## Title
The Development of Second-order Social Cognition and its Relation with Complex Language Understanding and Memory



Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Arslan, Burcu
Hohenberger, Annette
Verbrugge, Rineke

Publication Date
2012

Peer reviewed

# The Development of Second-order Social Cognition and its Relation with Complex Language Understanding and Memory

**Burcu Arslan (barslan.cogs@gmail.com)**
**Annette Hohenberger (hohenberger@ii.metu.edu.tr)**
Department of Cognitive Science, Middle East Technical University, Üniversiteler Mahallesi,
Dumlupınar Bulvarı, No: 1, 06800, Ankara, Turkey

**Rineke Verbrugge (l.c.verbrugge@ai.rug.nl)**
Institute of Artificial Intelligence, University of Groningen, P.O. Box 407,
9700 AK Groningen, The Netherlands

**Abstract**

In this study, the development of second-order social cognition and its possible relationship with language and memory were investigated. For this reason two second-order false belief tasks (FBT_2), a short term memory task (WST), a complex working memory task (LST), a linguistic perspective-taking test (PTT) and a double-embedded relative clause task (REL_2) were used with 21 Turkish kindergarten children (aged 4-5 years), 47 primary school children (aged 6-12 years) and 10 adults. A general developmental trend was found for all tasks. However, a multiple linear regression showed that once age was partialed out, none of the other tasks could predict FBT_2 scores. Our findings are consistent with the modularity view that Theory of Mind (ToM) is a faculty of the human mind that does not share intrinsic content with other faculties such as language and memory (Leslie et al., 2004) and also with Apperly's (2011) 'two-systems' account of Theory of Mind. However, it develops together with those other faculties which may constrain the expression of children's false belief reasoning as a manifestation of their social cognitive abilities.

**Keywords:** Second-order Social Cognition; Cognitive Development; Theory of Mind; Language; Memory

## Introduction

In daily life, we are constantly in interaction with other agents, such as co-workers, friends and family members. As a result of this interaction, we form models pertaining to the different mental states of other agents. Social cognition of individuals is shaped based on these models. The ability to understand that different agents have different mental states, such as desires, beliefs, knowledge and intentions, which can be different from one's own, is called Theory of Mind (ToM) (Premack & Woodruff, 1978). Zero-order, first-order, second-order and higher-order reasoning are different levels of social cognition. The objects of zero-order reasoning are the rules of nature and our real-life environment. For instance, if David knows: "There is an apple on the table", he is applying zero-order reasoning. However, in daily life we are not just talking about world facts. Social interaction covers statements such as "David thinks Jessica knows that there is an apple on the table". In this situation David is applying first-order reasoning by attributing a mental state to Jessica. In addition to first-order reasoning, there are even more complex social situations like "Jack thinks David knows that Jessica knows that there is an apple on the table". This time, Jack is applying second-order reasoning by attributing first-order reasoning to David who attributes a mental state to Jessica who reasons about an object in the real world; therefore we, in turn, are attributing third-order reasoning to Jack. In this study we follow Verbrugge (2009) in using the term 'second-order social cognition' in the same sense as 'second-order theory of mind'.

First-order ToM develops between the ages of three and five (Wimmer & Perner, 1983). Interestingly, second-order ToM develops much later, between the ages of six and nine (Perner, 1988). The reason for this gap is not entirely clear yet and attracts the curiosity of researchers who are working on ToM. In Verbrugge (2009), it is hypothesized that the developmental latency between first and second-order social reasoning is due to children's constraints on serial processing rather than limitations in simple working memory capacity. More explicitly, 6 year-old children may have the ability to represent another person's mental state about their own mental state. However, they cannot apply this ability because of the lack of efficiency in applying the related mental processes serially (cf. Hendriks, Van Rijn, & Valkenier, 2007).

One of the most widely applied verbal paradigms for studying ToM is the false-belief task (FBT), which has first been studied by Wimmer and Perner (1983). The main idea of the FBT is to examine whether children can attribute a false belief to another agent in a given story where they know the reality and the other agent has a false belief. Using language comprehension tasks is another verbal paradigm for studying the development of social cognition. These tasks generally test listeners' semantic and/or pragmatic inferential abilities. In these tasks, the listener has to take the speaker's linguistic alternatives and his/her choice into account to understand the correct meaning of the sentence. In the present study, a complex language comprehension task was used to test children's ability to meet a questioner's expectations of an appropriate answer to his / her questions in a given context, by taking the questioner's perspective.

