

Why do Young Children Fail in False Belief Tasks:
Linguistic Representations and Implicit Processing

by

Li Yi

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

Gary Feng, Supervisor

Reiko Mazuka

Kevin Pelphrey

Misha Becker

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Psychology and Neuroscience in the Graduate School
of Duke University

2009

ABSTRACT

Why do Young Children Fail in False Belief Tasks:
Linguistic Representations and Implicit Processing

by

Li Yi

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

Gary Feng, Supervisor

Reiko Mazuka

Kevin Pelphrey

Misha Becker

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Psychology and Neuroscience in the Graduate School
of Duke University

2009

Copyright by
Li Yi
2009

Abstract

Despite recent evidence that infants under one year of age have implicit understanding of theory of mind, three-year-old children repeatedly fail in traditional false belief tasks. A series of 4 studies investigated two possible sources of errors. First, children's comprehension of theory of mind questions was tested in an elicited imitation task. Second, their understanding of mental events was measured using anticipatory eye movements in non-verbal tasks. Results showed that young children's performance in verbal false belief tasks is limited by their understanding of linguistic representations of beliefs and their ability to monitor mental states in real-time. This implies the limitations of young children in keeping track of complex social events in real time and in understanding language conventions in real time.

Contents

Abstract	iv
List of Tables	ix
List of Figures	x
1. Introduction	1
1.1 Performance in Traditional Theory of Mind Tasks	8
1.1.1 Traditional Theory of Mind tasks	8
1.1.2 Challenges to the Traditional ToM Tasks	9
1.2 Non-verbal ToM reasoning	11
1.2.1 Theories on Non-verbal ToM reasoning	11
1.2.2 Evidence for Non-verbal ToM reasoning	13
1.2.3 Challenges to Onishi-Baillargeon Study	17
1.2.4 Design of the Current Project	22
1.3 Linguistic Representations of Theory of Mind	23
1.3.1 Linguistic Markers of Theory of Mind	23
1.3.2 Acquisition of ToM markers	26
1.3.3 Linguistic Representations of ToM and ToM Performance	27
1.3.4 The Current Study	31
1.4 Other Issues Addressed in the Present Studies	33
1.4.1 Second- and Higher-order Theory of Mind	33
1.4.2 Cultural Differences	38

1.5 Overview of the Present Studies	38
2. Study 1: Children’s Understanding of First-order Beliefs.....	40
2.1 Method	41
2.1.1 Participants.....	41
2.1.2 Material.....	42
2.1.3 Procedure.....	51
2.2 Results and Discussions.....	51
2.2.1 Verbal Answers to ToM Questions.....	51
2.2.2 Elicited imitation	54
2.2.3 Eye Movements in the Change-of-Location Task	58
2.2.4 Eye Movements in Unexpected-content Task	80
2.3 General Discussion for study 1.....	102
3. Study 2: Children’s Interpretations of ToM Questions.....	104
3.1 Methods	105
3.1.1 Participants.....	105
3.1.2 Material.....	105
3.1.3 Procedure.....	106
3.2 Results	108
3.3 Discussion.....	110
4. Study 3: Understanding of Second-Order Beliefs	113
4.1 Method	113
4.1.1 Participants.....	113

4.1.2 Material.....	114
4.1.3 Procedure.....	117
4.2 Results.....	118
4.2.1 Verbal answers to ToM questions.....	118
4.2.2 Elicited imitation.....	120
4.2.3 Eye Movements.....	122
4.3 Discussion.....	134
5. Study 4: Chinese Children’s Interpretations of ToM Questions and Implicit ToM	137
5.1 Methods.....	137
5.1.1 Participants.....	137
5.1.2 Material and Procedure.....	138
5.2 Results.....	139
5.2.1 Elicited imitations.....	139
5.2.2 Verbal answers to ToM questions.....	140
5.2.3 Eye movements in the Change-of-location task.....	143
5.2.4 Eye movements in the unexpected-content task.....	149
5.3 Discussion.....	153
6. General Discussion.....	156
6.1 Performance in Traditional ToM Tasks.....	158
6.2 Non-verbal ToM Reasoning.....	158
6.3 Linguistic Representations of ToM.....	160
6.4 Development of ToM.....	162

6.5 ToM, Language, and Culture.....	164
6.6 Conclusions	165
References	167
Biography	174

List of Tables

Table 1: Descriptions of each scene of the videos used in the unexpected-content task of study 1	49
Table 2: Descriptions of each scene of the videos used in study 3.....	115

List of Figures

Figure 1: Two processes influencing children’s ToM performance	5
Figure 2: Correct rates of all ages answering different types of questions in the change-of-location task of study 1.....	52
Figure 3: Correct rates of all ages answering different types of questions in the unexpected-content task	53
Figure 4: Percentages of making errors in ToM markers and other errors in the elicited imitation task	54
Figure 5: Mean looking time to the left and right in the TB trial of study 2	60
Figure 6: Eye movements of all age groups across time in the true belief trial of study 261	
Figure 7: Eye movements of adults across time in the critical moment in the true belief trial of study 2.....	62
Figure 8: Eye movements of 4-year-olds across time in the critical moment in the true belief trial of study 2	62
Figure 9: Eye movements of 3-year-olds across time in the critical moment in the true belief trial of study 2	63
Figure 10: Mean looking time to the left and right in the first FB trial of study 2	64
Figure 11: Eye movements of all age groups across time in the false belief non-complete trial of study 2.....	65
Figure 12: Eye movements of adults across time in the critical moment in the false belief non-complete trial of study 2	66
Figure 13: Eye movements of 4-year-olds across time in the critical moment in the false belief non-complete trial of study 2.....	66
Figure 14 Eye movements of 3-year-olds across time in the critical moment in the false belief non-complete trial of study 2.....	67
Figure 15: Mean looking time to the left and right in the second FB trial of study 2	68

Figure 16: Eye movements of all age groups across time in the false belief complete trial of study 2.....	69
Figure 17: :Eye movements of adults across time in the critical moment in the false belief complete trial of study 2.....	70
Figure 18: Eye movements of 4-year-olds across time in the critical moment in the false belief complete trial of study 2.....	70
Figure 19: Eye movements of 3-year-olds across time in the critical moment in the false belief complete trial of study 2.....	71
Figure 20: Eye movements of 4-year-olds across time in the false belief non-complete trial (pass – orange line; fail-purple line).....	72
Figure 21: Eye movements of 4-year-olds who passed the traditional ToM questions across time in the false belief non-complete trial of study 2.....	74
Figure 22: Eye movements of 4-year-olds who failed the traditional ToM questions across time in the false belief non-complete trial of study 2.....	74
Figure 23: Eye movements of 4-year-olds who passed the elicited imitation task across time in the false belief non-complete trial of study 2.....	75
Figure 24: Eye movements of 4-year-olds who failed the elicited imitation task across time in the false belief non-complete trial of study 2.....	76
Figure 25: Mean looking time at the crayon and candle in the first FB trial of the unexpected-content task of study 2.....	81
Figure 26: Eye movements of all age groups across time of the critical moment in the familiarization trial of study 1.....	82
Figure 27: Eye movements of adults per 100 ms of the critical moment in the familiarization trial of study 1.....	83
Figure 28: Eye movements of 4-year-olds per 100 ms of the critical moment in the familiarization trial of study 1.....	84
Figure 29: Eye movements of 3-year-olds per 100 ms of the critical moment in the familiarization trial of study 1.....	84

Figure 30: Mean looking time at the crayon and candle in the second part of the FB-FB condition of the unexpected-content task of study 1	85
Figure 31: Eye movements of all age groups across time of the critical moment in the second part of the FB-FB condition of study 1	86
Figure 32: Eye movements of adults per 100 ms of the critical moment in the second part of the FB-FB condition of study 1	87
Figure 33 Eye movements of 4-year-olds per 100 ms of the critical moment in the second part of the FB-FB condition of study 1	88
Figure 34: Eye movements of 3-year-olds per 100 ms of the critical moment in the second part of the FB-FB condition of study 1	88
Figure 35: Mean looking time at the crayon and candle in the first part of the FB-TB condition of the unexpected-content task of study 1	89
Figure 36: Eye movements of all age groups across time of the critical moment in the first part of the FB-TB condition of study 1	90
Figure 37: Eye movements of adults per 100 ms of the critical moment in the first part of the FB-TB condition of study 1	91
Figure 38: Eye movements of 4-year-olds per 100 ms of the critical moment in the first part of the FB-TB condition of study 1	92
Figure 39: Eye movements of 3-year-olds per 100 ms of the critical moment in the first part of the FB-TB condition of study 1	92
Figure 40: Mean looking time at the crayon and candle in the second part of the FB-TB condition of the unexpected-content task of study 1	93
Figure 41: Eye movements of all age groups across time of the critical moment in the second part of the FB-TB condition of study 1	94
Figure 42: Eye movements of adults per 100 ms of the critical moment in the second part of the FB-TB condition of study 1	95
Figure 43: Eye movements of 4-year-olds per 100 ms of the critical moment in the second part of the FB-TB condition of study 1	96

Figure 44: Eye movements of 3-year-olds per 100 ms of the critical moment in the second part of the FB-TB condition of study 1	96
Figure 45: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the belief question	98
Figure 46: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the belief question	98
Figure 47: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the overall ToM questions.....	99
Figure 48: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the overall ToM questions.....	100
Figure 49: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the elicited imitation task	101
Figure 50: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the elicited imitation task	101
Figure 51: Percentages of making errors in ToM markers and other errors in the elicited imitation task in study 2.....	109
Figure 52: Correct rates of answering the prediction questions in the three conditions of study 3	119
Figure 53: Correct rates of answering the questions with one or two ToM markers in the three conditions	120
Figure 54: percentages of making errors in ToM markers and other errors in the elicited imitation task of study 3.....	121
Figure 55: Mean looking time at the crayon and candle in the first-order FB trial of study 3	123
Figure 56: Eye movements of all age groups across time of the critical moment in the first-order FB condition of study 3	124
Figure 57: Eye movements of adults per 100 ms of the critical moment in the first-order FB condition of study 3	125

Figure 58: Eye movements of 6-year-olds per 100 ms of the critical moment in the first-order FB condition of study 3.....	125
Figure 59: Eye movements of 5-year-olds per 100 ms of the critical moment in the first-order FB condition of study 3.....	126
Figure 60: Mean looking time at the crayon and candle in the FB-TB trial of study 3....	127
Figure 61: Eye movements of all age groups across time of the critical moment in the first-order FB trial.....	128
Figure 62: Eye movements of adults per 100 ms of the critical moment in the FB-TB FB condition of study 3.....	129
Figure 63: Eye movements of 6-year-olds per 100 ms of the critical moment in the FB-TB FB condition of study 3.....	129
Figure 64: Eye movements of 5-year-olds per 100 ms of the critical moment in the FB-TB condition of study 3.....	130
Figure 65: Mean looking time at the crayon and candle in the second-order belief trial of study 3.....	131
Figure 66: Eye movements of all age groups across time of the critical moment in the second-order belief condition of study 3.....	132
Figure 67: Eye movements of adults per 100 ms of the critical moment in the second-order belief condition of study 3.....	133
Figure 68: Eye movements of 6-year-olds per 100 ms of the critical moment in the second-order belief condition of study 3.....	133
Figure 69: Eye movements of 5-year-olds per 100 ms of the critical moment in the second-order belief condition of study 3.....	134
Figure 70: Percentages of making errors in ToM markers and other errors in the elicited imitation task in study 4.....	140
Figure 71: Correct rates of all ages answering different types of questions in the change-of-location task of study 4.....	141

Figure 72: Correct rates of all ages answering different types of questions in the unexpected-content task of study 4.....	142
Figure 73: Mean looking time to the left and right in the TB trial of study 4.....	143
Figure 74: Eye movements of all age groups across time in the true belief trial of study 4.....	144
Figure 75: Mean looking time to the left and right in the second FB trial of study 4.....	145
Figure 76: Eye movements of all age groups across time in the false belief complete trial of study 4.....	147
Figure 77: Mean looking time to the left and right in the first FB trial of study 4.....	148
Figure 78: Eye movements of all age groups across time in the false belief non-complete trial of study 4.....	149
Figure 79: Mean looking time at the crayon and candle in the FB trial of the unexpected-content task of study 4.....	150
Figure 80: Eye movements of all age groups across time of the critical moment in the FB trial of study 4.....	151
Figure 81: Mean looking time at the crayon and candle in the FB-TB trial of the unexpected-content task of study 4.....	152
Figure 82: Eye movements of all age groups across time of the critical moment in the FB-TB trial of study 4.....	153

1. Introduction

Theory of Mind (ToM) refers to a system of inferences used to impute mental states of oneself or others in order to explain and predict behaviors (Premack & Woodruff, 1978). Dennett (1978) claimed that the minimal justification of a person possessing a theory of mind is that he/she holds beliefs of a simple factual nature. The “acid test” of Theory of Mind, the false belief tasks, measure children’s predictions or explanations of others’ behaviors based on false beliefs (Frith & Frith, 1999). Preschool children’s performance in the false belief tests was found to be positively correlated with their social interaction skills (Astington & Jenkins, 1995; Dunn & Cuting, 1999; Lalonde & Chandler, 1995; Watson, Nixon, Wilson, & Capagne, 1999), understanding of deception (Peterson & Siegal, 2002), and moral development (Baird & Astington, 2004). Moreover, preschoolers’ understanding of false belief predicts their later social understanding skills and peer relations (Astington, 2003).

Traditionally children’s understanding of false beliefs is typically assessed using change-of-location and unexpected-content tasks (Gopnik & Astington, 1988; Wimmer & Hartl, 1991; Wimmer & Perner, 1983). In these tasks, children are verbally presented a story in which the protagonist holds a false belief, and are asked – in most cases verbally – to either predict the protagonist’s behavior or mental state. Children’s answers to these questions are taken as a measure of their ToM competence. Using these verbal tests,

researchers found that the understanding of others' false beliefs emerge in preschool years (Gopnik & Astington, 1988; Wimmer & Hartl, 1991; Wimmer & Perner, 1983).

Children under 4 year old failed to predict others' behaviors based on false beliefs, while most 5-year-old children succeed in the tasks. This pattern has been replicated in hundreds of studies and consistently (Wellman, Cross, & Watson, 2001). This age gap of ToM development has been the source of a lot of discussion.

In the past 20 years, theories on Theory of Mind (ToM) have been developed to explain the age shift of understanding false belief in preschool years. A dominant theory claimed that children construct a theory-like understanding of mind (Perner, 1991; Wellman, 1990). Preschool children experience a fundamental change from a "mentalistic theory of behavior" to a "representational theory of mind" (Perner, 1991). In contrary, theorists holding the simulation perspective attribute 3-year-old children's failure in the traditional tasks to their limited cognitive abilities in simulating others' mental states (Gordon, 1986; Greenwood, 1999; Harris, 1989).

Recent evidence showed that infants as young as 15 month old showed early understanding of false beliefs revealed by their looking time (Onishi & Baillargeon, 2005; Southgate, Senju & Csibra, 2007). These studies challenged the traditional view of the development of ToM and may suggest an innate basis of ToM (Scholl & Leslie, 1999).

One challenge the infants' research on ToM received is the validity of the task: infants' responses may be based on behavioral rules instead of understanding mental states (Perner & Ruffman, 2005; Ruffman & Perner, 2005). The issue whether the implicit task is a measure for ToM competence is still under debate (Csibra & Southgate, 2006).

This new trend of research on ToM also suggests a reconsideration of the debate on the development of ToM, especially the gap between 3 and 5 year olds' ToM in traditional false belief tasks. In other words, why do 3-year-olds still fail in the traditional false belief tasks considering they already acquire the concept of mind in 15 months? I hypothesized that, if infants obtain ToM competence, the performance of 3-year-olds in the traditional tasks should be due to the following two aspects: their real time processing of mind situations involved in the false belief tasks, and their interpretations of the mind questions.

In order to pass traditional false belief tests (Gopnik & Astington, 1988; Wellman, et al., 2001; Wimmer & Perner, 1983), children have to succeed in the following two tasks: they must keep track of the mental state of protagonists as they witness a social event, and they must be able to correctly comprehend questions about mental states and respond appropriately, both of which have to be achieved in real-time (see Figure 1). The first component, ToM reasoning, is an automatic process to encode people's actions in terms of mental states and other aspects of the ToM. It develops from the early

understanding of false beliefs showed in infants' research (Onishi & Baillargeon, 2005; Southgate, et al, 2007) and does not necessarily involve language. However, language can be used in reasoning about others' mental states, revealing different underlying processes of mental states from the nonverbal and automatic ToM reasoning. Linguistic representations of ToM are required in communicating messages about mental states. In the context of traditional false belief tests, children must correctly interpret questions about mental states as such. Failing to encode real-time events in a framework of mental states will result in erroneous answers in traditional ToM tasks, so is a misunderstanding of verbal questions that come after the story. The primary goal of the current study is to investigate the source of young children's difficulties with traditional ToM tasks, i.e., whether they have troubles with ToM reasoning in real-time, or with linguistic representations of ToM.

