Modelling the executive components involved in processing false belief and mechanical/intentional sequences

Hiromi Tsuji, and Peter Mitchell

1Osaka Shoin Women's University
2University of Nottingham

Word count (exc. figures/tables): 4767

***Requests for reprints should be addressed to Hiromi Tsuji, 4-2-26 Hishiyanishi, Higashiosaka, Osaka 577-0807, Japan (e-mail: tsuji.hiromi@osaka-shoin.ac.jp**

Abstract

To understand the executive demands of the FB task relative to an alternative theory-of-mind (or mechanical causality) task, picture sequencing, the present study used path analyses. 166 children between 3 to 6 years old completed the FB and picture sequencing tasks, three executive function tasks (updating, inhibition and shifting) and receptive language. The model with the best fit indicated that FB performance entailed a direct contribution from shifting of attention and inhibitory control, which was independent of the significant contribution made by picture sequencing. This model clearly indicates that FB inference requires more executive processing than picture sequencing, which is used as an alternative task to measure theory-of-mind.

Keywords: theory of mind, mentalising, executive functions, picture sequencing task, false-belief task

A long-standing debate has surrounded how best to explain children's developing ability to pass a test of false belief, such as the unexpected transfer test in which child participants are asked to predict where a protagonist will search for a coveted object (Wimmer & Perner, 1983). If children predict that the protagonist will look in the place s/he last saw the object, they are credited with an understanding of false belief in that they apparently individuate between what they know is true (the object is in Location A) and what they think the protagonist believes is true (s/he falsely believes the object is in Location B). Some researchers argued that development is driven by children acquiring a concept of belief at about 4 years of age (Gopnik & Wellman, 1992; Perner, 1991; Wellman, Cross, & Watson, 2001). In contrast, some researchers argued that children already possess a concept of belief by 4 years of age but still give the 'wrong answer' in an unexpected transfer test due to various task demands apart from those relevant to having a concept of belief (Lewis & Osborne, 1990; Mitchell & Lacohée, 1991; Russell, Mauthner, Sharpe, & Tidswell, 1991; Siegal & Beattie, 1991). Specifically, Russell et al were the first to suggest that young children might err not because they have difficulty with the concept of belief but primarily because of the executive demands of the unexpected transfer test; these authors suggested that developing an ability to pass an unexpected transfer test depends on the maturation of executive processes that the task draws upon.

Bloom and German (2000) went as far as suggesting that the unexpected transfer test is not fit for the purpose of investigating development of a 'theory of mind'; at the very least, they suggested, a priority is to understand what kind of demands are imposed by the unexpected transfer test (apart from those demands that draw upon the child's concept of belief). The purpose of the current research was to begin to address this question by testing children on a variety of tests of executive processing in addition to an unexpected transfer test. The aim was to generate a statistical model that would optimally explain variance in performance data associated with the unexpected transfer test and an alternative measure that also putatively taps into the child's understanding of inner states.

To understand the cognitive demand of the false-belief task, defining sources of information that a participant processes during the task is important. Achim, Guitton, Jackson, Boutin, and Monetta (2013) outlined the 8 sources of information framework (8-SIF) to characterize tasks that are currently available for a mentalising judgment. According to the framework, mentalising tasks can involve information for agents and for contexts. Both types of information can be drawn immediately from a given task or from a person's memory, or from both. The immediate information for the agent and for the context can be presented by perceptual and linguistic means. The information drawn from a person's memory of the agent and of the context includes specific or general information. Specific information includes the details that the participants already know about the protagonist's characteristics as well as the surrounding environment related to the task, whereas general information refers to general knowledge about people and contexts (e.g. different categories of people and situations that can lead to certain assumptions), which the participants can recall and use spontaneously. Achim et al. suggest that these sources of information could contribute to, and influence, the accuracy of mentalising judgments.

A classic false-belief task involves a verbal presentation of a scenario with a puppet enactment in which a protagonist (i.e. the puppet) comes to hold a false-belief about the location of an object due to the unexpected transfer of that object. In this task, participants are required to process the perceptual and linguistic information in the story (immediate information) and then they need to access social knowledge about what people would normally do in a case of the unexpected transfer of an object (stored information) to make a correct judgment.