The development of ToM has been intensely investigated and documented in the literature. However, one of the debatable issues is still if other factors contribute to ToM understanding during development. There is one influential factor, namely language development (Astington & Baird, 2004; Flobbe, Verbrugge, Hendriks, & Krämer, 2008): Does language have an effect on acquiring ToM, or not? In order to elucidate the relationship between language and social reasoning during development, two language tasks were used in this study. The first one is a complex language comprehension task in which the morphological structure in particular zero- *vs*. accusative-marked nouns had to be mastered and the second one is a double-embedded relative clause task in which complex syntactic structure was required. Generally, complement clauses are studied in the literature in order to investigate the relationship between the syntactical component of language and ToM. Unlike complement clauses, relative clauses do not necessarily involve mental state predicates. Using relative clauses instead of complement clauses allows us to specifically focus on the structural format of 2-way embedding. This is a purely structural parallel between second-order embedding in the thought domain and second-order embedding in the language domain.

## Method

### Participants

A total of 68 (35 female, 33 male, M=7.53 yrs, SD=2.53) children and 10 (5 female, 5 male, M=33.48 yrs, SD=10.00) adults participated in the experiments. Children's grades varied from kindergarten to fifth-grade, and their age range varied from 4 to 12 years. There were 21 kindergarten children (M=4.43 yrs, SE= .07), 17 first graders (M=6.99 yrs, SE= .09), 15 third graders (M=9.01 yrs, SE= .08) and 15 fifth graders (M=11.00 yrs, SE= .10). A group of 10 adults served as a control group.

### Design

A cross-sectional study with the four above-mentioned age groups was conducted. All subjects participated in the following five tests: word span task, second-order false belief task, perspective-taking test, second-order relative clause task, and listening span task.

All of the tests were completed in one session, which varied from 25 minutes to 35 minutes. Children were tested in a quiet empty classroom at their school.

### Materials and Procedures

**Word Span Task (WST)** Children's short term memory was tested with Ünal's (2008) Turkish version of the WST. Mono-syllabic Turkish words such as "saç, tuz" and "yurt" (hair, salt and country) were selected considering their frequency in daily usage and easy pronunciation. There are a total of seven sets, which consist of 2 to 8 words. Each set comprised 3 sub-sets. An example of a set of 2 words is as follows: i) köşk muz (manor banana); ii) pil üst (battery upper); iii) buz dört (ice four).

The words from these sets were read to the participants starting from the 2-word set. After reading one set (i.e. köşk muz), the participant repeated the words in that order. If the participant made two errors, i.e., any error in two of the three sub-sets of that level, the experiment was terminated. If he/she made fewer than two errors, the subsequent, next higher, set was read (i.e. the 3-word set). The word span equals the correct number of words at the respective level at which the child makes fewer than two errors. Thus, in the analysis the word span range may vary between 1 and 8.

**Second-order False Belief Task (FBT_2)** This task consists of two different second-order false belief stories, namely the 'Birthday Puppy' Story and the 'Chocolate Bar' Story. Both stories were adapted from English to Turkish from Flobbe et al. (2008) with the authors' permission. These stories were told to the subjects by presenting Flobbe et al.'s (2008) drawings. Second-order embedding structures such as "Mary thinks that John thinks the chocolate is in the drawer" were not used in the stories.

For both stories, the drawings were shown to the participants while the stories were being told. The order of stories in the false belief task was balanced. If a participant gave correct answers to the reality control (Where is the chocolate now?), first-order ignorance (Does John know that Mary has hidden the chocolate in the toy chest?), linguistic control (Does Mary know that John saw her hide the chocolate?), second-order false belief (Where does Mary think that John will look for the chocolate?) and justification (Why does she think that?) questions, the participant's score of the first story was 1. The total score for both of the false belief stories is therefore minimally 0 and maximally 2. Since the questions preceding the second-order false belief question are control questions, they need to be answered correctly in order for the false belief question to be possibly answered correctly.