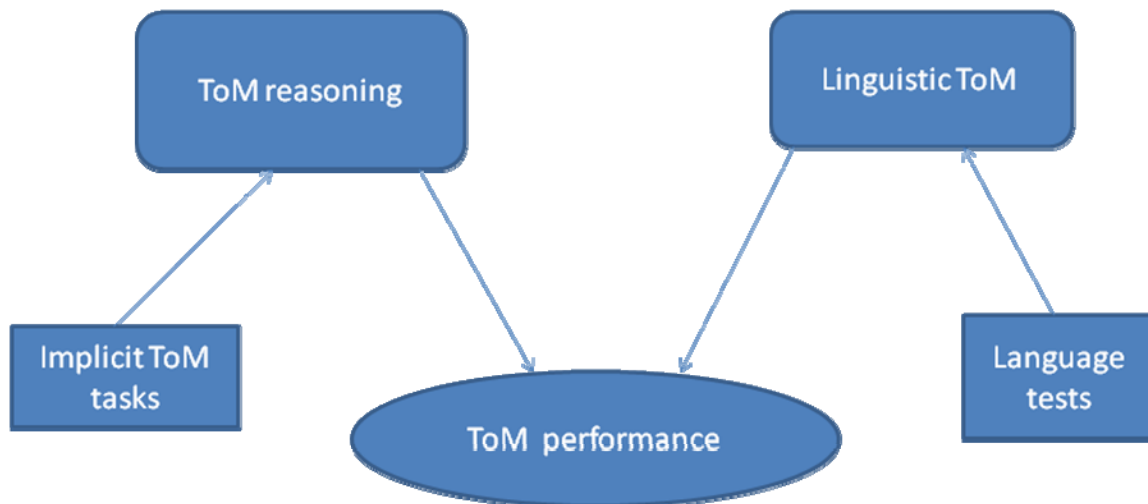


Figure 1: Two processes influencing children’s ToM performance

The notion that the ToM includes an automatic component and a linguistic aspect is broadly consistent with current theories of ToM. A number of theories proposed an independent component of ToM that may be “unverbalizable”, “unconscious” and prelinguistic (Clements & Perner, 1994; Onishi & Baillargeon, 2005; Scholl & Leslie, 1999; Southgate, et al, 2007; Tager-Flusberg, 2001). Clements & Perner (1994) made a distinction between “implicit” and “explicit” ToM, and developmentally the implicit understanding of beliefs precedes the explicit understanding of beliefs. Liu and colleagues (2004) proposed that processing of mental states is very fast, in real time, and automatic, and they used ERP to measure this automatic process (Liu, Sabbagh, Gehring, & Wellman, 2004). More details about previous theories and research in this field will be reviewed in Section 1.2 of this chapter.

There is also a large literature on the relation between ToM and language (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 2003; Lee, et al., 1999; Lohmann & Tomasello, 2003; Milligan, et al., 2007; Tardiff, et al., 2004). A meta-analysis showed that children's performance in various traditional false-belief tasks is related to their language abilities (Milligan, et al., 2007). Directly relevant to the present study, prior research has shown that certain linguistic structures – lexical items (mental verbs such as *think* and *know*) and syntactic structures (such as the tensed sentential complement as in *She thought he was in Paris*) – are consistently used to mark mental states (de Villiers & de Villiers, 2000). In fact, communicating ideas about complex mental states such as a false belief would be very difficult without such linguistic devices. This implies that young children must acquire these linguistic ToM markers in order to succeed in traditional ToM tasks, which almost invariably involve verbal questions about mental states. Evidence of the relation between ToM and language, particularly the importance of linguistic markers of ToM, will be reviewed in Section 1.3.

This study introduces two methodological innovations. First, to track children's real-time ToM reasoning, an eye-movement-based paradigm was developed, in which events that involve mental state changes were shown in animations and videos, with minimal linguistic requirements. Critically, there is a period of time in each video when participants are implicitly cued to make anticipatory eye movements. Based on their

understanding of the event, participants will look at a particular location on the screen sooner and longer. This implicit measure of ToM reasoning helps to disentangle the linguistic demand from traditional ToM tasks.

In addition, an elicited imitation task was introduced to assess participants' linguistic representations of ToM. Elicited imitation is usually used to assess children's language acquisition, especially for acquisitions of precise syntactic structures (Ambridge & Pine; 2006; Keller-Cohen, 1981; Lust, Flynn & Foley, 1996). The task involves listening to an utterance and repeating it back. The current task is also designed to overload the working memory and central processing capacity of participants, so that they will make errors. Misremembering a name or an action does not change the gist of the sentence. However, dropping a linguistic marker of ToM will necessarily change the nature of the sentence. The interesting question is: do children and adults treat the ToM markers as any other words, or do they try to preserve these markers at the expense of other words?

In Chapter 1, I plan to review the main questions asked in this paper and previous attempts to answer these questions. Chapter 2 investigates whether children do ToM reasoning in non-verbal ToM tasks, and how well children understand ToM markers in ToM questions, especially whether they pay attention to the ToM markers when they hear questions and statements about mind. Meanwhile, their performance in

the traditional ToM questions will be correlated with their non-verbal ToM reasoning and linguistic ToM abilities. Chapter 3 tests how well children understand ToM markers in questions used in previous ToM research. Chapter 4 examines how children understand others' second-order beliefs. Chapter 5 replicated the studies in Chapter 2 with Chinese-speaking participants. Chapter 6 is a general discussion section.

1.1 Performance in Traditional Theory of Mind Tasks

1.1.1 Traditional Theory of Mind tasks

An understanding of false belief requires making a distinction between mental states and reality. And this is the essence of traditional False Belief tasks. In a classic false belief study, the change-of-location task (also known as Sally-Ann task), Wimmer and Perner (1983) told the following story to preschool children: Maxi, a young boy, puts some chocolate in the kitchen cupboard and leaves the room. While he is away (and cannot see) his mother moves the chocolate from the cupboard to a drawer. After he returns, children are asked "where will Maxi look for the chocolate?" Children older than 4 typically answer the cupboard, even though the chocolate is, in fact, in the drawer. Younger children often fail this task, judging that Maxi will look for the chocolate in the drawer, the physical location the object is last seen. This finding suggested that Theory of Mind is acquired gradually, and the critical transition happens around age 4.

Another widely used false belief task is the unexpected-content task, also known as the Smarties task (Gopnik & Astington, 1988; Saltmarsh, Mitchell, & Robinson, 1995; Wimmer & Hartl, 1991). Children are showed a crayon box filled with candies and asked what another person, who has never looked inside the box, will think is inside the crayon box (“What will Mary think is in here, crayons or candies?”). Findings using this task are generally consistent with the change-of-location task (Gopnik & Astington, 1988; Wellman, et al., 2001).

The change-of-location task and the unexpected-content task are among the most widely used assessments of children’s ToM. A meta-analysis by Wellman, et al., (2001) showed that preschool children follow a consistent developmental pattern from below-chance performance to above-chance performance in the traditional ToM tasks, regardless of the type of task, type of question, nature of the protagonist, or the nature of the target object. Findings from the meta-analysis suggested that children do not acquire an understanding of false belief until 4 years old. However, the notion that children’s performance in these traditional ToM task directly reflects their underlying ToM competence does not go unchallenged.

1.1.2 Challenges to the Traditional ToM Tasks

Although children’s performance in the traditional false belief tasks are consistent regardless of types of tasks and questions, there are some factors that can

significantly improve their performance: making the motive of the protagonist explicit, not showing or destroying the object, making the protagonist's mental states salient, showing the child a picture representing the protagonist's belief (Wellman, et al, 2001). This suggests that children's performance in the false belief tasks may be modified by some experimental factors, and if these factors are manipulated to simplify the task, the understanding of false belief may be found in younger children.

A strong criticism is that false belief tasks are highly language-dependent so that passing the traditional false belief tasks may require additional information processing resources, more than the capacity of Theory of Mind (Bloom & German, 2000). Children have to accurately interpret the meanings of the instructions and questions in order to answer them (Bloom & German, 2000). Although many of the false belief studies used dolls, real objects, cartoons, or animations as visual aids in presenting the task, every one of the 178 studies reviewed in Wellman, et al., (2001) used verbal instructions and questions. Therefore, young children's competence of Theory of Mind may be masked by the linguistic complexity of the story and the questions (de Villiers & de Villiers, 2000). That is, their failure in the false belief task may be attributable to language deficits but not Theory of Mind deficits.

There are two approaches to disentangle the relations between language and ToM. One is to develop tasks that are independent of language, which has the added

benefit of allowing testing ToM understanding among pre-verbal infants (Clements & Perner, 1994; Onishi & Baillargeon, 2005; Southgate, et al., 2007). The other approach is to try to explicate the nature of the dependency between language and ToM (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 2003; Hughes, 1998; Lee, et al., 1999; Lohmann & Tomasello, 2003; Milligan, et al, 2007; Roth & Leslie, 1998; Tager-Flusberg & Sullivan, 2000; Tardiff, et al., 2004). The following two sections go into more details in each approach.

1.2 Non-verbal ToM reasoning

1.2.1 Theories on Non-verbal ToM reasoning

There is increasing evidence that children show non-verbal ToM processes in implicit measures (Clements & Perner, 1994; Scholl & Leslie, 1999; Tager-Flusberg, 2001), although the origins of this non-verbal ToM is still under debate.

According to some theorists children are endowed with the ability to automatically encode social events in a mentalistic framework. Scholl & Leslie (1999) proposed an “innate encapsulated, and domain-specific” module for Theory of Mind, the Theory of Mind Mechanism (ToMM). ToMM is a brain mechanism specialized to allow children to attend to mental states. Examples of such a mechanism include eye-gaze following, (Butterworth & Cochran, 1980; Butterworth & Grover, 1990) and attending to pointing gesture (Povinelli, 1997). This theory suggests an automatic

process in children to encode social information, which may build the foundation for the future ToM process.

Automatic ToM reasoning is also an important component in the componential view of Theory of Mind by Tager-Flusberg (2001), Tager-Flusberg and Sullivan (2000) and Tager-Flusberg and Joseph (2005). According to this view ToM includes two distinct components: the social-perceptual component and the social-cognitive component. The social-perceptual component of Theory of Mind is based on the innate tendency to attend to the social information in the environment, such as face, voice, and body gesture. Understanding of other people's minds develops from this innate mechanism and through social interaction with other people. The social-cognitive component of Theory of Mind is a representational understanding of mind, based on not only perceptual cues but also from the sequence of events over time. This social-cognitive component of Theory of Mind develops from the social-perceptual component of Theory of Mind during the preschool years.

Similarly, Clements and Perner (1994) defined the concept of "implicit" ToM as "unconscious" and "unverbalizable" knowledge about other people's mental state. The implicit ToM can be measured in non-verbal tasks involving ToM reasoning, while explicit ToM is the traditionally measured ToM. Their studies showed that this implicit ToM precedes explicit ToM.

1.2.2 Evidence for Non-verbal ToM reasoning

One of the earliest attempts to measure non-verbal ToM reasoning was by Clements & Perner (1994). In the *anticipatory looking* paradigm, children listened to a typical false belief story and watched the scene being played out on a cardboard. The agent's (Sam, the mouse) movements were invisible when it went through a V-shaped occluder until it reappeared at one of the two exits (mouse holes). After as the prompt "I wonder where he's going to look?", the experimenter gave children 1-2 seconds to anticipate Sam's reappearance at one of the two mouse holes. Children's anticipatory looking was videotaped and analyzed frame-by-frame during the anticipation period. Results showed that children between 2 years 5 months and 2 years 10 months looked longer at the location where the object really was, suggesting neither implicit nor explicit understanding of false belief. This was consistent with their verbal answers. Most of the children between 2 years 11 months and 4 years gave the real location as the answer to the verbal question, but they anticipated the protagonist to go to where the agent thought the object was, suggesting an implicit but not explicit understanding of false belief. This distinction between implicit and explicit understanding of false belief suggested the onset of implicit form of understanding of mind is about 2 years 11 months, earlier than what traditional ToM tasks show.

Although the measurement of implicit knowledge, the anticipatory looking, itself is independent of language in Clements & Perner's (1994) study, the stories were told verbally, and a verbal prompt was given before children made the anticipatory looking. Therefore, children's responses may be triggered by the verbal instructions or questions they heard. Therefore, the measurement for implicit understanding of false belief used in Clements and Perner's (1994) study is not fully independent of language.

Onishi & Baillargeon (2005) devised a completely nonverbal false belief task, in which 15-month-old infants watched an event similar to the story told in the traditional false belief tasks. The violation-of-expectation paradigm was used to assess infants' looking time. Infants were randomly assigned into three conditions. In the first condition, they were shown an event in which (1) an actor grasped one of two toys (the target) from two boxes, played with it for a while, and put it back in the original location; (2) then the two toys exchanged locations without being witnessed by the actor, (3) the actor returned and reached for the target toy (unexpected condition) or non-target toy (expected condition). When the toys were in the boxes, they were invisible to children. In another false belief (FB) condition the actor witnessed a location change of the target, and then without the actor's witness, another location change happened which made the target go back to the original location. In the true-belief (TB) condition, the actor witnessed the location change (or no change). If infants expected the actor to search for

the toy according to her beliefs (true or false) about its location, regardless of its actual location, then they should look reliably longer when that expectation was violated.

Their results confirmed this prediction by showing that 15-month-old infants expected the actor to reach where she believed the toy should be, and infants looked longer when the reaches did not match their expectations (e.g. reaching to the actual location in the FB condition) than when reaches matched their expectations. According to Onishi & Baillargeon (2005), this showed that 15-month-old infants can (1) distinguish the perception of their own and others; (2) track what the actor can see and cannot see; (3) understand that the actor's perception, but not their own, guides her actions. This study provides some of the first evidence for early understanding of false belief.

Onishi & Baillargeon (2005)'s finding was confirmed by Southgate, et al., (2007) with 2-year-old infants using similar design. Instead of using Onishi & Baillargeon (2005)'s violation-of-expectation paradigm, Southgate, et al., (2007) used an anticipatory-looking paradigm, in which the participants' anticipatory looking indicates their expectations of the agent's actions. They showed 25-month-old children videos in which an actor witnessed a toy being hidden, and then the toy was later removed by a puppet while the actor turned away from the scene and did not witness the removal of the toy. When the actor turned back, the two windows, one of which the actor would open and retrieve the toy, were lighted (and with a simultaneous sound) to signal the retrieval

event. The signal and the retrieval event were familiarized with children in the familiarization trials before the test trials.

Southgate, et al., (2007) used two measurements of children's eye movements. One was the locations of the first saccade following the illumination of the windows, which showed most of infants firstly gazed at the correct window. The other one was the amount of time looking at each window during the first 1750 ms, which showed that infants looked at the correct window longer than the incorrect one. These findings showed that 25-month-old infants can correctly anticipate the action of the actor according to the actor's false belief.

In their design, Southgate, et al., (2007) considered the problem of "reality bias", which occurs when children's knowledge of the actual location of the object is in contrast to the location where children expect the agent to go. This becomes a problem when children have difficulty inhibiting or ignoring this information. Southgate, et al.. (2007) solved this problem by taking out the object from the screen at the end. The authors also designed several conditions to ensure that children did not respond according to the lower-level cues. To argue against Perner & Ruffman (2005)'s claim of children responding according to the last position of the object, they included a condition in which after the actor turned away from the scene, the puppet put the toy in the other box, and then removed it from the scene. To ensure that children's responses

were not due to the last position of the actor's attention, they included a condition that with the toy in a box, the puppet went to the other box to close the lid, which was witnessed by the actor; and then the actor turned away from the scene, and the puppet remove the toy from the scene.

Studies using non-verbal measures, particularly the recent research on infants' ToM reasoning, strongly suggest that traditional ToM tasks underestimate the competence of young children. Nonverbal techniques such as eye movement recording can be a powerful tool to unveil implicit, automatic ToM reasoning.

1.2.3 Challenges to Onishi-Baillargeon Study

Given that the Onishi-Baillargeon study is by far the most influential evidence for implicit ToM, it is not surprising that a debate is ongoing as to whether or not its findings reflect a true understanding of false belief or a demonstration of young babies' clever behavioral adaptation. Onishi and Baillargeon's (2005) influential findings were interpreted in different ways which inspired a debate on the nature of the knowledge used by young children in this task.

Perner and Ruffman (2005) and Ruffman and Perner (2005) argued that an assumption that a mind guides the behavior is not necessarily involved in explaining the Onishi and Baillargeon's (2005) findings. Instead, they proposed some simple ways to interpret the findings without going through the assumption of mind: (1) "low level

similarity to previous encodings” hypothesis which claims that infants form a three-way actor-object-location association in the brain which allows shorter looking time to old combinations and longer looking time to new ones.; (2) “behavior rules” hypothesis which argues that infants may expect the actor to act according to a particular rule, such as to search for the object where they last saw it; (3) “teleological understanding” hypothesis that argues that infants understand behaviors to be goal-directed and rational; (4) “mental understanding” hypothesis emphasizes different perspectives of a circumstance for understanding false belief. According to Perner and Ruffman (2005) and Ruffman and Perner (2005), Onishi and Baillargeon’s (2005) findings provide evidence for infant’s understanding of behaviors but still leaves the question of the infant’s understanding of mind. That is, infants’ performance in nonverbal false belief tasks can be explained by simpler understanding of behavior according to Perner and Ruffman (2005) and Ruffman and Perner’s (2005) arguments.