The picture-sequencing task (Baron-Cohen, Leslie, & Frith, 1986) is another type of task intended to assess children's mentalising ability (Achim et al., 2013). In this task, participants are asked to rearrange four snapshots of an event into sequential order such that it is coherent in relation to an agent's mental states, including their false-beliefs. This task is a good point of comparison for an unexpected transfer task because it imposes different processing demands while still tapping into the child's understanding that others can hold false beliefs. According to the 8 sources of information framework (8-SIF) by Achim et al. (2013), these two tasks (false belief and picture sequencing) impose similar processing demands except for one major difference. Both tasks include the sources of information drawn from memory for the general assumptions about the agent and for the context but do not include specific information for the agent or the context. However, these tasks differ in that uniquely for the false belief, task the immediate information for the agent and for the context is presented using both perceptual and linguistic modalities, whereas no immediate linguistic information is available for the picture-sequencing task. In summary, a false-belief task requires participants to engage in the dual tasks of processing two different sources of information, which would place an extra cognitive load on participants when processing linguistic information; this extra load is absent from the picture-sequencing task.

When characterising mentalising based on the 8-SIF, it becomes clear that tackling a false-belief task requires processing two sources of information simultaneously. Thus to make a correct mentalising judgment in a false-belief task, children would need to have sufficient executive processing capacity. Even adults are reported to find the cognitive load challenging in a false-belief task. In their study of implicit false-belief processing, Schneider, Lam, Bayliss, and Dux (2012) found that adult participants who were subject to the dual-task condition failed to track the false-location based on a protagonist's false belief, whereas those who were under the no-load condition correctly tracked the false-location. This result suggests that mentalising at the implicit level is also disrupted when extra cognitive-load associated with an unexpected transfer test is applied.

To cope with the cognitive-load in the context of unexpected transfer, certain levels of executive processes may be required. Development of Executive functions (EFs) has become one of the foci in the discussion of how children acquire their capacity to exercise mentalising judgments in the FB task. EF encompasses a wide range of abilities such as working memory (updating), inhibitory control (inhibition) and cognitive flexibility (shifting), and provides the foundation for reflective thinking, reasoning and self-control of behaviours (Blair, 2016; Zelazo et al., 2013). When engaged in most cognitive tasks, at least some of these functions are at work, although the exact role of these functions for mentalising is not yet clear.

Most studies that examine the role of executive processes in mentalising use some of the components of EF to measure this capacity. However, there is lack of consensus on the exact causal relationships between EFs and ToM. This may be due to the lack of a clear operational definition of EF (Barkley, 2012), which has led to a variability of tasks employed in various studies, and to methodological issues including a low level of internal-consistency and retesting reliability across EF tasks (van der Ven, Kroesbergen, Boom, & Leseman, 2013). The findings from a recent meta-analysis suggest a weak to moderate association between executive processing capacity and false belief judgment across different cultural contexts (Devine & Hughes, 2014).

As for the exact role of EFs in false-belief judgments, Devine and Hughes support a hybrid *emergence-expression* hypothesis, which is derived from both the *emergence* account that EFs play a functional role in the conceptual development of false-belief judgment (Russell, 1996) and the *expression* account that false-belief tasks place incidental demands on EFs (Russell et al., 1991). In an effort to tease out a *pure* incidental task demand in false-belief judgment, several studies (Carlson, Claxton, & Moses, 2015; Perner, Lang, & Kloo, 2002) used modified tasks such as the ‘think-know’ task (Moore, Pure, & Furrow, 1990) and the explanation task (Robinson & Mitchell, 1995) that have different degrees of cognitive demands. These authors found relationships between EFs and a modified version of mentalising judgments, which seemed to support the emergence account. However, irrespective of the degree of conceptual difficulty in tasks, all the tasks are akin to a classic form of false-belief context, where the sources of information are concerned. On the 8-SIF(Achim et al., 2013), these tasks require processing of immediate information conveyed through both a linguistic and a perceptual channel. In this respect, even conceptually easier tasks may still require a certain level of executive processing, which led to the identification of significant relationships between EFs and these modified tasks. Thus it is unclear that these observed relationships are due to common demand derived from conceptual overlap or cognitive-load.