**Perspective-taking Test (PTT)** The perspective-taking test is a complex language comprehension task including two close-ended questions with two options. In the story, Ali tells Ayşe that he is planning to go to the bookstore today. Ayşe wants Ali to buy a storybook. Ali goes to the bookstore and buys the book. While Ali is going back home, he sees his friend Mehmet on the road. Mehmet asks Ali what he did today. At this point, the experimenter asks the participant which of the following answers, (a) or (b), Ali gives to Mehmet:

a) Kitab-ı al-dı-m. (I bought the book.)
   Book-ACC buy-PAST-1PSg
b) Kitap al-dı-m. (I bought a book.)
   Book buy-PAST-1PSg

After that, the experimenter continues to tell the story. Ali goes back home. Ayşe opens the door and asks Ali what he did today. This time the experimenter asks the participant which of the following answers, (a) or (b), Ali gives to Ayşe: a) Kitabı aldım or b) Kitap aldım. The order of the answers to the close-ended questions provided to the subjects was balanced across participants. Since Mehmet asks the question without having been introduced to the book before, the expected answer for the first question was the answer with zero-marking: “Kitap aldım” (referring to “a book”). The reasoning behind this answer is as follows: Ali knows that Mehmet does not know that Ali went to the bookstore to buy a storybook for Ayşe. However, the expected answer for the second question was “Kitabı aldım” (referring to “the book”) rather than “Kitap aldım”, since Ayşe had told Ali that she wanted him to buy a storybook, The reasoning behind this answer is as follows: Ali knows that Ayşe wants to know whether Ali bought the storybook or not.

If the participant gave the expected answer to the two questions, s/he received a score of 2 points.

**Double-embedded Relative Clause Task (REL_2)** This task is related to the comprehension of relative clauses (RC) in Turkish. This task was adapted from Özge (2010) with the author’s permission. The questions and the drawings were modified to double-embedded RCs in order to analyze the participants’ second-order language embedding abilities, on a par with their second-order ToM abilities. One practice and 6 experimental items were used. Figure 1 demonstrates the drawings for one of the questions related to this task. The critical positions for finding the correct answers were equally distributed across the drawings (3 times in the first row and 3 times in the second row) and between right (2 times), left (2 times), and central position (2 times).

First, introductory pictures were shown to the participants in order to familiarize them with the animals in the action by telling the name of the animals and the actions (e.g., “this is a pushing sheep”). After that, the pictures representing the questions were shown one by one (see Figure 1).

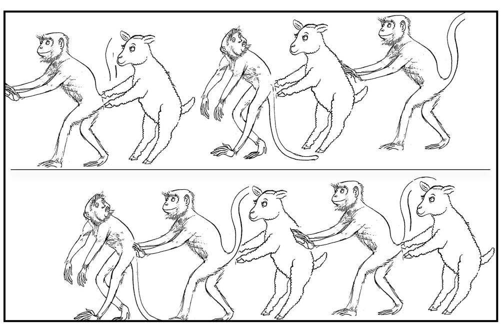

Figure 1: Hangi resimde kuzuyu iten maymunu iten bir kuzu var? (“In which picture there is a sheep pushing a monkey that is pushing a sheep?”)

The first and second rows of the picture were pointed out in order to make it clear that there are two separate lines of pictures by saying, “This is the first picture and this is the second picture”. In the trial session, it was explained that the participants were required to point out the row with the animals related to their answer. If they could not answer correctly in the trial, the correct animals were pointed out by the first author with necessary explanations. The sentences were repeated up to 4 times. Participants’ scores could range from 0-6.

**Listening Span Task (LST)** In order to measure complex working memory, Ünal’s (2008) English-to-Turkish adaptation of the original LST was used with the author’s permission. The task consists of sets of sentences read out to the participants one by one. There are a total of five collections, each of which consists of six sets of sentences. The first collection contains six sets of two sentences each, the second collection contains six sets of three sentences each, and so forth, until the fifth collection, which contains six sets of six sentences each. An example of a 3-sentence set of LST is as follows: i) Muzlar bisiklete biner. (Bananas ride bicycles); ii) Elimiz beş parmaklıdır. (Our hands have five fingers); iii) Soğan acıdır. (Onions are hot).