Csibra and Southgate (2006) argued against Ruffman and Perner (2005) that their explanation failed to explain the results regarding the following issues (1) 15-month-olds are very unlikely to form the particular episodic three-way associations in the brain; (2) the teleological model is not appropriate for explaining Onishi and Baillargeon’s (2005) results since it can only take into account actual states of reality; (3) children younger than 2 years old failed in both explicit and implicit false belief tasks in Clements and

Perner's (1994) study but succeeded in a similar implicit tasks without being asked the questions in Onishi and Baillargeon's (2005) study.

Onishi & Baillargeon (2005)'s study was further challenged by another study conducted by Surian, Caldi, & Sperber (2007) which used a similar paradigm but found no reliable evidence for 13-month-old infants' understanding of others' false beliefs.

Surian, et al., (2007) also used the violation-of-expectation paradigm to test if 13-month-olds can attribute people's beliefs. In their familiarization trials, there were two opaque screens, one on the left and one on the right. One apple was put in one of the screens and a piece of cheese was put behind the other (through the familiarization trials they were always put behind the same screens). An agent was present when the objects were hidden. Then the agent went behind the same screen to chew on the same object. After four familiarization trials like this, one test trial was presented to infants. Participants either saw a test trial in the "seeing condition", in which the screens were very low so that the agent can see what were behind the screens, or a "not-seeing" condition, in which the screens were the same as in the familiarization trials. In the test trials, the locations of the two objects were reversed. After the objects had been hidden, the same agent appeared. There were two possibilities of the agents' action: in the old-goal event the agent went to the same object as in the familiarization trials, and in the new-goal event the agent went to the different object. The hypothesis was that if

participants can attribute other people's beliefs, they would look longer at the new-goal event than the old-goal event in the "seeing condition", but would look at the old-goal event longer than the new-goal event in the "not-seeing condition" (since they would go to the same side of the screen as in the familiarization trials). Results showed that they did look longer at the new-goal event in the "seeing condition", showing their surprise of seeing the agent chose a new object; however, in the "not seeing condition" they did not show a difference of looking time in the old-goal vs. new-goal events.

There were two additional conditions in the second experiment reported by Surian, et al., (2007): the "knowing event" in which the agent was present when the two objects were hidden, and the "not-knowing event" in which the agent was absent when the two objects were hidden. In both conditions, the screens were high so that the agent could not see what were behind the screens. The same old-goal and new-goal events were both presented. The hypothesis was that if participants can attribute other people's beliefs, they would look longer at the new-goal event than the old-goal event in the "knowing condition", and look at the old-goal event longer than the new-goal event in the "not-knowing condition". Results showed that they did look longer at the new-goal event in the "knowing condition", but in the "not knowing condition" they did not show a difference of looking time in the old-goal vs. new-goal events. When the responses in the not-knowing condition were combined with the not-seeing condition and 3 outliers

were excluded from the analysis, infants looked longer at old-goal than the new-goal events.

In summary, using implicit measurements recent studies provide evidence for non-verbal ToM reasoning. However, there are still no conclusive results from this trend of research: Onishi and Baillargeon (2005) and Southgate et al. (2007) found 15- and 25-month-old infants showed early understanding of false beliefs, while 2.5-year-old children in Clements and Perner (1994)'s study did not show the implicit understanding of false belief; there is no clear evidence of 13-month-old infants' implicit understanding of false belief in Surian et al. (2007)'s study using a similar but somewhat different experimental setting. This may suggest that although infants respond correctly in the false belief tasks in some circumstances, they may only do so in narrow and constrained environment. This is consistent with the findings from Wellman et al. (2001)'s meta-analysis that children's performance in the traditional false belief task may be influenced by subtle differences of the experimental settings. These subtle differences in the experimental settings may influence also children's implicit understanding of false belief, which requires a real-time and automatic processing based on the available information.

The goal of the current project is to understand the failure of young children in the traditional change-of-location and unexpected-content tasks. Compared to the much simplified tasks used in recent infant studies, the stories told in the traditional ToM tasks

are much more complex. It is possible that 3-year-olds fail to keep track of ToM events even in non-verbal tasks: young children may be too overwhelmed by the real-time processing demand to encode rich and complex events, despite having an elementary ToM understanding.

1.2.4 Design of the Current Project

In the following studies, I will exam whether preschool children and adults show ToM reasoning in non-verbal tasks. The studies were designed to address this question by showing participants videos involving minimum language input, and recording their eye movements as an index of their understanding of false belief.

A two-alternative anticipatory eye movement paradigm was used to reveal participants' ToM reasoning in a revised version of change-of-location task. Similar to the design of Clements & Perner (1994), during a critical event in the video, participants are presented with two locations, and are cued implicitly (and nonverbally) to predict the action of a prognostic. McMurray and Aslin (2004) also used similar paradigm to assess infants' auditory and visual categories and showed that 6-month-olds can anticipate movements of stimuli under occlusion according to certain cues. In addition, children's explicit ToM performance was also assessed with traditional ToM questions and will be correlated with the ToM reasoning performance.

Children's implicit understanding of false belief was also tested in the revised version of unexpected-content task, which is different from the change-of-location task in details of experimental designs. First, there are verbal prompts to initiate a critical moment of prediction in the unexpected-content task, but not in the change-of-location task. Second, as argued by Southgate et al (2007), the reality bias may influence children's performance in the false belief tasks. The object was not removed from the screen in the change-of-location task, but it was not a problem for the unexpected-content task. Third, the change-of-location task used animated cartoon characters and the unexpected-content task used the real person videos. Fourth, the "not knowing" condition was indicated by the character masking her eyes in the change-of-location task, and by the character leaving the room in the unexpected-content task. These details of experimental settings may influence children's performance and may lead to different results from previous studies.

1.3 Linguistic Representations of Theory of Mind

1.3.1 Linguistic Markers of Theory of Mind

Although ToM reasoning can be non-verbal, the communication of ToM has to be mediated by language. A ToM marker is defined as a linguistic expression without which a ToM statement or question becomes a factual statement or question. In English, for example, beliefs are often marked with mental-state verbs ("think", "know") and

tensed sentential complements (de Villiers & de Villiers, 2000). Children have to understand these ToM markers to succeed in the traditional ToM questions.

Consistent linguistic marking for ToM is necessary to communicate mental states. Without a mutual understanding of such markers, it would be impossible to tell whether a linguistic message is about reality or about a mental representation/state about reality. In order to ask a question about people's private state of mind (e.g., beliefs and desires) about an event X, the question has to include additional linguistic elements (e.g., additional clauses and/or mental verbs such as think or want) so that the question can be distinguished from simply a question about X. These additional linguistic structures may create at least two challenges for young children.

In order to succeed in traditional ToM studies, children must correctly interpret a ToM question as a question about mental states, rather than about reality. A child may fail because s/he has not acquired the mental terms or corresponding syntax (e.g. the sentential complement) to describe other people's minds. It is also possible that children may be limited by their cognitive capacity in processing of long sentences and are thus forced to disregard linguistic markers of mental states. Both of these possibilities suggest that young children may not mentally represent these critical linguistic markers of mental states. Children who fail to comprehend the ToM marker in these questions will

misinterpret the questions about the mind as questions about reality. If so, they will perform as if they have no ToM ability.

For example, in the unexpected-content task (Gopnik & Astington, 1989), children are asked about their past false belief ("what did you think is in the box before you open it?"), and a 3rd person's false belief ("What will Mary think is in here, crayons or candies?"). Children may misunderstand these questions, failing to interpret these questions as questions about mental states. This communicative barrier may also underlie the failure, to date, to find evidence of false belief understanding among nonhuman primates: although reliable nonverbal communication of beliefs is not completely impossible, the conditions under which this can be achieved may be subtle and fragile. For example, in children's limited experiences with such questions including "do you think", most of the time the mental verb phrase can be ignored without any changes to the meaning. For example, the answer to the question "what do you think it is?" is the same as the answer to "what is it?" The marker "do you think" is non-contrastive in this context, making it difficult for the child to infer the pragmatic meaning of this structure. By extension, the child may choose to ignore similar structures such as "did you think". In doing so, s/he will not interpret it as a question about mind but a question about reality.

1.3.2 Acquisition of ToM markers

The acquisition of these linguistic labels, however, is not a trivial task for several reasons. First, the referents of these markers are intangible and abstract. There are no specific objects or actions related to the concept of "belief". For example, children learn the concept of "apple" by seeing a pointing to a real apple, but they cannot learn the concept of "belief" by seeing a pointing to a head. Second, learning ToM markers requires that their meanings be contrastive, i.e., changing (or dropping) the markers will change the meaning of the utterances. However, in most of cases, ToM markers are used to represent someone's true beliefs (e.g. "I think the door is closed"). Only a small portion of cases ToM markers represent false beliefs (e.g. "I thought the door was closed"). This makes the acquisition of ToM markers difficult for children because in most cases, the true belief situation, these markers may be ignored by children without affecting their understanding of the utterances (what do you think is in the box? → what is in the box?). While in false belief situation, these markers cannot be ignored, otherwise the meaning of the utterance will be changed (what did she think is in the box → what is in the box?). Unfortunately, there may not be enough cases of false beliefs in young children's daily conversations for them to acquire ToM markers. An unpublished corpus study conducted by me using the CHILDES database showed that in a child's daily language input before age 4, only around 30% of time "do you think", the most

frequently used ToM markers, was used in the situations where it cannot be ignored from the questions without changing the meaning of the whole questions. The percentages of ToM markers that cannot be ignored increase significantly from age 3 to 4.

These factors make the acquisition of mental vocabulary relatively late comparing to other semantic categories. From 2.5 years, mental-state terms such as “think”, “know” and “remember” start to appear in children’s spontaneous speech (Limber, 1973; Shatz, Wellman, & Silber, 1983). Two- and three-year-old children begin to acquire mental-state terms to first refer to perception, emotion, and desire (e.g. see, hear, happy, love, want), and then to knowledge and beliefs (e.g. know, think, believe; Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982). Three- and four-year-old children can comprehend the subtle differences of implications of “think” and “know” by giving different answers to the questions using “think” and “know” to ask about people’s false belief (Johnson & Maratsos, 1977).

1.3.3 Linguistic Representations of ToM and ToM Performance

As discussed before, children have to acquire linguistic representations of ToM to succeed in traditional ToM tasks. Previous research showed that linguistic representations are important for children’s false belief understanding (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 2003; Lee, et al., 1999; Lohmann & Tomasello, 2003; Milligan, et al., 2007; Tardiff, et al., 2004). Many studies

focus on specific linguistic structures or features that are critical for understanding false beliefs and mental states. Most of these linguistic features are consistent with the definition of linguistic ToM markers in the current context. Other researchers look at a broader connection between ToM and general language abilities.

Some researchers argued that language is an essential precursor of false-belief representation (Astington & Baird, 2005). In a strong version of “Linguistic determinism”, the development of Theory of Mind relies on the acquisition of the syntax of complementation, which provides the representational tools for understanding false belief. For example, in the sentence “*He thought the door was closed*”, it is possible that the whole sentence is true but the embedded complement is false. Therefore, the complement can work as a representational structure for belief. However, at an early age, complements of belief verbs are treated as true “realis” clauses (de Villiers & de Villiers, 2000). Children begin to realize that the complements embedded under the verb “think” is false compared to the real world, with analogy to the complement of the verb “say”, which shares the same discourse contexts and syntactic structures. It is much easier to see and understand that people “say” things different from the reality. Therefore, the verb “say” may provide a bootstrap and understanding of syntax of complement of belief verbs.

This linguistic determinism theory is supported by research evidence from longitudinal studies, training studies, studies with deaf children and autistic children. A longitudinal study by de Villiers and Pyers (2002) measured 3- to 5-year-old children's understanding of false belief and spontaneous language production and comprehension over a year. Regression analysis showed that preschool children pass the false belief tasks after they master the structure of verbs with complements.

In a training study of Theory-of-Mind acquisition (Lohmann & Tomasello, 2003), a large group of 3-year-old children were shown a series of deceptive objects (e.g. a candle with an apple-like shape) and participated in several conditions of language training, including training on the mental state terms and training on sentential complements that do not involve mental state terms. Results showed that children's performance in false belief tasks was improved in these language trainings, even the training on the sentential complement without mental-state terms.

Another training study was conducted by Hale and Tager-Flusberg (2003). In this study, 60 preschoolers who failed in the false belief pretest participated in the training on false belief, sentential complements or relative clauses. The results showed that children who got the sentential complement training not only improved in the sentential complement tests, but also in the false belief tests; children who received false belief training only improved in the understanding of false belief but not in language. This

finding suggests a causal role of sentential complement in the development of understanding belief.

Lee, Olson, and Torrance's (1999) investigated the effects of mental verbs in Chinese on Chinese-speakers' understanding of false belief. "*Xiang3*", "*yi3wei2*", and "*dang4*" are mental verbs to represent beliefs in Chinese. "*yi3wei2*", and "*dang4*" are more likely to be used to talk about belief when it is false than "*xiang3*" which is more neutral. Results showed that children perform better in "*yi3wei2*", or "*dang4*" questions, than "*xiang3*" questions. This finding suggests that preschool children can detect subtle differences within a category (belief) in language and make use of the differences to understand beliefs.

Tardif, Wellman, and Cheung (2004) tested whether the type of Cantonese verbs used in the test questions of the false belief task makes a difference in Cantonese-speaking children's performance. Cantonese represents beliefs in two ways: either neutral belief (*nam5*) or explicitly false (*ji5wai4*). Results showed a strong advantage in the performance of false belief tasks for children questioned with explicitly false belief words than children questioned with neutral belief words. This finding implied that children's ToM performance is mediated by their linguistic ToM.

On the other hand, the relations between language and ToM are discussed more broadly in the literature: not only the specific linguistic structures, but the linguistic

experiences or general language abilities are also related to the false belief understanding (Brown et al, 1996; Cheung, Chen, Creed, Ng, Wang, & Mo, 2004; Dunn et al, 1991; Milligan, et al., 2005; Ruffman et al, 2002). Dunn & Brophy (2005) claimed that the development of ToM is through the participation in linguistic exchanges. This claim is supported by studies which found positive correlations between family conversational experiences in early years and their later ToM performance (Brown et al, 1996; Dunn et al, 1991; Ruffman et al, 2002). However, these general language experiences or abilities are not equivalent to ToM markers mentioned in the current studies.

1.3.4 The Current Study

I have argued that in order to succeed in a traditional ToM task, children and adults must correctly interpret the ToM markers in verbal questions. However, a direct assessment of their understanding of the ToM marker, for example, by asking questions or making predictions, will be hopelessly confounded with children's ToM ability.

The approach taken in the current study is to assess the representational strength of ToM markers without invoking ToM processing. Here I adopt an elicited imitation task, in which participants are asked to repeat verbatim what they hear (Ambridge & Pine; 2006; Flynn & Lust, 1981; Keller-Cohen, 1981; Lust, et al., 1996). The sentences are questions taken verbatim from previously published ToM studies or made in the same

way. In other words, they are representative of the type, length, and complexity of questions children will encounter in traditional ToM tasks. Pilot studies show that preschool aged children have tremendous difficulties in repeating the questions. Almost invariably they will drop words, phrases, and other linguistic structures in their responses. The key question is whether they are more or less likely to drop ToM markers compared to other elements in the questions.

It is important to note that I do not imply that if a child fails to repeat a ToM marker, s/he necessarily misinterpreted the ToM question as a factual question. Children drop words for many reasons. Missing a ToM marker in the elicited imitation task is not diagnostic of misunderstanding. Rather, children's performance in the elicited imitation task is taken here as an index of the representational strength of a linguistic structure. The logic is that if one understands a question as asking about the mental state, rather than reality, then he or she is likely to make an effort to preserve the ToM marker, possibly at the expense of other errors when the working memory is overloaded. On the other hand, if one does not attach any significance to the ToM marker, then it is as likely to be dropped as other elements.

In study 1, children and adults are asked to repeat the questions in the context of a traditional ToM task. The goal is to test the hypothesis that young children (e.g., 3-year-olds) do not appreciate the special nature of linguistic ToM markers in these

questions and therefore they are likely to drop them. This also affords an item-level analysis, where the performance in the elicited imitation task can be correlated with their performance in the same ToM tasks. The hypothesis is that being able to preserve ToM markers in elicited imitations should be predictive of better ToM performance. In study 2 children and adults are administered a standalone version of the elicited imitation task by using questions from previous research. Study 3, replicated study 1 in the second-order belief understanding task. Study 4 replicated study 1 with Chinese-speaking counterparts.

1.4 Other Issues Addressed in the Present Studies

1.4.1 Second- and Higher-order Theory of Mind

Most traditional Theory of Mind research investigates children's understanding of other people's representations of physical events, also known as first-order beliefs. Some more complex situation, such as deception, requires the higher capacity to attribute second-order, or embedded, mental states (e.g. "she thinks that he thinks ..."). Children need to understand not only people's thoughts about the physical events, but also people's thoughts about other people's thoughts (second-order beliefs), and even what people think that others think about their thoughts (higher-order beliefs).