The picture-sequencing task, which requires participants to process perceptual information, is likely to have a different role in the relationship with EFs if the cognitive-load in executive processing is lower than in false-belief tasks. To clarify this relationship, the present study uses the EF tasks to measure working memory (updating), inhibitory control (inhibition) and cognitive flexibility (shifting the focus of attention) and to examine the degree to which each measure contributes to cognitive processing in the picture-sequencing task and in judging false belief.

As argued by Bloom and German (2000), if a FB task is more cognitively demanding than another kind of mentalising task, then we could find different relationships between the false-belief and picture-sequencing tasks. To examine these relationships, the present study conducted path analyses of the following two models. The first model is a non-recursive model in that both the FB and picture-sequencing tasks are set as endogenous variables with a feedback loop between them. The relationship between these endogenous variables and the exogenous variables of EFs, receptive language and age are examined. The second model is recursive in that a path is only drawn from picture-sequencing to false-belief according to the assumption that false-belief is measuring more than mentalising (Bloom & German, 2000). The magnitude of the direct effects from EFs and receptive language and age, as well as indirect effects from these variables via the picture-sequencing task was examined.

We predict that a false-belief task, which requires more cognitive-load in processing sources of information, is likely to place more demands on the EFs of inhibitory control and shifting components than a picture-sequencing task.

Method

Participants.

One-hundred-and-sixty-six children (94 boys) between 3 and 6 years old (M age = 5;0, SD = .86, range 3;5 – 6;5) participated in this study. They were from a catchment area of middle to lower-middle class communities located on the outskirts of Osaka city. Ethical approval was granted by the author’s academic institution and written informed consents were obtained from the parents of the participants.

Design and Materials.

Mentalising abilities were assessed using picture-sequencing and false-belief tasks. Executive functions (EFs) were assessed using the pointing-stroop task (Berger, Jones, Rothbart, & Posner, 2000), the Dimension Change Card Sorting (DCCS) task (Zelazo, 2006), and the Digit span task. Receptive language was assessed using a Japanese version of the PVT (Ueno, Nakoshi, & Onuki, 2008).

Picture-sequencing task.

Four pictures, depicting scenes that represented an event involving causality or mental states, were presented to the children. The children were asked to put them in order to make a story. Four sets of task stimuli were chosen from the picture sequencing test (Baron-Cohen et al., 1986). These included two sets for causality events: a person causally interacting with an object and an object causally interacting with an object (mechanical condition in the original study); and two sets of stimuli involving the attribution of people's mental states for a missing teddy bear and a missing bar of chocolate (intentional condition). The children's ability to represent events was assessed based on the correct sequencing of pictures to make the story. The scoring system used by the original study was followed.

Executive function (EF) tasks.

The Pointing-stroop task (Berger et al., 2000) and Dimension Change Card Sorting (DCCS) task (Zelazo, 2006) were presented to the participants on a computer with a touch panel screen. The Digit span task that measured working memory (WM) was administered manually. The order of administering the tasks was counterbalanced.

Pointing-Stroop task.

This task was based on Berger et al. (2000) and comprises three blocks: practice (2 trials), compatibility (8 trials), and incompatibility (8 trials). The sequence of stimuli presentation is shown in Figure 1. The trial was started by touching a ‘focal point’ on the screen which initiated the display of cat and dog pictures with simultaneous audio presentations of the animal sound “meow”/ “woof”. No inter-stimuli intervals were included so that the child could initiate the next trial when s/he was ready. For scoring, accuracy was used (number of correct responses in incompatibility trials). Speed of response was not used as it was found not to be reliable in a previous study (van der Ven et al., 2013).

[Figure1]

Dimensional Change Card Sorting (DCCS) task.

This task was developed from the original manual presentation of DCCS (Zelazo, 2006). The children’s correct responses after introducing a new sorting rule was measured. The trial sequence of stimuli presentation is shown in Figure 2. A block of two practice trials, which was repeated until the child understood the task, preceded the pre-switch block (6 trials) and the post-switch block (6 trials). The task was initiated by touching a “star”, which caused the target card to be shown below the fixed picture cards. When the child touched one of the pictures (hat or bag) a sound was made to acknowledge the action regardless of accuracy. The child initiated the next trial when s/he was ready by touching the focal point on the screen. For scoring, accuracy (number of correct responses for the post-switch condition) was used.

[Figure2]

Digit span task

Following the instructions of KABC (Kaufman & Kaufman, 2004), a series of digits were presented to the child orally. The child needed to repeat back the digits to the experimenter in a forward order.