The participants were expected to first judge the truthfulness of the sentences by saying “Yes” or “No”. Secondly, they had to recall the last word of all the sentences of a set told to them, in reverse order. After they gave an answer to the first sentence, the next sentence was told to them. For example, for the 2-sentence set, if the first sentence is “Muzlar bisiklete biner.” (Bananas ride bicycles), the participants were required to say “Hayır[1];biner”. After that, if the second sentence is “Elimiz beş parmaklıdır.” (Our hands have five fingers), they were required to say “Evet[2];parmaklıdır, biner.”. If the participant made less than two mistakes in a sentence collection, the subsequent sentence set, which comprised one more sentence, was told to the participant. The score of the participants equaled the number of sentence collections in which they did not make more than one mistake. Participants’ scores could range from 0-6.

## Hypotheses

We hypothesized main effects of grade for all tasks. Children’s performance should increase with increasing grade. Furthermore, we hypothesized that FBT_2 could be predicted by the remaining tasks, in particular by the complex language tasks, PTT and REL_2, to the extent that those share variance with it.

## Results

First, the statistical analyses of children’s responses to the five tasks are presented. Later, the results of the fifth graders will be compared with the results of adults. The p values are two-tailed, unless stated otherwise (in which case they are one-tailed). In order to analyze the developmental trend in the tasks used in the experiment, the data was divided into four groups according to the children’s grades (kindergarten, 1st, 3rd, 5th grade). Since the data violates normality assumptions, non-parametric Kruskal-Wallis and Mann-Whitney tests were used. Since six Mann-Whitney Tests were used to test the difference across the grades, the

[1] ‘Hayır’ means ‘No’.
[2] ‘Evet’ means ‘Yes’.

alpha level for the Bonferroni correction was set to .008. Although the data was not normally distributed, linear regression analysis was used to investigate the relationship between the second-order false belief task and the other tasks. Error bars in Figures 2-6 represent standard errors.

### FBT_2

There is a significant difference in performance between the grades ($\chi^2$ (2) = 40.22, p= .000). According to the Mann-Whitney Tests, while there is a steady increase in performance, there is no significant difference between the first and third grades and between the third and fifth grades. However, there is a significant difference between kindergarten and grade one (Z= -3.73, p= .000), kindergarten and grade three (Z= -4.73, p = .000), kindergarten and grade five (Z= -5.36, p = .000), and grade one and five (Z= -2.99, p = .003). Figure 2 shows the mean values of the FBT_2 score according to the grades. Since all of the adults and all of the fifth graders answered all of the questions correctly, there is no significant difference between the adults' and fifth graders' FBT_2 performance ($\chi^2$ (2) = 0.00, p= 1.00).

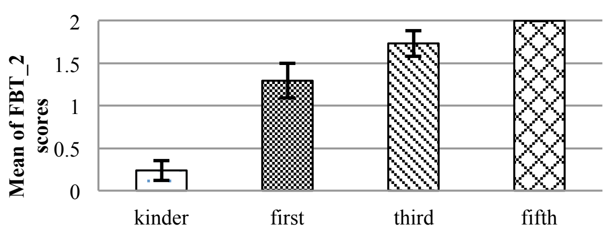


Figure 2: The development of FBT_2

### WST

There is a significant difference between the grades ($\chi^2$ (2) = 24.67, p= .000). According to the Mann-Whitney Tests, there is no difference between the first, third and fifth grades, while there is a significant difference between kindergarten and grade one (Z= -3.06, p= .002), kindergarten and grade three (Z= -4.14, p= .000), and kindergarten and grade five (Z= -3.59, p= .000). Figure 3 shows the mean values of the Word Span Task score according to the grades. The analysis showed that there is a significant difference between the adults' and fifth graders' performance ($\chi^2$ (2) = 8.925, p= .003).

