979). T-K Developmental Scale for Young Children. Tokyo: Taken Publishing.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.