False-belief (FB) tasks

Two unexpected transfer tasks (Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983) in which a protagonist holds a false-belief about the location of an object were used to measure theory-of-mind ability. Each task included a memory question for the initial location of the item, a reality question for the present location of the item and a false-belief question for the protagonist's belief about the location of the item. All three questions had to be answered correctly to pass the task.

Procedures.

Children were tested individually in a quiet room in one session, which lasted 25 to 30 minutes. The order of the tasks was counterbalanced with the exception of the PVT, which was always administered last because the duration of this task varies dramatically depending on the individual's language ability and this may have had an effect on the child's motivation for any subsequent tasks.

Results

Descriptive statistics for all the measures are summarised in Table 1. There was no gender difference except for DCCS in which girls scored higher than boys: $t(164) = 3.23$, two-tailed $p < .001$. Cohen's $d = .51$.

[Table1]

For the picture-sequencing task, separate one-way ANOVAs on the scores for the domains of mechanical and intentional picture sequencing were conducted. The children's scores increased with age for both domains: Welch's $F(3, 83.10) = 31.50$, $p < .001$, $\omega^2 = .36$ for the mechanical domain; and Welch's $F(3, 83.60) = 28.50$, $p < .001$, $\omega^2 = .33$ for the intentional domain. Games-Howell post hoc tests (a prior alpha level was set at .05) suggest that most groups differed from one another with older groups performing significantly better than younger groups (ps < .01), with the exception of mechanical domain for the 3 and 4 year-old groups ($p = .19$), and for the 5 and 6 year-old groups ($p =.052$). The intentional domain also showed a similar trend (ps <.05) with the exception of the 5 and 6 year-old group ($p = .077$).

For executive functions (EFs), separate one-way ANOVAs were conducted with Stroop, DCCS and WM as dependent variables. Stroop differed with age: Welch's $F(3, 79.30) = 11.80$, $p < .001$, $\omega^2 = .16$. Games-Howell post hoc tests suggest that significant improvements were observed between 3 year-olds and 5 & 6 year-olds, and also between 4 year-olds and 6 year-olds: ps < .01. DCCS differed with age: Welch's $F (3, 80.00) = 10.76$, $p < .001$, $\omega^2 = .15$. Games-Howell post hoc tests suggest significant improvements between 3 & 4 year-olds and 5 & 6 year-old groups: $ps < .05$. WM differed with age: Welch's $F (3, 82.53) = 12.39$, $p < .001$, $\omega^2 = .17$. Games-Howell post hoc tests suggest significant improvements between 3 & 4 year-olds and 5 & 6 year-old groups: $ps < .05$.

For false-beliefs, a one-way ANOVA indicated a significant increase in the children's performance between age groups: Welch's $F (3, 86.31) = 38.88$, $p < .001$, $\omega^2 = .41$. The Games-Howell post hoc tests suggested significant improvements in scores between all age groups ($p < .03$) with the exception of the 5 and 6 year-old groups ($p = .38$).

Pearson's correlation coefficients between the measurements are summarised in Table 2. All the variables interrelated positively with a moderate to strong degree of association. The largest correlation in the table, excluding PVT and age, is between the intentional and mechanical domains in the picture-sequencing task. This is greater than the correlation between the intentional picture sequencing and FB, which are supposedly within the same domain for the measurement of mentalising ability. Therefore, based on these data, it is possible that intentional and mechanical domains measure the same thing, suggesting that the boundary between intentional and mechanical is spurious in this particular case.

[Table2]

Path analysis for the FB task and the picture sequencing task as endogenous variables with executive functions and receptive language as exogenous variables

To examine how performance measured by the FB and the picture-sequencing tasks was related to executive functions (EFs) and receptive language, path analyses using AMOS with maximum-likelihood estimation[1] were conducted. Two components (intentional/mechanical) for the picture-sequencing task correlated significantly each other (see Table 2), and inter task-item analysis found no justification for separating these variables of picture sequencing into component parts. Thus, these scores were combined to make a single measurement (see Supporting information 1). A total score for the FB and the picture-sequencing task were respectively set as endogenous variables while WM, Stroop, DCCS, receptive language and age were set as exogenous variables for each of the two possible model options: recursive and non-recursive. The analysis of the non-recursive model revealed that the feedback loop between endogenous variables did not converge. A path was found from picture sequencing to FB in the recursive model.