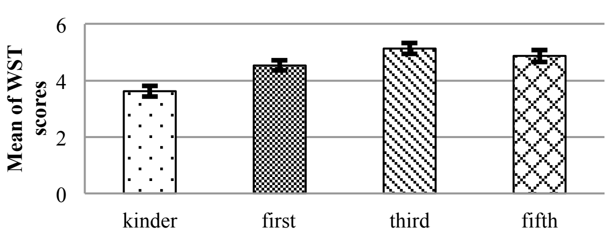


Figure 3: The development of WST

### PTT

There is a significant difference between the grades ($\chi^2$ (2) = 8.53, p= .036). According to the Mann-Whitney Tests, there is no difference between the kindergarten and grade one, grade one and three, grade one and five, grade three and five, while there is a significant difference between the kindergarten and grade five (Z= -2.473, p= .006, one-tailed). Figure 4 shows the mean values of the perspective-taking test score according to the grades. The analysis showed that there is no significant difference between the adults' and fifth graders' performance ($\chi^2$ (2) = 1.778, p= .182).

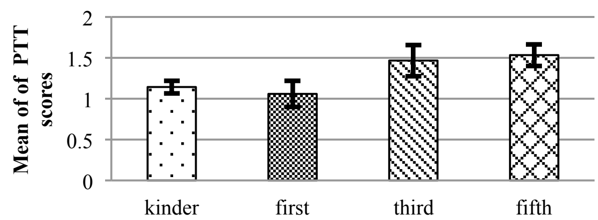


Figure 4: The development of PTT

### REL_2

There is a significant difference between the grades ($\chi^2$ (2) = 27.37, p= .000). In order to see which grades differ significantly from each other, Mann Whitney Tests were used. Figure 5 shows the mean values of the double-embedded relative clause score according to the grades. The analysis showed that there is a significant difference between the adults' and fifth graders' performance ($\chi^2$ (2) = 6.096, p= .014).

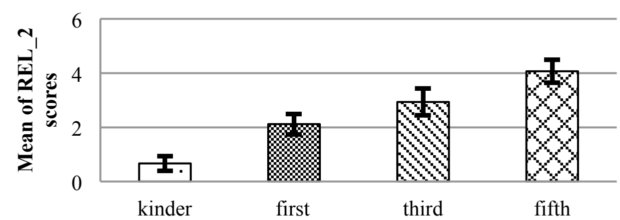


Figure 5: The development of REL_2

### LST

There is a highly significant difference between the grades ($\chi^2$ (2) = 30.87, p= .000). According to the Mann-Whitney Tests, there is no difference between the kindergarten and first grade, nor between third and fifth grades, while there is a significant difference between the kindergarten and grade three (Z=-3.53, p= .000), kindergarten and grade five (Z= -4.64, p = .000), grades one and three (Z= -2.92, p = .003), and grades one and five (Z=-4.08, p = .000). Figure 6 shows the mean values of the Listening Span Task score according to the grades. The analysis showed that there is a significant difference between the adults' and fifth graders' performance ($\chi^2$ (2) = 4.729, p= .030).

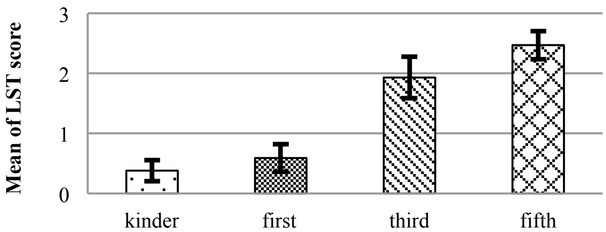


Figure 6: The development of LST

### PTT Predicting FBT_2

Since the data violates normality, the non-parametric Spearman's Rank Order Correlation was used to test the relationship between total FBT_2 and PTT scores. This analysis showed that there is no significant relationship between total FBT_2 and PTT ($r_s$ = .19, p= .126). Partial correlation was also used in order to control the other variables. When age is controlled for, the previous correlation of $r_s$= .19 between PTT and FBT_2 drops to -.095 (p= .444); when WST is controlled for, the correlation drops to .12 (p= .922), when REL_2 is controlled for, the correlation drops to .036 (p= .772) and when LST is controlled for, the correlation drops to -.22 (p= .860).