A classic second-order belief task (Perner & Wimmer, 1985) told a story about two children, John and Mary, saw an ice-cream van in a park. They were independently

informed that the ice-cream van was moved from the park to the church, and did not know the other child was also informed. At the end of the story, children between age 5 to 10 were asked “Where does John think Mary will go to buy an ice-cream?” and “Does John know that Mary knows where the ice-cream man is now?” Results showed that the second-order belief understanding begins to emerge around 6 years of age (Perner & Wimmer, 1985).

Sullivan, Zaitchik & Tager-Flusberg (1994) used a simpler story than Perner & Wimmer (1985) by making the story shorter, involving fewer characters, locations and episodes. Instead of using tape recorders as in Perner & Wimmer (1985)’s study, Sullivan, et al. (1994) used an experimenter telling the similar stories, asking children probe question in the middle of the stories, giving them feedback for each question, and providing memory aids. Sullivan, et al. (1994) simplified the second-order belief story by using deception. In the new story, “birthday puppy”, Peter’s mom misinformed Peter about his birthday gift in order to surprise him. However, Peter later discovered his true gift in the basement. At the end of the story, Peter’s grandma called mom and asked her “does Peter know what his gift is?” A question was asked to children “What does Mom say to Grandma?” Results showed that around 40% of 4-year-olds and 90% of 6-year-olds can attribute second-order beliefs in this version of story.

Traditional research on second-order belief understanding also used story-telling paradigm, and thus the problem with language and working memory dependency is even more aggravated than research on the first-order ToM. Particularly, young children may have difficulty dealing with double-embedded propositions to answer the questions like “Does John know that Mary knows where the ice-cream man is?” (Sullivan, et al., 1994). Children are much more likely to misrepresent one or more ToM markers in these questions.

The Current Study

The current study is an attempt to investigate why children younger than 6 year old fail in the second-order belief understanding tasks. By this time they should have acquired the first-order belief understanding, and there may be two reasons why they fail in the classic second-order belief understanding tasks. First, they can linguistically represent ToM markers in the questions with only one marker, but not in the ones with two or more markers. The questions asked in the second-order belief tasks are more complex than the questions for the first-order belief questions, not only because of the numbers of markers, but also they are in a more complex syntactic structure- the double-embedded complements. Second, children younger than 6-year-olds may not possess an automatic and real-time reasoning for the second-order beliefs. This will be tested with

implicit measurement, the anticipatory eye movements, with minimum language involvement.

In study 3, 5- to 6-year-olds' interpretations of second-order ToM questions and their real-time process of second-order beliefs in a nonverbal task will be addressed. However, due to the double-embedded nature of the classic questions for second-order beliefs ("Does John know that Mary knows where the ice-cream man is?"), they are very difficult to be asked non-verbally. In order to test children's second-order ToM in a non-verbal story, I designed a story to test the second-order beliefs using the deception paradigm.

The paradigm of children's deception behaviors was revised to fit in the current purpose. Young children's deception behaviors are typically studied in a paradigm developed by Lewis, Stanger and Sullivan (1989). Children were instructed not to peek at a toy when they are left alone in a room. Most 3-year-old children could not help peeking at the toy but later when asked, most of them denied having done so or did not answer the question. This denied of peeking is related to their understanding of first order false belief in a Theory of Mind task (Polak & Harris, 1999).

I combined this task with unexpected-content task to put a candle in the crayon box, and asked two players, A and B to play this game in the videos. In the video, B is not allowed to look in the box but she looks while A leaves. When A asks about the

content of the box, B has to point to the crayon although she knows it is a candle in the box to get the reward for not looking in the box. A training part is shown before the main part of the video to emphasize B's desire to get the reward and the motivation of her deception. Children's anticipatory looking will be recorded to reveal their predictions and understanding of the deceiving behaviors. This paradigm can be used to test children's second-order belief understanding since in order to make correct predictions of B's actions (pointing to the crayon instead of the candle), participants need to understand that B tries to pretend not have looked in the box to make A believe that she did not look in the box. It is predicted that adults, as experienced mind-reader, are able to attribute second-order beliefs by looking longer at the crayon than the candle; while children may not show the second-order belief understanding in their eye movements.

However, this video tells a story more complex than the stories told in the classic second-order ToM tasks. The current version of the second-order ToM story involved deception. In order to make correct predictions of the actor's actions, participants need to take both the actor's second-order beliefs and her motivation to cheat into consideration. This may require extra process than the classic second-order ToM stories.

1.4.2 Cultural Differences

ToM tasks have been tested with children from different countries speaking different languages. A meta-analysis showed that children in all cultures exhibited the same developmental trajectory of ToM, although children of the same age from different countries may perform differently (Wellman, et al., 2001). Children from Australia and Canada performed somewhat better than U.S. and U.K. children, while children from Austria and Japan were somewhat worse than U.S. and U.K. children (Wellman, et al., 2001). The differences may be due to linguistic and cultural differences, but the developmental pattern of ToM is universal. In the present study, the same tasks were also conducted in China with Chinese-speaking children and adults, to see whether the findings with English-speaking participants will be expanded to another culture.

1.5 Overview of the Present Studies

The overarching goal of this project is to understand why young children fail to answer traditional ToM questions correctly. Study 1 (Chapter 2) is designed to achieve two goals. First, it tests whether English-speaking children adequately represent linguistic ToM markers in ToM questions or statements. If young children fail to encode ToM markers, they are at risk of misinterpreting ToM questions as factual questions. Also, it asks whether answers to traditional ToM questions can be predicted – at the item level – by whether or not the ToM marker is correctly repeated in the elicited imitation

of the same question. That is, in addition to answering questions in traditional ToM tasks, participants are also asked to repeat the same ToM questions. A positive correlation is predicted if the linguistic representation of ToM marking contributes to performance in traditional ToM tasks.

In addition to examining the role of linguistic representations of ToM, study 1 also measures real-time ToM reasoning by tracking participants' eye movements during traditional ToM tasks. A number of recent studies suggest that young children may be able to automatically and implicitly understand social event in terms of mental states. On the other hand, the dominant view in developmental psychology is that children do not acquire an adult-like theory of mind until 4 or 5 years of age.

Study 2 (Chapter 3) also tests the linguistic representations of ToM but using questions from previous studies, to explain children's failure in the traditional false belief tasks in terms of linguistic representations. Study 3 (Chapter 4) is similar to study 1 but instead focuses on second order beliefs. Study 4 (Chapter 5) reports a series of studies with Chinese-speaking children and adults, replicating major findings in the English studies. Implications of these findings will be discussed in the General Discussion in Chapter 6.

2. Study 1: Children's Understanding of First-order Beliefs

Study 1 examined two processes involved in the traditional false belief tasks to reason about other people's false belief. First, study 1 investigates whether young children pay attention to the ToM markers when they hear ToM questions. In this study I will examine young children's linguistic representations of complex ToM questions, particularly whether they omit linguistic ToM markers, when the complexity of the questions exceeds their memory and cognitive capacities. Three- and four-year-old children participated in an elicited imitation task in which participants were asked to repeat verbatim questions they heard. All questions were asked in the same way as mind questions in previous false belief tasks. The hypothesis is young children do not linguistically represent ToM markers so that they will drop the markers when the questions exceed their working memory or cognitive capacities. It is also hypothesized that this linguistic representation can predict their traditional ToM performance.

Second, young children may also fail traditional ToM tasks if they fail to perceive and reason about a mind-related event in a ToM framework. Study 1 assesses children's non-verbal ToM abilities by tracking their eye movements while they watch a ToM event unfold. In all the tasks there is a critical moment in which the participants will make anticipatory eye movements. Depending on how they perceive the event, their

eyes will be directed to different visual targets. Because no explicit instruction or verbal question was involved, these anticipatory eye movements will shed light on how children and adults encode ToM events in real-time, and whether 3-year-old children fail traditional ToM tests because they do not spontaneously encode these events as ToM events. In addition, study 1 also examines the relation between eye movements and the performance in the verbal ToM tasks, and with the elicited imitation tasks.

Study 1 used the change-of-location and the unexpected-content paradigms, the two dominant research paradigms in studying children's ToM (Gopnik & Astington, 1989; Wellman, et al., 2001; Wimmer & Perner, 1983). One of the purposes of this study is to investigate whether young children show ToM reasoning in non-verbal ToM tasks, and its relations with explicit ToM understanding. According to Clements and Perner (1994)'s findings, implicit understanding of false belief indicated by looking time precedes explicit understanding of false belief shown by verbal answers to ToM questions. It is hypothesized that some of the children will make correct predictions in their eye movements but give incorrect answers to ToM questions in the current tasks.

2.1 Method

2.1.1 Participants

Thirty-two 3-year-old (average age=41 months, 15 females, 17 males), twenty-six 4-year-old (average age=54 months; 14 females, 12 males) English-speaking children,

and twenty-nine English-speaking adults (average age=19 years; 15 females, 14 males) were recruited from community preschools and Duke University. The adults also participated in the studies in studies 2 and 3; 3- and 4-year-olds children also participated in study 2; and 5- and 6-year-olds also participated in study 3.

2.1.2 Material

Elicited imitation task

This task was designed to see which part of the question, if any, is omitted when the length and complexity of the sentence is beyond the memory/cognitive capacity of the participants. In order to test the hypothesis that young children fail traditional ToM tasks in part because they do not reliably understand the mind-related nature of critical ToM questions, it is also important to measure, at a trial-by-trial level, how they encode the ToM question and how they answer them. This requires children to answer and repeat each question as part of the traditional ToM task. Pilot studies indicated that it is difficult for young children to repeat the question before answering it. The procedure used in the present study was to answer the question first – which typically took little time – and then repeat the question.

The experimenter (a non-native but fluent English speaker) asked participants to play a game with a puppet named “Fifi”. A cover story was told that Fifi is very “silly” and “forgetful”, so that “Whenever I say something to her, she always forgets”. Then the

child was asked to help Fifi: "could you help her? You just need to reminder her what I say. Could you repeat/copy what I say?"

To make sure the child understand the task, some training was provided. The training started with single words like "apple" and "banana", and was then extended to more complex phrases and sentences like "I like apples". Then the child was trained to repeat questions such as "do you like apples?" or "what is your name?" Some children tended to answer instead of repeat the question. In this case, the experimenter reminded them: "Remember you need to remind Fifi what I just said. Could you repeat (or say, copy) what I say?" or "could you ask Fifi the same question?"

Then children were trained in this study to answer some verbal questions and then to verbatim repeat the questions (as trained before). Children were shown a picture book. The experimenter pointed to some animals or objects in the book and asked questions such as "what is this?" or "what color is this?" or "how many flowers are there?" After children answered the question, they were asked to repeat the question back: "What did I just ask you?" or "Could you remind Fifi that question?" There were no training for older children and adults. They were simply asked to answer the questions and then to repeat the questions back to the experimenter. Participants' answers were recorded verbatim on a recording sheet and also in audio recording tapes.

Change-of-location Task

Children in this study were shown the videos telling the stories in both the change-of-location task and the unexpected-content task (see the next section). Their anticipatory eye movements were recorded as an index of their ToM reasoning.

Four animated short cartoon movies were made by a professional artist using Adobe Flash. The style and format of the animations resemble TV programs familiar to preschool aged children. All children found the animations fun and engaging.

The animation started with an introductory episode, which showed a girl named Lisa, who was playing in a park and would hide her toys (cars, balls, trucks, and trains) and later retrieve them. Then each episode depicted an independent game in which Lisa hid a different toy. There were two possible hiding locations, a lawn chair and a large rock, at the upper-left and upper-right corners of the screen. Lisa had to follow a Y-shaped path to go to either place. At the center of the screen, the path went through a forest and consequently one could not see which way Lisa would go until she re-emerged from either the left or the right path. Each time, Lisa would go through the path and disappear behind the forest. After a 5-second delay, she reappeared either on the left side of the path to hide the toy behind the chair, or on the right side of the parkway to hide the toy behind the rock.

After hiding the toys, Lisa came back to the entry of the parkway and covered her eyes with a mask. This reassures the participant that Lisa was unable to see. During the time she covered her eyes, the toys either moved to the other side on their own (from chair to rock or from rock to chair; the false belief trials), or stayed in where they were (the true belief trials). These self-propelled motions were credible because the toys were wind-up cars and balls, and the movements were accompanied by the sound of wind and blowing leaves. Then Lisa uncovered her eyes, turned, and went back to the forest, with a clear intension to retrieve the toy as she said she would in the introductory episode. She would disappear behind the woods for five seconds, and reappear either on the chair side or the rock side to find her toys. Participants' anticipatory looking during these five seconds was recorded by the eye tracker and then analyzed as an index of their implicit understanding of beliefs.

Besides the introductory trial, there was a TB trial, in which the car did not move, and two FB trials in which the car moved, either from the chair to the rock, or from the rock to the chair. The first FB trial was a non-complete trial, which did not show the ending of the story: the movie stopped when Lisa disappeared in the forest and never came out of the forest. The second FB trial was a complete trial, which showed the ending of the story: Lisa went to where she originally put the car, but not where the car actually was.

Traditional ToM questions about this cartoon movie were asked to at the end of each trial, including: (1) probe questions to test whether the participants follow the story and their memory of the key events, e.g. “Where should Lisa go to look for the car?”; (2) belief questions: “Where did Lisa think the car was when she opened her eyes?” and “Where did Lisa think the car was before she put on the mask?” (3) a prediction question in the non-complete FB trial (“Where will Lisa go to look for the car?”). After they answered the question, they were also asked to do elicited imitation of the questions by being asked: “what did I just ask you?”

Unexpected-Content task:

Four real-person video clips were shown to the participants in this task. In these videos, two pairs of persons played with two different sets of toys. Before watching each video, child participants were given opportunities to play with the same toys as they would later see on the videos. It is important that they were familiar with the toys prior to the eye tracking study, so that they did not show a strong preference to look at one toy versus another. The videos follow a basic script, in which protagonist A, alone, is first showed a familiar container (e.g., a crayon box) with a surprising content in it (e.g., birthday candles). A then showed another protagonist B the box before leaving the room, and later came back to ask B what was in the box. During the intervening time B may or

may not open the container to see the actual content, i.e., B may or may not held a false belief about the content of the box.

A detailed script of events in these videos can be found in Table 1. These videos were in two conditions, with two parts of videos for each condition. In the FB-FB condition B did not open the box so that she held a false belief in both two parts of the video; while in the FB-TB condition, B's false belief in the first part was updated into a true belief in the second part by opening the box and discovering the content of it. As shown in the Table 1, the FB-FB condition tells the same story as the FB-TB condition except that in the FB-TB condition the content of the box is discovered by B, while in the FB-FB condition, the content of the box is unknown to B.

There is a critical moment in which the participants' anticipatory eye movements were most revealing of their real-time understanding of the event. When B was asked about the contents of the container, a still shot of two objects (the expected and the actual content, one on the left and one on the right) were shown for five seconds, as B said in a hesitant voice "hmm..." as if she was thinking and deciding. The actions leading to this scene, the hesitation of B, and the close-up shot of the two objects were all invitations for the participants to enter a mental simulation in order to predict which object B was likely to point to.

At the end of the anticipatory period, the participant saw a finger (presumably that of B) pointing to one of the objects. Participants were then asked various factual and ToM questions as in traditional ToM tasks.

During each five second delay participants' anticipatory looking was recorded to assess their prediction of the actor's pointing actions. Then an ending (B's pointing actions) was shown and some probe and belief questions were asked after showing each video. The descriptions of each scene and questions are listed in the Table 1.

Table 1: Descriptions of each scene of the videos used in the unexpected-content task of study 1

	FB-TB condition	FB-FB condition
Part 1	A, alone in the scene, shows a crayon box, takes out the crayon in the box, and puts a candle in it.	same as FB-TB condition, but replace crayon with cookie, and candle with plate
	B comes in the room, and A asks B the content of the box	same as FB-TB condition
anticipatory looking scene 1	scene of two objects for 5 s	same as FB-TB condition
predictions 1	"which one will she point to?"	same as FB-TB condition
pointing 1	B points to the crayon	B points to the cookie
verbal questions (answer and then repeat the sentence)	Justification question: "Why did she point to the crayon?"	Justification question: "Why did she point to the cookie?"
	Probe question: "What is in the box?"	same as FB-TB condition
	belief question 1: "What does she think is in the box?"	same as FB-TB condition

Table 2 (continued)

Part 2	While A leaves, B looks into the box and is surprised to see the candle.	While A leaves, B does not look in the box.
	A returns and asks B to point to one of two objects (crayon and candle) to indicate what is in the box.	same as FB-TB condition
anticipatory looking scene 2	scene of two objects for 5 s	same as FB-TB condition
predictions 2	"which one will she point to?"	same as FB-TB condition
pointing 2	B points to the candle	B points to the cookie
verbal questions (answer and then repeat the sentence)	Justification question: "Why did she point to the candle?"	Justification question: "Why did she point to the cookie?"
	Probe question: "What is in the box?"	same as FB-TB condition
	belief question 1: "What does she think is in the box?"	same as FB-TB condition
	belief question 2: "What did she think was in the box before she opened it?"	N/A

2.1.3 Procedure

Participants watched all cartoons and videos. Participants' eye movements while they watched the videos were recorded by a Tobii X50 eye tracker. The eye tracker was connected to a 17" monitor with 800 X 600 resolutions. Stimuli were presented on this monitor via a computer running Tobii's Clearview software.