The initial recursive model that included WM did not have a good fit and WM did not have any significant effects on either of the endogenous variables. Thus another model that excluded the paths from WM was considered. This model indicated a lower ACI than the previous model with a better model fit: X^2 (2) = .72, p = .70, AGFI = .98, RMSEA < .0001, AIC = 52.72. In this model, a path from the DCCS task to the picture-sequencing task was

[1] univariate distributions were within the recommended ranges for skewness < 2 and kurtosis < 7 (West, Finch, & Curran, P. J. ,1995)

not significant and thus was removed. The final model had an improved model-fit index, which had a lower value of AIC: X^2 (3) = 1.95, p = .58, AGFI = .97, RMSEA < .0001, AIC = 51.95.

This model also had better fit than the other model, including gender: X^2 (6) = 8.87, p = .18, AGFI = .86, RMSEA = .54, AIC =103.87. This final model is shown in Figure 3. For picture sequencing performance, in addition to age (b = .43) and receptive language (b = .25), Stroop (b = .17) made a significant contribution, explaining 51% of the variance. This picture-sequencing task (b = .28) made a significant contribution to FB, with additional unique and significant contributions from receptive language (b = .19), Stroop (b =.21) and DCCS (b =.15), explaining a total of 52% of variance in FB. Age did not make a unique or significant contribution to FB. When the picture-sequencing scores for only the intentional trials were used then this gave similar results for model fit index: X^2 (3) = 1.94, p = .59, AGFI = .97, RMSEA < .0001, AIC = 51.94 (A full description is given in Supporting information 2).

[Figure3]

Discussion

The present study compared two tasks, false-belief and picture sequencing, and their relationships to executive functions (EFs) and receptive language. The model that included a path drawn from the picture-sequencing task to the FB task gave the best fit to the data; it modelled the relationship between these two tasks together with other exogenous variables. In this model, age, receptive language and Stroop (inhibition) were significant contributors to the picture-sequencing task, explaining 51% of variance. However, the executive functions of DCCS (shifting) and WM (updating) did not make significant contributions to the picture-sequencing task. The picture-sequencing task subsequently contributed to the FB task, with independent contributions from EFs (Stroop and DCCS), age and receptive language, accounting for 54% of the variance. These results suggest that FB performance could be explained by additional unique variance in DCCS and Stroop in addition to the picture sequencing task; this supports the claim by Bloom and German (2000) that the FB task measures more than just mentalising ability. Another way of interpreting the data would be that the FB task more strongly draws upon executive processes than does the picture sequencing task, irrespective of whether or not that task has intentional content.

For the EFs, the DCCS task was used to measure the flexible shifting of the focus of attention, whereas the Stroop task was used to measure inhibitory control. Additional variance in both inhibitory control and shifting of the focus of attention were needed to pass the FB task, suggesting that these two functions appeared to be key elements that are required to pass the FB task. Although, the flexible shifting of the focus of attention as measured by the DCCS task seems to be critical for passing the FB task, it is not essential for the picture-sequencing task. In contrast, inhibitory control as measured by the pointing-stroop task appeared to contribute to both tasks (FB and picture sequencing). Thus, some forms of inhibitory control are likely to be involved in responding to the general executive demands of both tasks. This interpretation accords with the claim that relationships between theory-of-mind and executive function tasks are attributable to more than just a common problem of inhibitory control, and that different aspects of EFs are needed to explain their developmental relationships (Perner et al., 2002). The executive function tasks in the present study were Stroop (inhibitory control) and DCCS (shifting the focus of attention), which were administered using an analogous format of touching a PC screen. Therefore it is unlikely that differential effects of DCCS and Stroop on the FB and picture sequencing tasks derived from any procedural differences.

With respect to the 8 sources of information framework: 8-SIF (Achim et al., 2013), the false-belief task includes additional linguistic sources of information for the agent and for the context. Extra linguistic information included in the FB task may be helpful in guiding the participants to impute the agent’s mental states. However, this source of information could come at a cost because of the linguistic information about the changes in the context resulting from a translocation of the object. It could inadvertently direct the participants’ attention to the more salient (the true belief) location rather than the agent’s (false) belief, without a volitional flexible shift of attention. Thus the participants may require an important shift in the focus of attention from the context they hear last. Unique variance in inhibitory control and shift of attention that help explain the FB scores in the model support this interpretation. These findings suggest that a higher level of EFs is necessary to pass the FB task than to pass picture sequencing.