### REL_2 Predicting FBT_2

Since the data violates normality, the non-parametric Spearman's Rank Order Correlation was used to test the relationship between total FBT_2 and REL_2. This analysis showed that there is a significant relationship between total FBT_2 and REL_2 scores ($r_s$ = .54, p= .000). Bivariate regression was also used in order to predict the model of REL_2 score predicting the FBT_2 score. Using the enter method, the FBT_2 score could be predicted from the REL_2 score by the following formula: 0.24 X REL_2 + 0.673 ($F_{66,1}$= 26.196, p= .000, r= .533, $R^2$= .284). Partial correlation was also used in order to control the other variables. When age is controlled for, the previous correlation of r = .533 between REL_2 and FBT_2 drops to .10 (p= .42); when WST is controlled for, the correlation drops to .39 (p= .001); when PTT is controlled for, the correlation drops to .52 (p=.000) and when LST is controlled for, the correlation drops to .25 (p= .041).

In the light of the partial analyses, multiple linear regression was used with age and REL_2 scores as independent variables and FBT_2 as dependent variable. Using the enter method, the FBT_2 score could be predicted from age and REL_2 score by the following formula: FBT_2= 0.039 X REL_2+ 0.25 X age – 0.751 ($F_{65,\ 2}$= 42.091, p= .000, r= .751, $R^2$= .564). However, only the contribution of age is significant (β= .692, t= 6.47, p = .000).

### Multiple Linear Regression for FBT_2

Two models were constructed by using multiple linear regression to predict the FBT_2 score: first just with age as predictor and second with age and all tasks. Table 2 shows the correlations of all tasks for FBT_2. Using the enter method, the FBT_2 score could be predicted by age by the following formula: 0.27 X age – 0.814 ($F_{66,\ 1}$= 83.965, p= .000) and by age and all tasks by the following formula: 0.236 X age + 0.145 X WST + 0.045 REL_2 – 0.034 X LST – 0.130 X PTT – 1.098 ($F_{62,\ 5}$= 17.519, p= .000, r= .765, $R^2$= .586). However, only the contribution of age is significant (β= .655, t= 5.45, p = .000). Collinearity diagnostics showed that age (94%) and WST (90%) each load highly on a different single dimension. This means that age and the WST do not share variance with each other. On the other hand, PTT, REL_2 and LST share some variance with one another. Still they mainly load on their own distinctive dimension. This is because they are also related to different abilities. Moreover, the LST (60%) and the REL_2 (75%) load highest on the same dimension which shows that both tasks tap into the same cognitive ability.

Table 2: Correlations of all tasks and age for FBT_2

| Variable | Correlation (r) | p |
|---|---|---|
| Age | .748 | .000** |
| WST | .518 | .000** |
| PTT | .160 | .096 n.s. |
| REL_2 | .533 | .000** |
| LST | .503 | .000** |

## General Discussion and Conclusions

As can be seen clearly from Figure 2, a linear developmental trend was found for the second-order false belief task score from grade one (6- 7 years) to grade five (10- 12 years), preceded by a big step between kindergarten and grade 1. We can say that second-order false belief reasoning starts to develop around the age of 6, and reaches adult-like understanding at around the age of 9;5 (grade 5). These findings are compatible with Perner and Wimmer's (1985) study, which states that second-order false belief understanding occurs after the age of 6. Although kindergarten children failed in the second-order false belief task on average, there were three of them who succeeded in the 'Birthday Puppy' Story and one of them who succeeded in both the 'Birthday Puppy' Story and 'Chocolate Bar' Story. Since their Listening Span Task and double-embedded relative clause task scores were better than the others, this is compatible with the view that children before the age of 6 may indeed have second-order social cognition which may show itself if the respective cognitive resources have also developed.

The results of the Word Span Task showed a significant and clear developmental trend from kindergarten to third grade. Fifth graders' score was somewhat lower than that of the 3rd graders, but only insignificantly.

The results of the Perspective Taking Test showed that kindergarten children and first graders had scores around 1 which is the score expected by chance. The salient development occurs between 1st and 3rd grade. Making pragmatic inferences by picking up morpho-syntactic clues like case-marking is a very advanced meta-linguistic skill. Giving correct answers to the questions requires a comparison between the two case forms and a decision which of them is better suited for the given context. Even adults' performance was not perfect and did not significantly differ from that of 5th graders. However, unlike children, some of the adults changed their first wrong answer and gave a correct answer after hearing the second question. This shows that some of the adults took the hearer's perspective and/or the experimenter's intention of asking those questions into account.