After each video, children and adults were asked several questions about the video. After answering each question, they were asked to repeat the questions that they just heard.

All participants watched the videos for unexpected-content task first and then the cartoon videos for change-of-location task. In the unexpected-content task, all participants watched the video in the FB-FB condition first and then the one in the FB-TB condition. In the cartoon videos for the change-of-location task, the true belief condition was presented first, then the non-complete false belief condition, and a complete false belief condition.

2.2 Results and Discussions

2.2.1 Verbal Answers to ToM Questions

Participants' correct rates in answering verbal questions were calculated for each task and question type. Results were shown in Figure 2 (change-of-location task) and

Figure 2 (unexpected-content task). Results showed that 3-year-olds made significantly more errors than 4-year-olds when answering (1) prediction questions (“Where will Lisa go to look for the car?”) in false belief trials $F(1,78)=12.29$ $p<.001$; and (2) belief questions “where did Lisa think the car was before she put on the mask?” $F(1,79)=43.95$, $p<.0001$, and “where did Lisa think the car was when she opened her eyes?” in false belief trials $F(1,79)=22.29$ $p<.0001$. There was no difference found between 3- and 4-year-olds in answering the probe questions “Where should Lisa go to look for the car?” $F(1,80)=2.59$, $p=0.11$. This question can be answered based on the factual knowledge of where the object was.

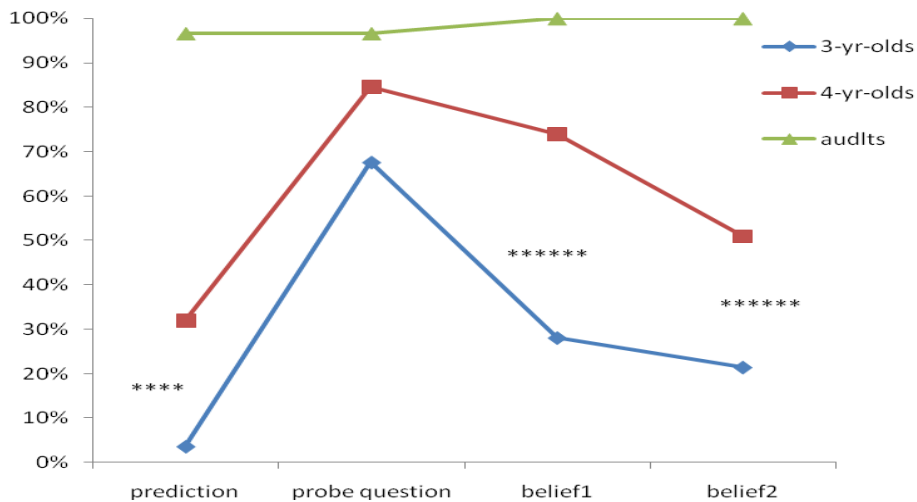


Figure 2: Correct rates of all ages answering different types of questions in the change-of-location task of study 1

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; ***** $p<.0001$

In the unexpected-content task, as shown in Figure 3, 3-year-olds made significantly more errors than 4-year-olds when answering belief questions in the present tense “What does she think is in the box?” $F(1,82)=39.40, p<.0001$, and belief questions in the past tense “What did she think was in the box before she opened it?” in false belief trials $F(1,80)=14.15, p<.0005$. There was no difference found between 3- and 4-year-olds’ performance in answering the prediction questions “which one will she point to?” $F(1,82)=0.83; p=.36$, and the probe questions “what is in the box?” $F(1,82)=0.45, p=0.50$.

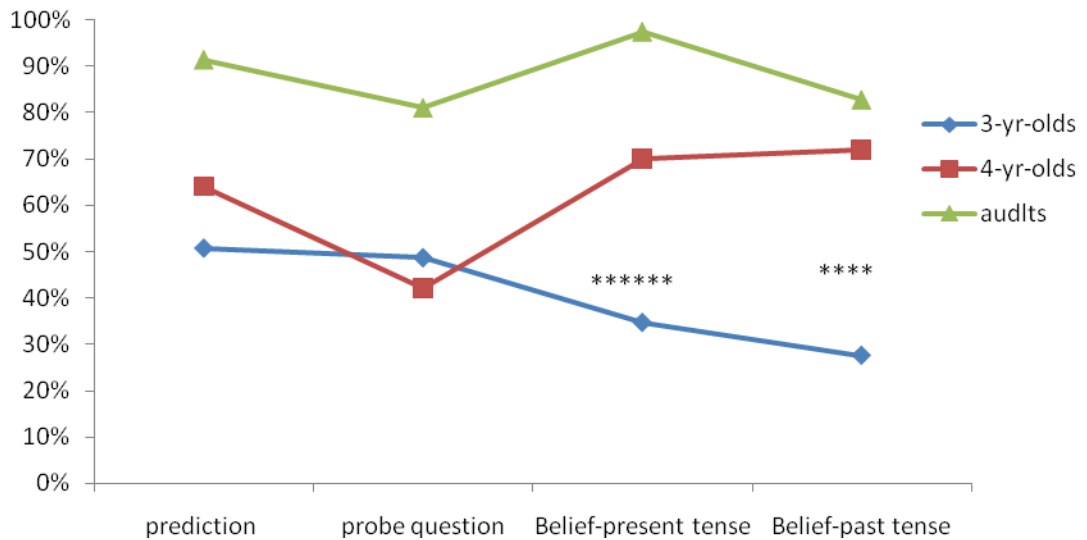


Figure 3: Correct rates of all ages answering different types of questions in the unexpected-content task

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; ***** $p<.0001$

2.2.2 Elicited imitation

Participants' performance in the elicited imitation task was coded in the same way as in the study 2. The reliability between the two coders was high *inter-coder reliability*=97.85%; *Scott's Pi*=96.05%.

Results showed that age differences in probabilities of omitting ToM markers: 3-year-olds omitted ToM markers over 90% of time; this rate dropped to around 60% at 4 year olds $F(1,41)=7.88, p<.01$. However, there was no difference of probabilities of making other errors among children $F(1,39)=0.40, p=0.53$. No adults made any kind of errors in this task and their performance was not depicted in the Figure 4.

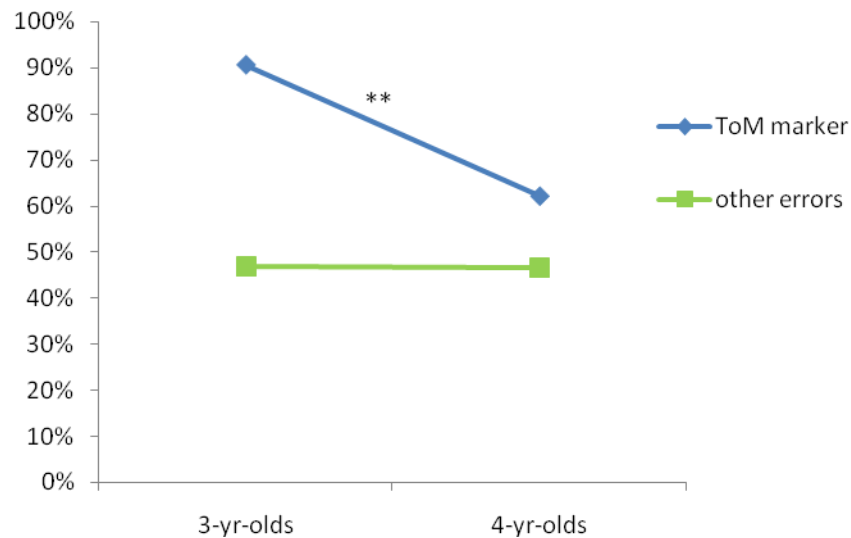


Figure 4: Percentages of making errors in ToM markers and other errors in the elicited imitation task

Notes: ** $p<.01$;

One hypothesis of the current study is that children's abilities to perceive and linguistically encode ToM events are the proximal mechanisms determining whether they succeed in traditional ToM tasks. The relations between children's performance of verbal answers to ToM questions and whether they can correctly repeat ToM markers were also tested. A chi-square test based on the combined sample of 3- and 4-year-old children showed that there was a correlation between children's verbal answers and whether they made errors in the ToM markers in the elicited imitation task *chi-square=18.83, p<.0001*. The same test was conducted to test the relations between children's performance on verbal answers and whether they made "other errors" and no such correlation was found *chi-square=2.65 p=0.10*.

However, this correlation may be partially attributed to age. Two analyses were conducted to address this concern. An ANCOVA was conducted with the age as a covariate to test the relations between children's verbal answers and whether they can correctly repeat ToM markers in the questions. Results of ANCOVA showed that children's performance on verbal answers is significantly predicted by whether they can correctly repeat ToM markers in the question after controlling age $F(1,323)=8.92, p<.005$. The same test was conducted between children's performance of verbal answers and

whether they made other errors and no such correlations were found $F(1,351)=2.80$, $p=0.10$.

To further test the correlation between the performance of repeating ToM markers and verbal answers, a chi-square test is conducted with the 4-year-old group only; a similar analysis could not be conducted with 3-year-olds because over 90% of them had difficulties representing ToM markers in the elicited imitation task. Results show that there was a correlation between 4-year-old children's verbal answer performance and whether they can correctly repeat ToM markers in the question $chi-square=4.19$, $p<.05$, but no correlation between 4-year-old children's verbal answer performance and making other errors in the elicited imitation task $chi-square=0.17$, $p=0.67$.

Interim discussion on the verbal answers to ToM questions and the elicited imitation task

Children's verbal responses to ToM questions replicated the findings in traditional ToM tasks (Gopnik & Astington, 1988; Wellman et al, 2001; Wimmer & Perner, 1983): 3-year-olds fail to show an understanding of false beliefs in their verbal responses. This similarity confirms the findings by a meta-analysis (Wellman, et al, 2001) that children's ToM performance are consistent in all types of tasks.

This current study was also designed to address the hypothesis that young children's failure in the verbal responses to ToM questions may be partially due to their

difficulty representing ToM markers. Findings of this study suggest that children younger than 4-year-olds fail to encode the ToM markers in traditional ToM questions, which may contribute to their failure in the traditional ToM tasks.

The findings in this study provided evidence for the framework proposed in Chapter 1, which claims that the lack of linguistic representations of ToM may limit the performance in answering traditional ToM questions. In order to understand the traditional ToM questions, children need to interpret the questions as questions about mind but not factual questions. Without linguistically representing ToM markers, children would misinterpret the ToM questions as questions about reality but not about mind. This is because in young children's daily language input, most of the time the ToM markers can be ignored in the utterances. Only in some rare cases of false belief situation, they have to pay attention to the ToM markers. This unreliability of ToM markers leads to the young children's insensitivity to ToM markers in questions.

Children's failure to correctly answer the prediction question ("Where will Lisa go to look for the car?") in the change-of-location task was not due to their memory limits: 3- and 4-year-olds answered the probe question ("where should Lisa go to look for the car?") as well as adults, while their performance in the prediction questions was still below chance. In the unexpected-content task, 3- and 4-year-olds were only correct in the probe questions ("what is in the box?") around 40% of time. Even adults were

only correct 80% of the time when answering this simple memory question. This may be primed by the pointing behaviors of the actor. At the end of the videos, the actor was asked “what is in the box?” and she pointed to the crayon in the false belief condition. This event may impact participants’ performance when they were asked the same question, so that they gave the answer the same as the actor did. This is an issue future designs should consider and address.

In summary, elicited imitations revealed that compared to older children and adults, 3-year-olds tend to consistently ignore ToM markers in the utterances. Linguistic representations of ToM are related with the traditional ToM performance regardless of age. This implies that linguistic representations of ToM make a unique contribution to the ToM performance.

2.2.3 Eye Movements in the Change-of-Location Task

Eye movements recorded during watching the cartoon videos were analyzed for all three conditions in the change-of-location task. First, participants’ first gazes after Lisa had disappeared were located (left or right paths). Second, the total looking time at the left or right paths was calculated and compared. Third, the average x position of participants’ gazes was calculated for each age group. The mean x is compared to the

midline ($x=373$ pixels), where Lisa disappears into the forest, to determine the preference of the age group.

True Belief Condition

In the true belief trial, the toy (e.g., car) was put on the right and never moved. Participants should look at the right side of the screen to predict Lisa's movement. Four areas of interests (AOI) were defined: the (1) left and (2) right paths where Lisa will reappear in the screen, and the (3) chair (left) and (4) rock (right). Results of the first gazes showed that 16 out of 25 3-year-olds, 16 out of 22 4-year-olds and 15 out of 24 adults made the first gazes towards the correct prediction by looking at the right path and the rock (AOIs 2 and 4) in the critical moment.

The total looking time at the left or right path was compared across subjects using a paired t-test (See Figure 5). Results showed that 3-year-olds $t(24)=2.54, p<.05$, 4-year-olds $t(21)=2.15, p<.05$ and adults $t(23)=3.50, p<.005$ all looked at the correct prediction longer than the incorrect prediction.

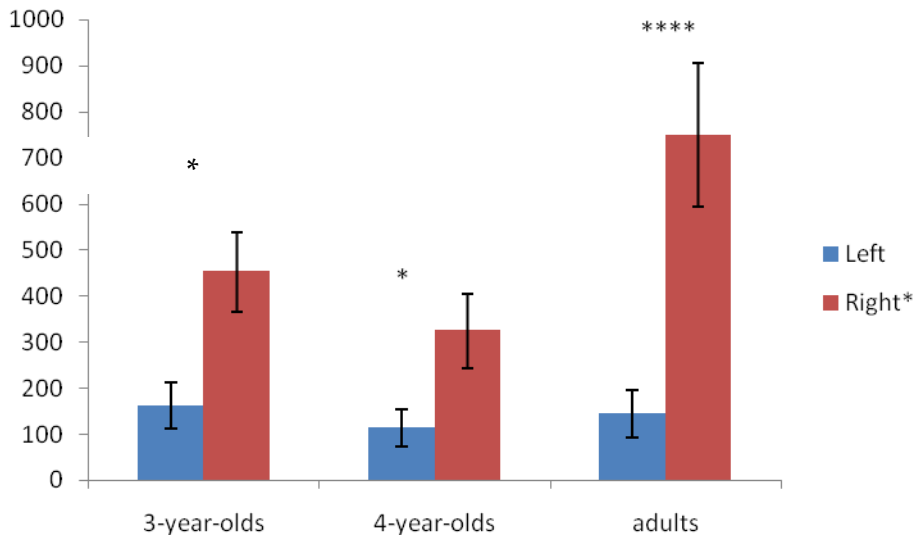


Figure 5: Mean looking time to the left and right in the TB trial of study 2

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

Figure 6 shows the average x positions of all age groups in the true belief trial. T-tests were conducted to test in the critical moment (excluding the first second) to test whether their average looking points were different from the middle line ($x=373$, where Lisa disappears in the forest). Results of t-tests showed that all ages looked longer at the correct prediction: adults: $mean=430.42$ $sd=57.99$, $t(25)=5.05$, $p < .0001$; age 4: $mean=445.96$, $sd=82.72$, $t(22)=4.23$, $p < .0005$; age 3: $mean=444.51$, $sd=103.31$, $t(27)=3.66$, $p < .005$).

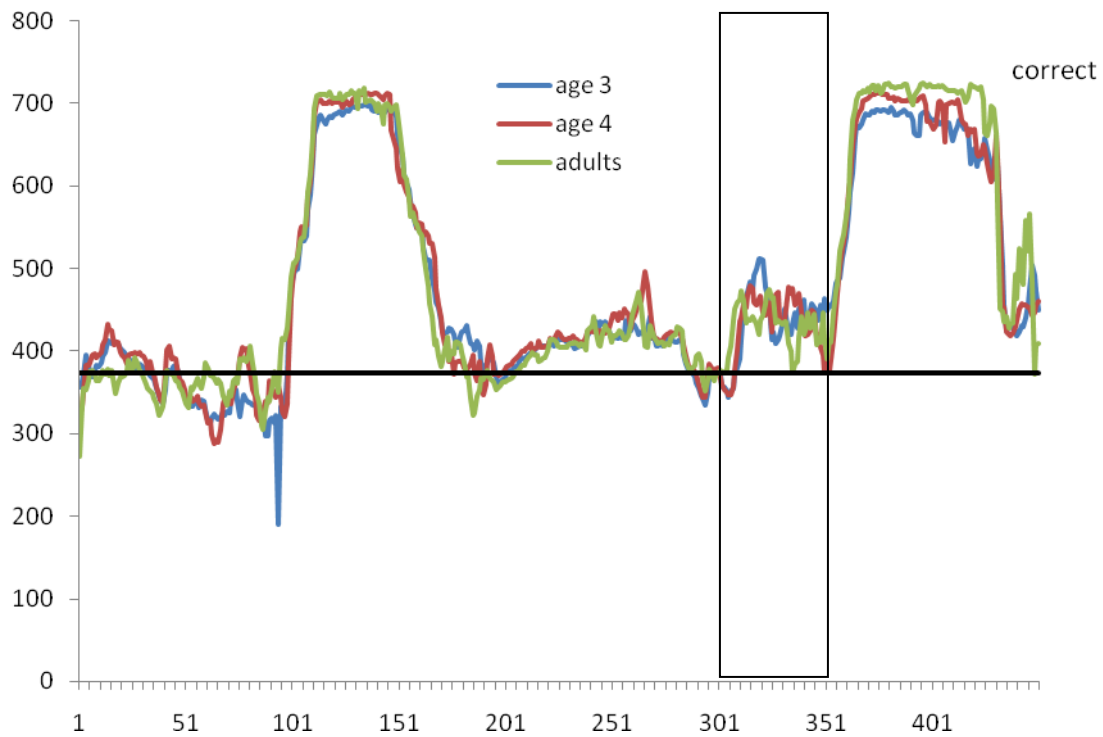


Figure 6: Eye movements of all age groups across time in the true belief trial of study 2

This test was conducted to each frame of the critical moment (per 100 ms, excluding the first second) in all ages. Error bars show the two standard errors form each looking point. Figures 7-9 show that all ages looked at the correct prediction most of the time during the critical moment.