Is it possible that the extra linguistic information given in the false-belief task could cause the participants to incorrectly impute the mental states of the agent? When the FB task is given without a verbal scenario to young children (Southgate, Senju, & Csibra, 2007), they performed well. In such a nonverbal FB task, the participants do not need to process linguistic information and therefore do not need to make an effortful shift in the focus of attention that may be required in the verbal FB task. Under such circumstances, the younger children were imputing the agent's mental state without interference from linguistic information. When verbal processing is involved, however, the linguistic representation could make it harder for the participants to shift their focus of attention to the original (falsely believed) location. Previous findings from adult participants whose implicit false-belief inferences were negatively affected by just listening to unrelated linguistic information (Schneider et al., 2012) raise the possibility that even a low-level of cognitive load could degrade a participants' performance.

In terms of the relationship with receptive language, both the false-belief and picture-sequencing tasks shared a similar degree of variance with linguistic competence. These results suggest that language is an important factor (Astington & Jenkins, 1999), even though the picture-sequencing task did not include a linguistic source of information. The picture-sequencing task is likely to require self-guided linguistic representations in order to interpret perceptual information and to construct a causal or intentional event that makes sense in the physical and social world.

The emerging picture for the FB task has important implications for the interpretation of existing studies in regard to the development of mentalising. For the relationship between executive functions (EFs) and theory of mind (ToM), EFs seems to be critical when mentalising ability is measured using FB tasks, whereas there is no clear relationship when other mentalising tasks are used, such as the intentional version of the picture-sequencing task. More precisely, the source of information included in mentalising seems to account for the possibility of finding a relationship between EFs and ToM. Differences in cognitive load in mentalising tasks, such as inclusion (high-load) or exclusion (low-load) of false belief, did not change their relationships with EFs (Carlson et al., 2015). Although cognitive load was manipulated, the tasks used by Carlson, Claxton and Moses (2015) use both linguistic and perceptual information, and are similar to each other with respect to the source of information based on 8-SIF (Achim et al., 2013). This seems to suggest that analyses of the tasks in terms of the sources of information are important when making a relevant selection of mentalising tasks.

This argument resonates with the present finding of a close association between the intentional and mechanical versions of the picture-sequencing task. From the original study (Baron-Cohen et al., 1986), intentional and causality versions of the picture-sequencing task were expected to elicit different levels of performance in children. Baron-Cohen et al. (1986) depicted atypical characteristics of mentalising ability in ASD children in comparison with typically developing children. However, it was not clear how typically developing children performed differently in the intentional and mechanical versions of the task. The present study found that task differences in the contents (intentional/ mechanical) may be very subtle when compared with the robustness of the executive processing required by the characteristics of the picture-sequencing task. Thus processing the sources of information in the picture-sequencing task is similar for intentional and mechanical versions, and differences in content had little impact on performance.

Conclusion

The false-belief (FB) task has been used in a large number of studies that investigate the development of theory of mind (ToM), and has almost become a signature task for the measurement of this ability. However, the path analyses in this study that examined the relationship between the FB task, the picture-sequencing task and executive functions suggests that passing the FB task requires considerably more than just mentalising ability, as claimed by Bloom and German (2000). Assessing the development of mentalising abilities (ToM) with a variety of mentalising tasks, giving due consideration to the source of information, could provide a better approach to understanding the development of this ability.

References

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On what ground do we mentalize? Characteristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment, 25*(1), 117.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relationship between language and theory-of-mind development. *Developmental Psychology, 35*, 1311-1320.