To the best of our knowledge this is the first time that a double-embedded relative clause task has been devised in a Turkish developmental study. Our result revealed a very strong developmental trend in the task. Also, adults outperformed fifth graders in this task.

In Listening Span Task, participants were expected to judge the semantic truth of the sentences, to report it, to remember the last word of that sentence, and then repeat the same steps again for the next sentence by also reporting the last word of the previous sentence, and so on. Since in Turkish the present form of the verb takes the suffixes –er, -ar, -ir, -ür, -ur for positive sentences and takes the suffixes –maz, -mez for the negative ones, the most challenging part of the task for children and even for some adults was to repeat the last word of the sentence when its semantic truth was false. So, they must inhibit the regular way of reporting, and have to report it in the instructed from. This inhibition in the Listening Span Task is thought to be similar to the inhibition necessary in false belief reasoning. The results showed a strong developmental trend, again particularly between first and third graders.

Even though second-order false belief scores could be significantly predicted by all other tasks (except for Perspective Taking Test), the regression analyses showed that only the contribution of age was significant. Once age was taken out, none of the other tasks could predict second-order false belief task anymore. In view of theoretical accounts of ToM, our findings are compatible with Leslie et al.'s (2004) modular account of ToM. He and his colleagues argue that ToM is a separate cognitive faculty as compared to language or memory. It is innate, i.e., in principle in place from early on, however, in order to manifest itself it may need to await the cognitive maturation of the child in those other domains. The "serial bottleneck" (Verbrugge 2009) might be one of those constraints that is overcome during development. Our findings are also compatible with Apperly's (2011) 'two-systems' account of ToM. Apperly posits that low-level efficient processing modules take care of simple ascriptions of perception, knowledge and belief, while high-level ToM makes use of less efficient and slower to develop general knowledge and inferential reasoning processes.

Since in our study we found concurrent development in all the cognitive abilities that we tested, that is, no delay between any of them, ToM may at any time have been supported just sufficiently enough to manifest itself at that level. Indeed, it might be impossible to prove the relation between ToM and the other cognitive domains in a cross-sectional study like ours, but only in a longitudinal study where such delays may be observed within rather than across individuals.

## Acknowledgments

Rineke Verbrugge is grateful to the Netherlands Organization for Scientific Research for Vici grant NWO-277-80-01.

## References

Apperly, I (2011). Mindreaders: The Cognitive Basis of "Theory of Mind". Psychology Press; Hove.

Astington, JW., and Baird, J. (2005). Why Language Matters for Theory of Mind. Oxford University Press; Oxford.

Flobbe, L., Verbrugge, R., Hendriks, P. and Krämer I. (2008). Children's application of theory of mind in reasoning and language. Journal of Logic, Language and Information, 17 (4), 2008, 417-442. Special issue on formal models for real people, edited by M. Counihan.

Hendriks, P., van Rijn, H., and Valkenier, B. (2007). Learning to reason about speakers' alternatives in sentence comprehension: A computational account. Lingua, 117(11), 1879–1896.

Leslie, Alan M., Friedman, Ori, and German, Tim P. (2004). Core mechanisms in 'theory or mind'. Trends in Cognitive Sciences, 12, 528-533.

Özge, D. (2010). Mechanisms and Strategies in the Processing and Acquisition of Relative Clauses in Turkish Monolingual and Turkish-English Bilingual Children. PhD thesis, Middle East Technical University, Ankara-Turkey.

Perner, J. (1988). Higher-order beliefs and intentions in children's understanding of social interaction. In J.W. Astington, P.L. Harris, and D.R. Olson, (Eds.). Developing Theories of Mind. Cambridge University Press, Cambridge. 271–294.

Perner, J. and 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, 5, 125-137

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 4, 515-526.

Ünal, G. (2008). Release from Proactive Interference and its relations to Executive Functions: A developmental study on Turkish children. Master thesis, Middle East Technical University, Ankara-Turkey.

Verbrugge, R. (2009). Logic and social cognition: The facts matter, and so do computational models. Journal of Philosophical Logic, 38 (6), 649-680.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13, 103–128.