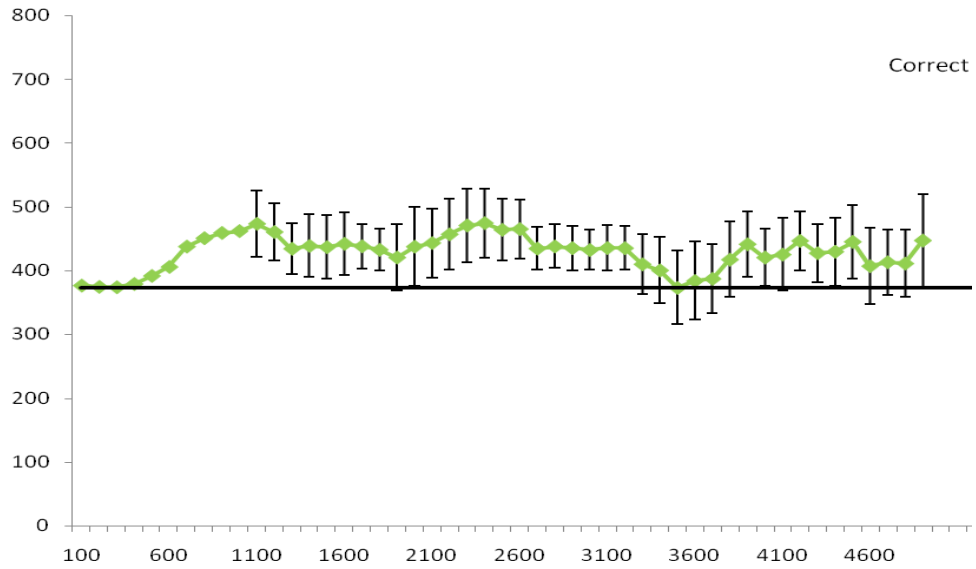


Figure 7: Eye movements of adults across time in the critical moment in the true belief trial of study 2

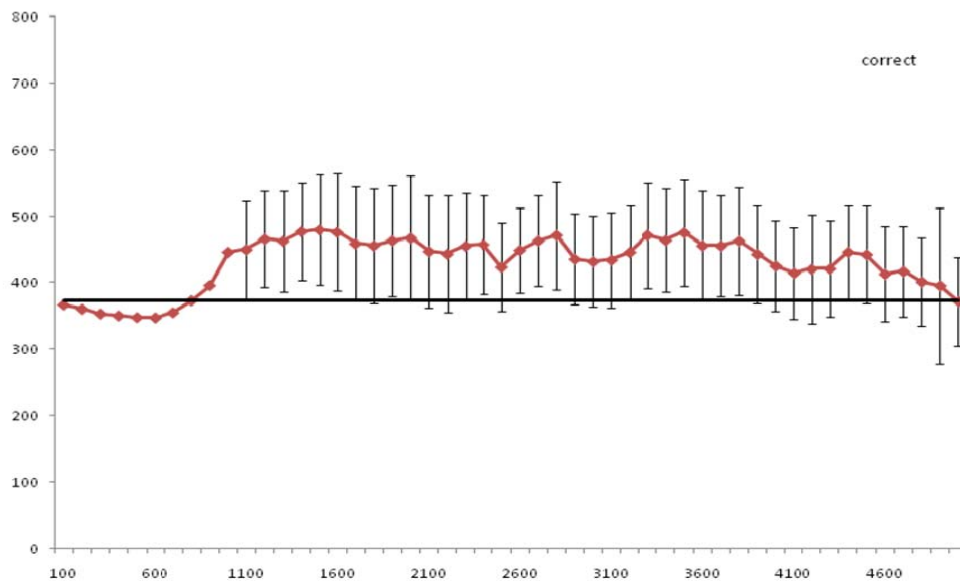


Figure 8: Eye movements of 4-year-olds across time in the critical moment in the true belief trial of study 2

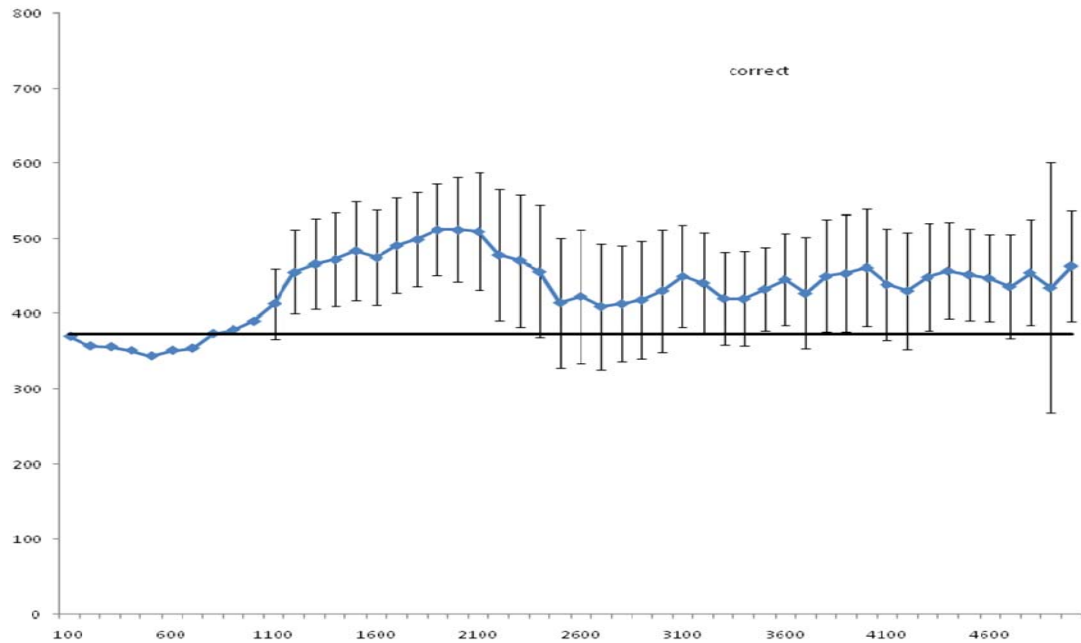


Figure 9: Eye movements of 3-year-olds across time in the critical moment in the true belief trial of study 2

False Belief non-complete Condition

In the second trial, the non-complete false belief condition, the car moved on its own from left to right. The correct prediction for Lisa's movement should be the left path. Results of the first gazes showed that 14 out of 25 3-year-olds, 15 out of 22 4-year-olds and 24 out of 27 adults made the first gazes towards the correct prediction in the critical moment.

Total looking time at the left or right were compared across subjects using a paired t-test (as shown in Figure 10). Results showed that adults looked at the correct prediction more than the incorrect prediction $t(21)=-2.97, p<.01$; 3-year-olds $t(24)=-1.66, p=.11$ and 4-year-olds $t(22)=-1.45, p=.17$. both showed no predictions to the left or right paths.

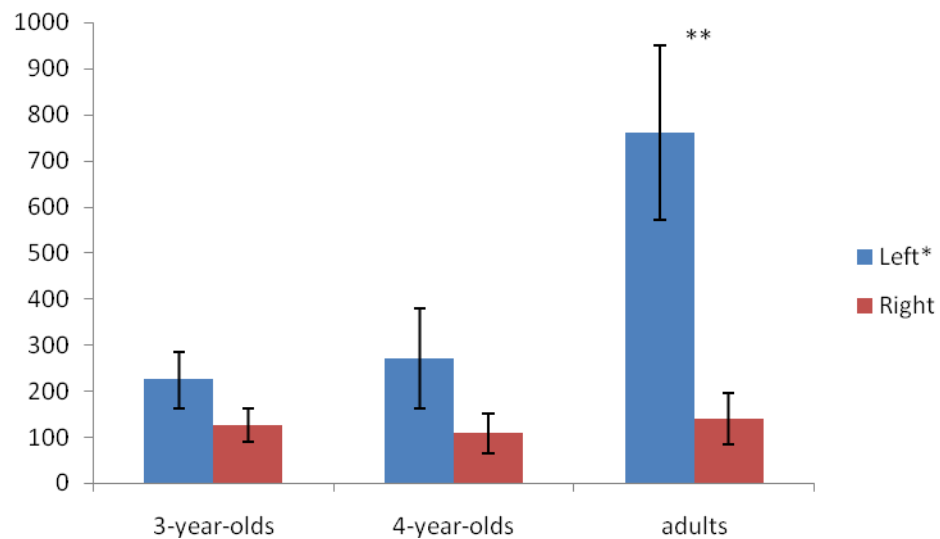


Figure 10: Mean looking time to the left and right in the first FB trial of study 2

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; ***** $p<.0001$

Eye movements across time were shown in the Figure 11. T-test showed that there is no clear predictions to the left or right paths for all ages: adults: $mean=352.1$; $sd=68.78, t(25)=-1.55, p=.13$; age 4: $mean=380.57, sd=146.06, t(22)=0.25, p=.81$; age 3: $mean=390.13, sd=154.79; t(28)=0.60, p=.56$.

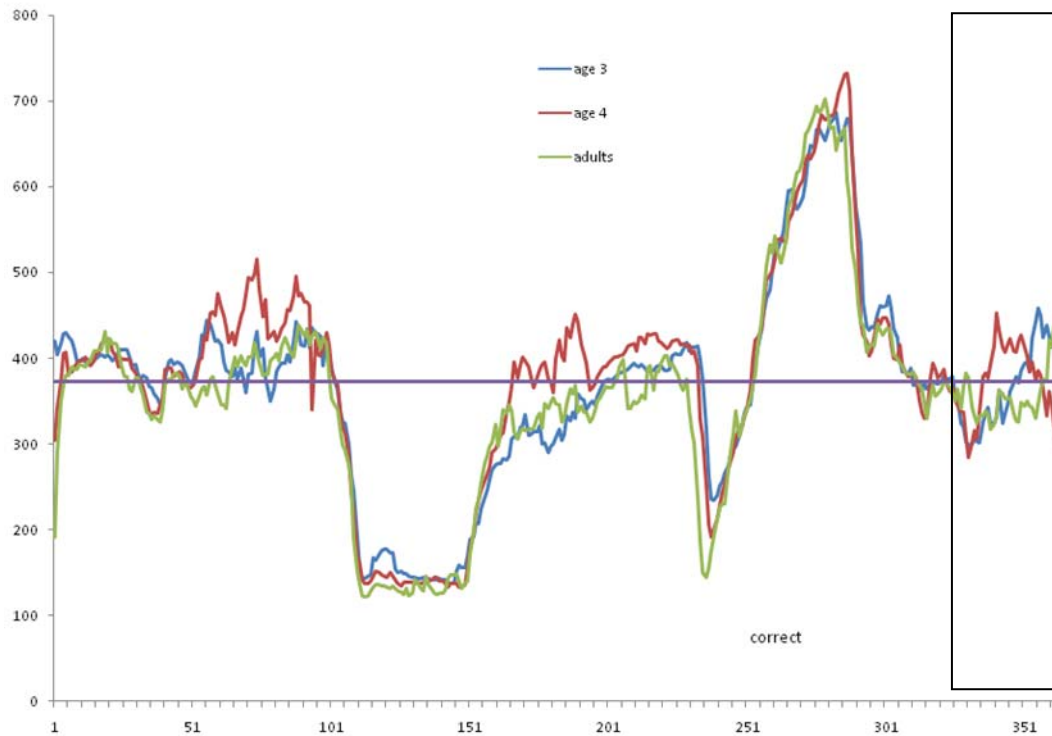


Figure 11: Eye movements of all age groups across time in the false belief non-complete trial of study 2

The results from the time-based analysis (Figure 12-14) showed that (a) adults looked at the correct answer at several time points during the critical moment; 4- and 3-year-olds showed no predictions to the left or right paths during the critical moment.

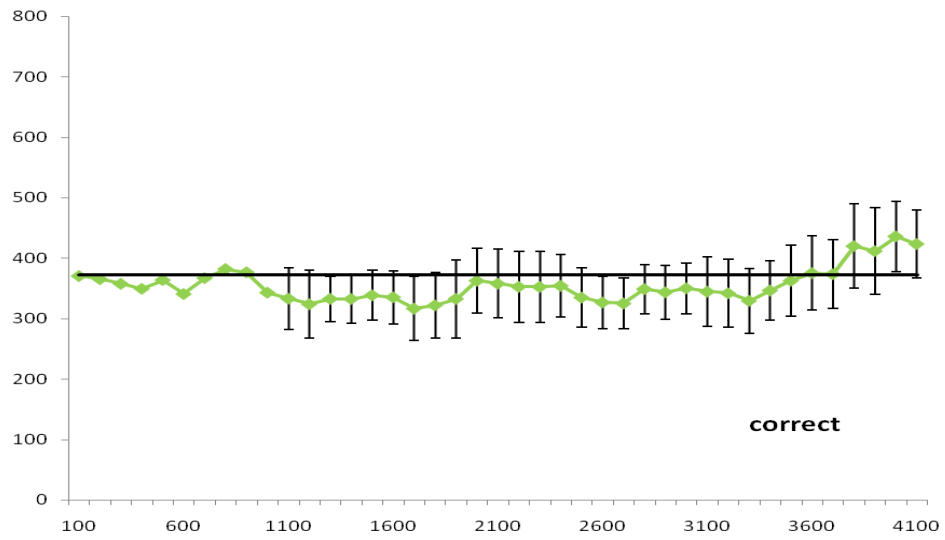


Figure 12: Eye movements of adults across time in the critical moment in the false belief non-complete trial of study 2

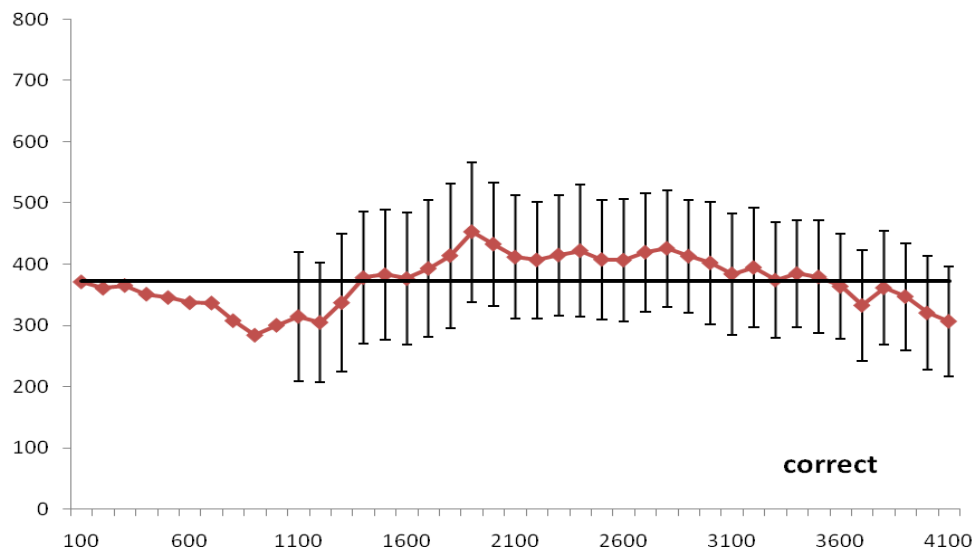


Figure 13: Eye movements of 4-year-olds across time in the critical moment in the false belief non-complete trial of study 2

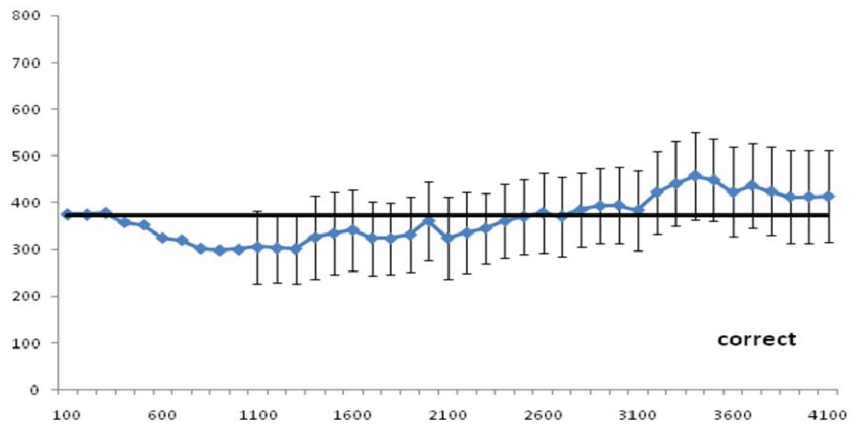


Figure 14 Eye movements of 3-year-olds across time in the critical moment in the false belief non-complete trial of study 2

False Belief Complete Condition (first trial of false belief condition)

In the complete false belief condition, the car moved from right to left. The correct prediction for Lisa's movement should be the **right** side of the screen. Results of the first gazes showed that 9 out of 28 3-year-olds, 10 out of 19 4-year-olds and 16 out of 26 adults made the first gazes towards the correct prediction in the critical moment.