Barkley, R. A. (2012). *Executive Functions: What they are, how they work and why they evolved*. NY: Guilford Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition, 21*(1), 37-46. doi:http://dx.doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and Intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology, 4*(2), 113-125. doi:10.1111/j.2044-835X.1986.tb01003.x

Berger, A., Jones, L., Rothbart, M., & Posner, M. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. doi:10.3758/BF03207798

Blair, C. (2016). Developmental Science and Executive Function. *Current Directions in Psychological Science, 25*(1), 3-7. doi:10.1177/0963721415622634

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*(1), B25-B31. doi:http://dx.doi.org/10.1016/S0010-0277(00)00096-2

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The Relation Between Executive Function and Theory of Mind is More Than Skin Deep. *Journal of Cognition and Development, 16*(1), 186-197. doi:10.1080/15248372.2013.824883

Devine, R. T., & Hughes, C. (2014). Relations Between False Belief Understanding and Executive Function in Early Childhood: A Meta-Analysis. *Child Development, 85*, 1777-1794. doi:10.1111/cdev.12237

Gopnik, A., & Wellman, H. M. (1992). Why the Child's Theory of Mind Really Is a Theory. *Mind & Language, 7*(1-2), 145-171. doi:10.1111/j.1468-0017.1992.tb00202.x

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children Second Edition*. Circle Pines, MN: American Guidance Service.

Lewis, C., & Osborne, A. (1990). Three - Year - Olds' Problems with False Belief: Conceptual Deficit or Linguistic Artifact? *Child Development*, *61*(5), 1514-1519.

Mitchell, P., & Lacohée, H. (1991). Children's early understanding of false belief. *Cognition, 39*(2), 107-127.

Moore, C., Pure, K., & Furrow, D. (1990). Children's Understanding of the Modal Expression of Speaker Certainty and Uncertainty and Its Relation to the Development of a Representational Theory of Mind. *Child Development, 61*(3), 722-730.

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

Russell, J. (1996). *Agency its Role in Mental Development*.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*(2), 331-349. doi:10.1111/j.2044-835X.1991.tb00881.x

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive Load Disrupts Implicit Theory-of-Mind Processing. *Psychological Science, 23*(8), 842-847. doi:10.1177/0956797612439070

Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition, 38*(1), 1-12.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 17*, 587-592.

Ueno, K., Nakoshi, S., & Onuki, S. (2008). *Picture Vocabulary Test-Revised*. Tokyo: Nihon Bunka Kagakusya.

van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology, 31*(1), 70-87. doi:10.1111/j.2044-835X.2012.02079.x

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.

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.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nat. Protocols, 1*(1), 297-301.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH TOOLBOX COGNITION BATTERY (CB): MEASURING

EXECUTIVE FUNCTION AND ATTENTION. *Monographs of the Society for Research in Child Development*, *78*(4), 16-33. doi:10.1111/mono.12032

Table 1. Descriptive statistics by age

Age groups		3-year	4-year	5-year	6-year	Total
Age		n = 30	n = 47	n = 54	n = 35	n =166
(42-78 months)	M	45.60	54.70	64.89	74.63	60.57
	SD	1.54	3.65	3.29	2.02	10.35
Picture sequencing						
Mechanical	M	1.17	1.85	3.02	3.63	2.48
(0-4)	SD	1.34	1.60	1.37	.84	1.61
Intentional	M	1.17	2.00	2.98	3.57	2.50
(0-4)	SD	1.23	1.49	1.30	.98	1.53
Executive functions						
Stroop	M	4.17	5.83	7.07	7.54	6.30
(0-8)	SD	3.30	2.88	2.04	1.29	2.71
DCCS	M	2.67	3.60	4.91	5.23	4.20
(0-6)	SD	2.50	2.40	1.78	1.61	2.28
WM	M	7.83	7.85	9.37	10.06	8.81
(0-15)	SD	2.39	2.21	2.31	1.53	2.33
False-beliefs						
(0-2)	M	.17	.94	1.44	1.66	1.11

	SD	.53	.89	.82	.73	.93
PVT						
(3-60)	M	12.40	16.15	23.81	31.09	21.11
	SD	7.87	7.21	10.38	10.71	11.31

Note. Ranges are indicated in parenthesis

Table 2. Pearson's correlation coefficients between the measurements and age

	1	2	3	4	5	6	7
1. Age	-						
2. PVT	.62**	-					
3. WM	.43**	.56**	-				
4. Stroop	.45**	.49**	.37**	-			
5. DCCS	.45**	.46**	.40**	.53**	-		
6. Mechanical	.59**	.54**	.39**	.44**	.38**	-	
7. Intentional	.59**	.53**	.39**	.43**	.41**	.61**	-
8. False-belief	.56**	.59**	.41**	.56**	.52**	.58**	.55**