Total looking time at the left or right exits were compared across subjects using a paired t-test. Results showed that adults looked at the correct prediction more than the incorrect prediction $t(25)=3.36, p<.005$; 3-year-olds looked at the incorrect answer more than the correct one $t(25)=-2.63, p<.05$; 4-year-olds showed no predictions to the left or right paths $t(18)=-1.82, p=.08$.

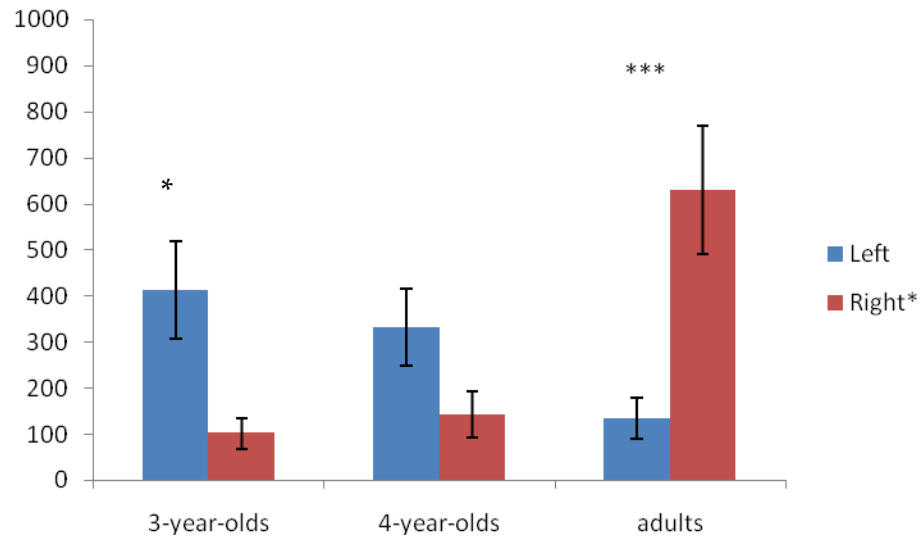


Figure 15: Mean looking time to the left and right in the second FB trial of study 2

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

Eye movements in the third trial across time were shown in the Figure 16. T-test showed that (a) adults looked at the correct answer more than the incorrect one $mean=446.14$; $sd=72.49$, $t(27)=5.43$, $p < .0001$; (b) 4-year-olds showed no predictions to the left or right side of the screen $mean=386.79$, $sd=107.33$, $t(22)=0.62$, $p=.54$; (c) 3-year-olds looked at the incorrect answer (the location of the car) more than the correct answer (where Lisa thinks the car is) $mean=329.31$ $sd=101.29$; $t(26)=-2.24$, $p < .05$.

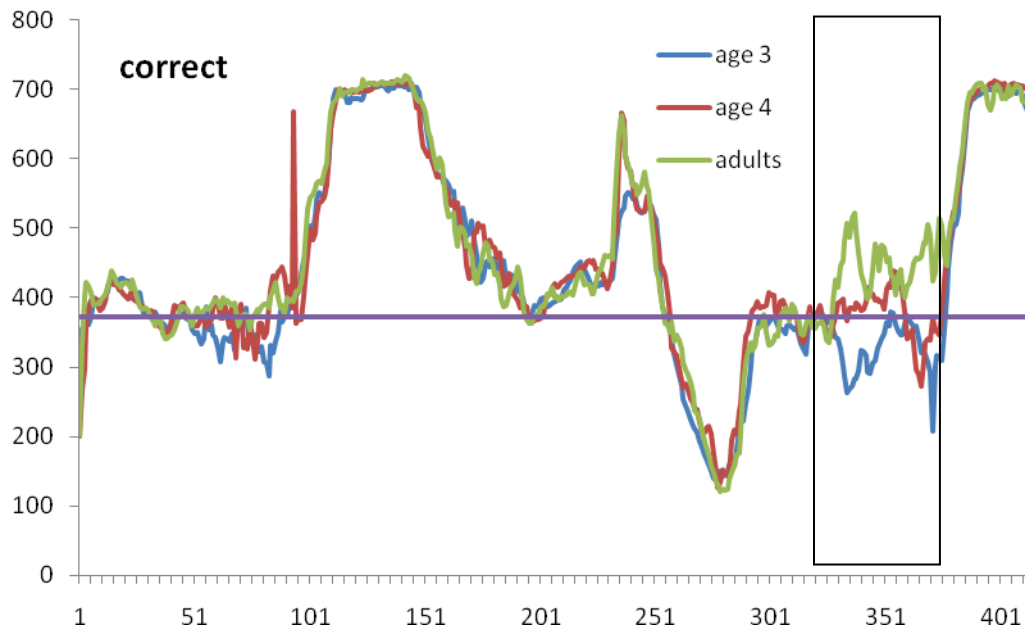


Figure 16: Eye movements of all age groups across time in the false belief complete trial of study 2

The results from the time-based analysis showed that (a) adults switched to the correct answer of the screen from middle line most of time during 1s to 5s; (shown in the Figure 17) (b) 4-year-olds showed no predictions before they looked at the incorrect answer only at 4400ms; (see Figure 18) (c) 3-year-olds looked at the incorrect answer (the location of the car) from middle line during 1100ms-1700ms and 2100-2300, and after 4600ms. (See the Figure 19)

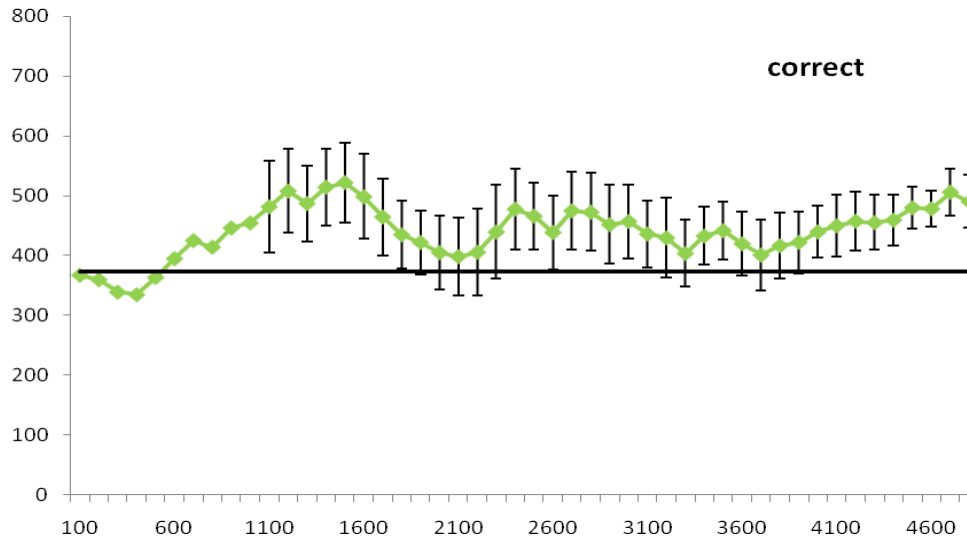


Figure 17: :Eye movements of adults across time in the critical moment in the false belief complete trial of study 2

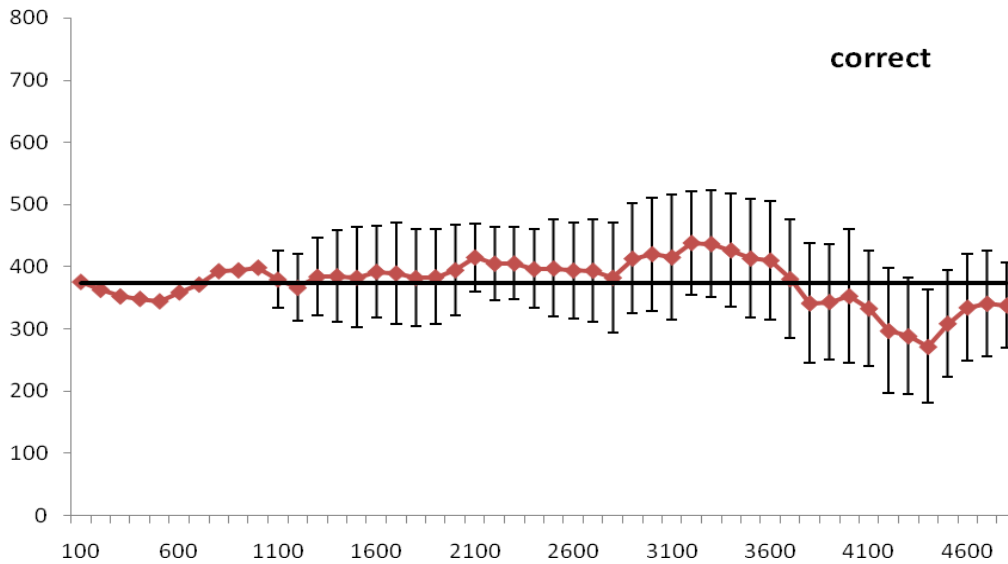


Figure 18: Eye movements of 4-year-olds across time in the critical moment in the false belief complete trial of study 2

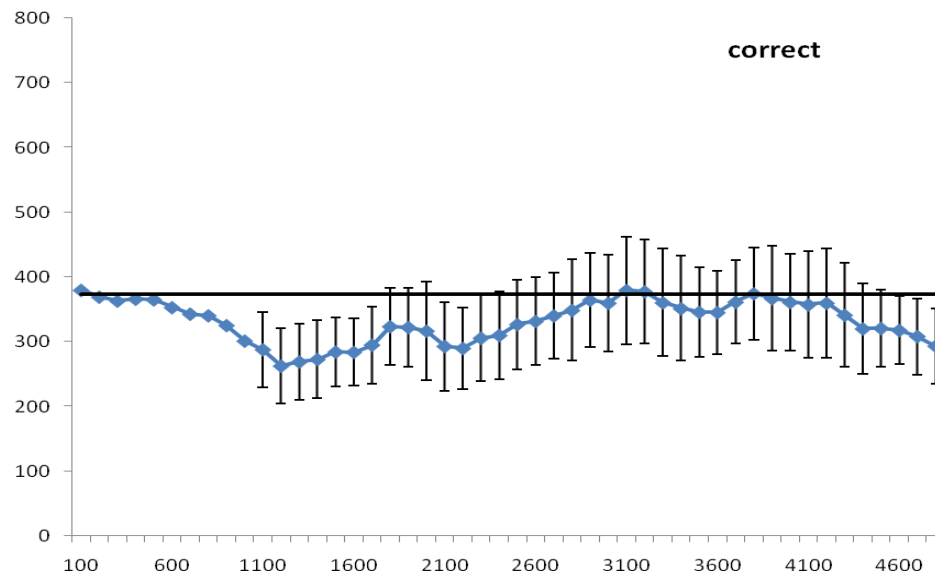


Figure 19: Eye movements of 3-year-olds across time in the critical moment in the false belief complete trial of study 2

Relations between performance in the ToM reasoning and traditional ToM performance

Eye movements of 4-year-olds did not show any predictions. One possibility is that they scanned the scene randomly. It is also possible that those who answered verbal ToM questions behaved like adults and those who failed verbal ToM questions behaved like 3-year-olds.

To test the second possibility, Figure 13 was separated into two lines: one for those who can correctly answer the prediction question, the key question in the traditional ToM task (“Where will Lisa go to look for the car?”), and one for those who cannot. Results showed in the Figure 20. Eight out of 25 4-year-olds correctly answered

this question and 17 gave incorrect answers. Results of an ANOVA showed 4-year-olds who gave correct answers to the prediction question looked more on the correct prediction according to Lisa's FB, $mean=273$, $se=52.97$, than the ones who failed to answer the prediction question $mean=413.03$, $se=31.47$, $F(1,21)=5.15$, $p<.05$.

Time-based analysis showed that there was dissociation from 1000ms to 3600ms between those 4-year-olds who passed the prediction questions and those who failed (as shown in Figure 20). Those 4-year-olds who passed the prediction questions made correct predictions like adults: they looked at the left path from 1100-1400ms, and 2600-3400ms. Those who failed the prediction questions made incorrect predictions: they looked at the right path (where the car actually is) from 1900-2100ms, and 2300-3100ms.

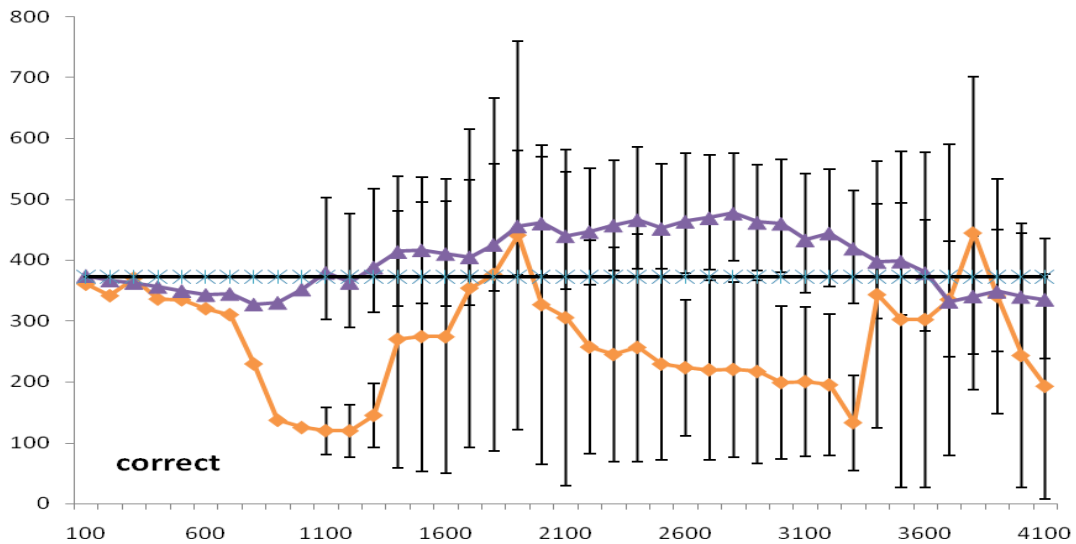


Figure 20: Eye movements of 4-year-olds across time in the false belief non-complete trial (pass—orange line; fail-purple line)

A chi-square test based on 4-year-old children's performance in the verbal answers and their eye movements in this trial showed that: there was a correlation between children's verbal answers to the prediction question and their total looking time at the correct answer (left AOIs: the chair and the left path) $\chi^2=3.70, p=.05$. Particularly, all 4-year-olds who made incorrect predictions in their eye movements (8 out of 22) also gave incorrect answers to the prediction question; for those who made correct predictions in their eye movements (14 out of 22), 9 of them answered the prediction question correctly, and 5 incorrectly. A similar, although not as significant, correlation, was found in 4-year-olds' verbal answers to the prediction question and their first gaze at the left AOIs (the chair and the left path) during the critical period $\chi^2=3.02, p=.08$.

This correlation was further tested by separating the Figure 13 with the general ability of explicit ToM and performance in the elicited imitation task. First, four-year-olds were divided into two groups: the Pass group whose performance of answering all mind-related questions was above average (11 out of 25, average of correct rates=.65), and Fail group who were below average (14 out of 25). Results of an ANOVA showed a marginal difference between the pass group and the fail group $F(1,21)=3.69, p=.068$. A time-based analysis showed the similar pattern (though not as strong as) as the dissociation divided by the prediction question (as shown in Figure 21 & 22).

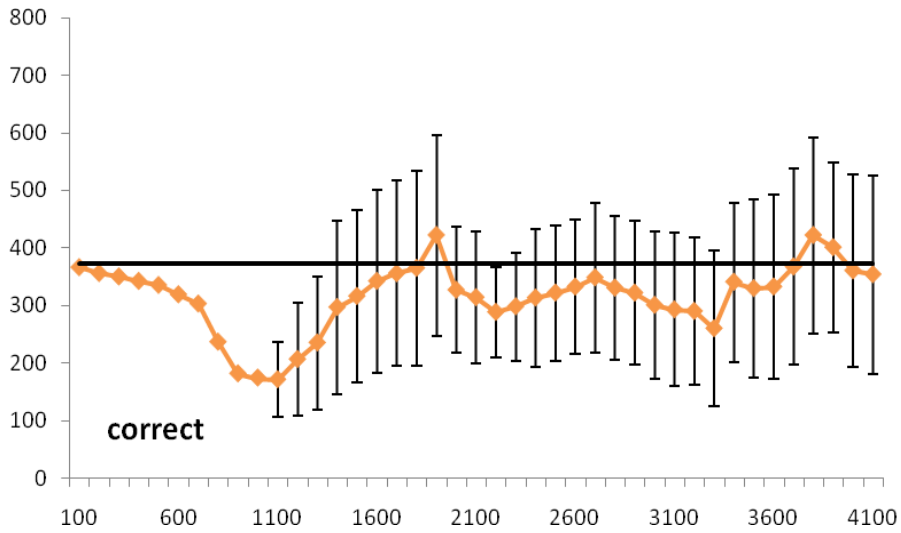


Figure 21: Eye movements of 4-year-olds who passed the traditional ToM questions across time in the false belief non-complete trial of study 2

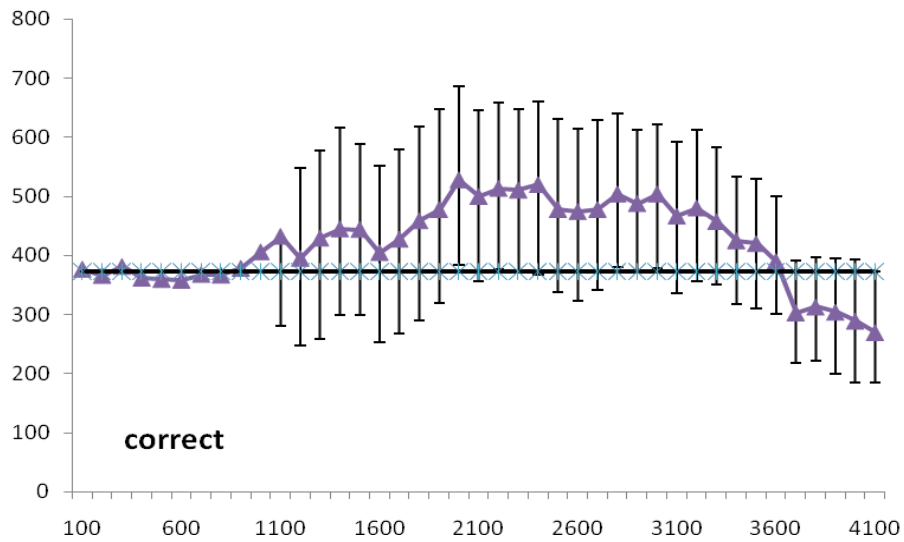


Figure 22: Eye movements of 4-year-olds who failed the traditional ToM questions across time in the false belief non-complete trial of study 2

Then four-year-olds were further divided into two groups: the Pass group whose elicited imitation performance was above average (12 out of 22, average correct rates=.46), and Fail group below average (10 out of 22). Results of an ANOVA showed no difference between the pass group and the fail group $F(1,18)=.04, p=.84$. A time-based analysis showed similar patterns of eye movements between the pass and the fail groups (as shown in Figure 23 & 24).

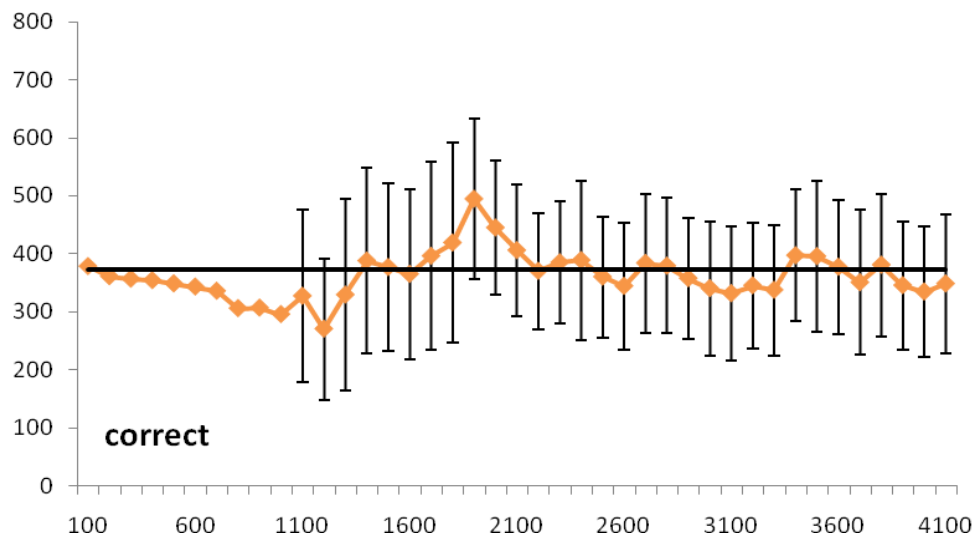


Figure 23: Eye movements of 4-year-olds who passed the elicited imitation task across time in the false belief non-complete trial of study 2

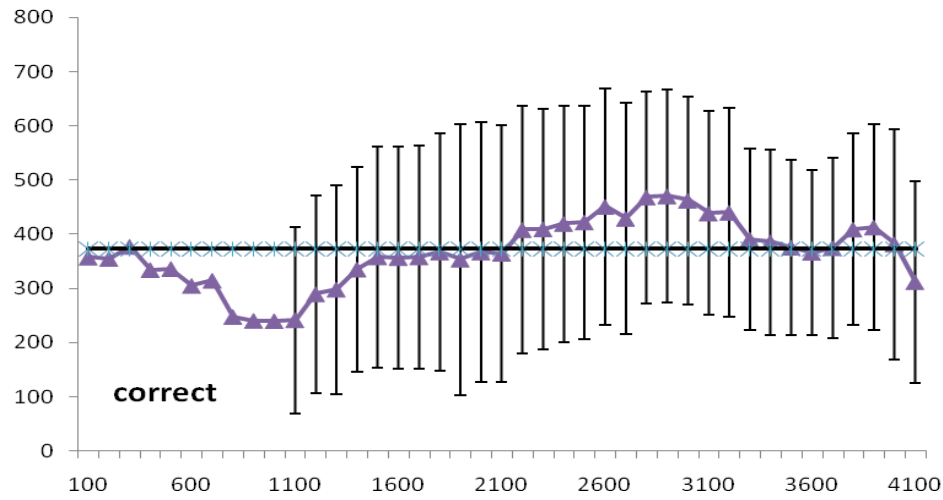


Figure 24: Eye movements of 4-year-olds who failed the elicited imitation task across time in the false belief non-complete trial of study 2

Interim discussion on eye movements in the change-of-location task

This current task was designed to examine children’s non-verbal ToM reasoning, and its relation to the traditional ToM performance. Anticipatory eye movements in this nonverbal change-of-location task provided a sensitive on-line measurement for the non-verbal ToM reasoning processes. Results showed that adults, but not 3- and 4-year-olds, showed a non-verbal ToM reasoning. Children’s anticipatory eye movement patterns in this task were associated with their performance in the traditional ToM questions. This suggested that part of the reason that young children failed in the traditional ToM questions is that they failed to keep track of social events in real time.

Although the starting age of implicit understanding of false belief in the current findings is different from findings of Clements & Perner (1994), their proposed developmental pattern that implicit understanding of false belief precedes the explicit one is confirmed in the current findings. All 4-year-olds who failed to show implicit understanding of false belief also failed to answer explicit ToM questions; some 4-year-olds showed implicit but not explicated understanding of false belief, and some of them show both types of understanding. The age difference between the current study and Clements & Perner (1994)'s finding suggests that children's performance of implicit task is influenced by environmental variables such as some details of experimental settings. For example, Clements & Perner (1994) used verbal prompts to indicate the critical prediction period, but current study did not use verbal prompts.

The current study is also different from previous studies on infants' understanding of beliefs (Onishi & Baillargeon, 2005; Southgate, et al., 2007) in terms of the purpose of the studies, the nature of the tasks and the measurements. First, Onishi-Baillargeon studies were designed to investigate the origin of the implicit understanding of beliefs, or the earliest signs of ToM. Compared to traditional ToM stories, they used a simplified way to show infants a series of events that may involve other people's mental states. The current study, on the other hand, was designed to investigate why young children fail in the traditional task. The hypothesis is that children have to possess an

automatic and real-time process, which cannot be shown in the verbal questions, to succeed in the traditional ToM tasks. For this purpose the story in this current task keeps the complexity of the traditional stories and thus may be in higher demand than the simplified stories for infants. The scenario of the cartoon movie was designed to tell the stories as close as possible to the stories in the traditional stories. Also the cartoons were shown in relatively natural settings. Compared to the Onishi-Baillargeon studies, the experimental settings in the current study were closer to the real social situation and the traditional ToM tasks. Therefore, the ToM reasoning shown in the current study was more likely to be an automatic process to a social situation.

Second, the current task was designed to assess a real-time process of ToM, which is done by using anticipatory eye movements. In contrast, Onishi & Baillargeon (2005) used the violation-of-expectation paradigm, in which infants' looking time reveals their surprise on the impossible events. This paradigm has to expose participants to some "after-fact incongruent events" to measure their recognitions of them (Southgate, et al., 2007). Therefore, the violation-of-expectation paradigm is not a measurement for the real-time process, which is a concern in the current study according to the purpose of it. On the other hand, the anticipatory-looking paradigm used in the current task assesses the predictions to actor's actions before the endings come out. The anticipatory eye movements measured in this paradigm reveal real-time processes of ToM reasoning.

Third, although Southgate, et al. (2007) also used the anticipatory-looking paradigm, they did not provide the time-based analysis for the real-time processes. Both of these two previous studies measured and compared total looking time over the correct prediction and the competitive one. In addition to this, the current study also provides a time-based analysis for real-time processing patterns of ToM reasoning across time. Southgate, et al. (2007) looked at the overall looking time during the first 1750 ms. However, as shown in the current study, participants even adults may need some process and response time so that they did not make any predictions in the first second.

Implicit understanding of beliefs has only been shown by a couple of studies and is still under controversy. The current findings suggest that the early understanding of false belief found in infants may be only applied to narrow and constrained situations. Infants may have difficulty in generalizing the understanding of mind in broader situations.

The non-verbal ToM reasoning was also addressed in another task revised from the traditional unexpected-content task (Gopnik & Astington, 1989). Anticipatory eye-movements were analyzed in similar ways (as shown in the following section) as in the current task to reveal the non-verbal ToM reasoning.

2.2.4 Eye Movements in Unexpected-content Task

Eye movements recorded during watching the four trials of unexpected-content videos were analyzed using similar procedures as in the change-of-location tasks. Participants' total looking time to the left and right sides of the screen were coded and compared using a paired t-test. Also the average x position at each time point was compared to the midline where $x=400$. The first second of the critical moment was excluded: because the critical moment starts with a new scene (close-up of two objects side by side), participants had to identify the objects before making predictions. Therefore, the first peak of left and the first peak of right (altogether around one second) will be excluded in this analysis.

There were two conditions in this task, FB-FB and FB-TB condition, each of which has two parts. In the FB-FB condition, actor B did not open the box so that she held a false belief in both two parts of the video; while in the FB-TB condition, B's false belief in the first part was updated into a true belief in the second part by opening the box and discovering the content of it. The first two trials were the two parts from the FB-FB condition, both telling FB stories, equivalent to the traditional unexpected-content task. The third trial, the first part of the FB-TB condition, is still a FB trial, also equivalent to the traditional unexpected-content task. The correct prediction for these three trials would be pointing to the crayon (right). The last trial (the part 2 of the FB-TB video) is a

true belief trial, which is harder than the first three trials, since it requests an understanding of update in mind so that a FB becomes a TB. The correct prediction in this trial should be pointing to the candle (left). All trials have the original object (“crayon” was used to represent crayon/chocolate/cookie) on the right side of the screen and the real object (“candle” was used to represent candle/marble/plate) on the left.

In the first FB trial, the total looking time at the left or right sides were compared across subjects using a paired t-test (as shown in Figure 25). Results showed that adults looked at crayon more than the candle $t(27)=2.70, p<.05$; 3-year-olds $t(28)=-.57, p=.58$ and 4-year-olds $t(23)=-.04, p=.9$ both showed no predictions for the pointing behaviors.

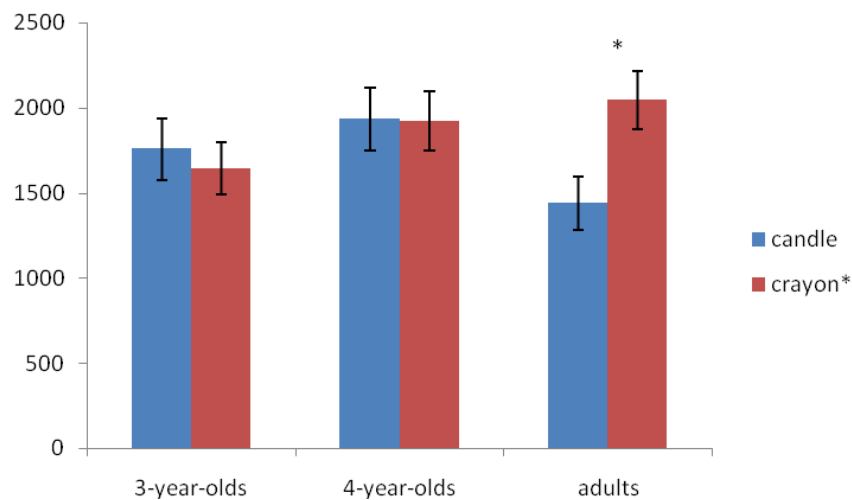


Figure 25: Mean looking time at the crayon and candle in the first FB trial of the unexpected-content task of study 2

Notes: * $p<.05$

T-tests were also conducted to test whether participants were looking to the left or the right side of the screen during the critical moment (excluding the first second). The results of the first FB trial, i.e., the familiarization trial, showed that, 3- and 4-year-olds and adults showed no predictions for the pointing over the critical moment (shown in the Figure 26).

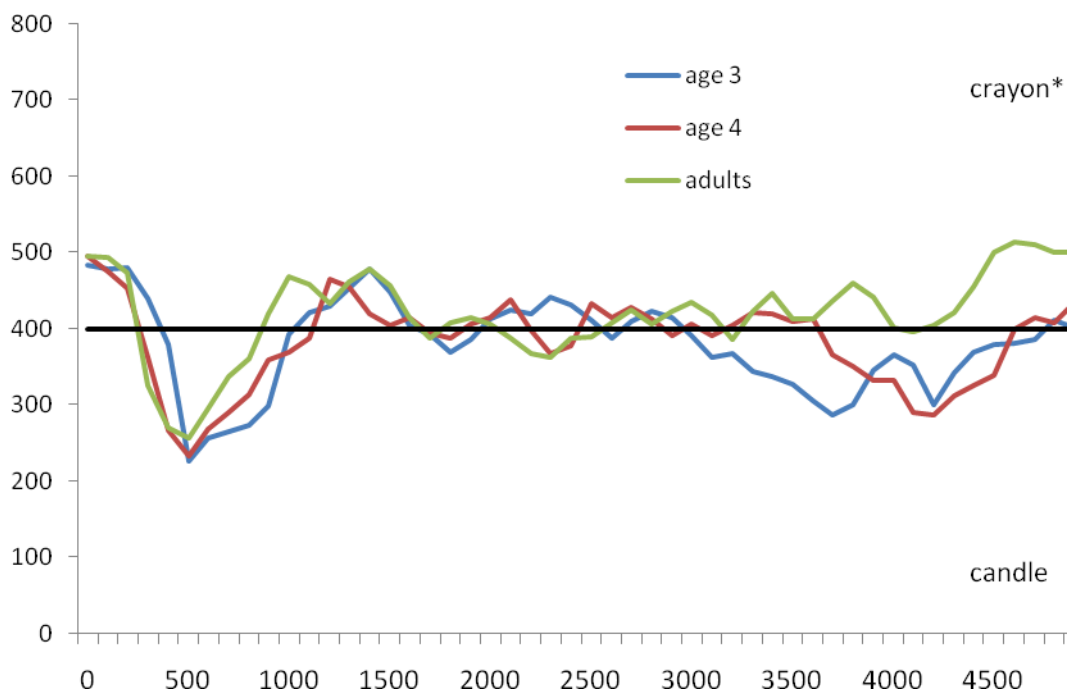


Figure 26: Eye movements of all age groups across time of the critical moment in the familiarization trial of study 1

T-tests were also conducted at every time point (per 100ms), excluding the first second (see Figure 27-29). Results showed that (a) adults did not show predictions of

pointing until 4500ms; after 4500ms, they looked at the correct answer more than the incorrect answer. In light of subsequent results, this is most likely because adults did not know what to expect in the familiarization trial; (b) 4-year-olds looked at the incorrect answer more than the correct answer during 4100-4300ms, and showed no predictions in rest of time over the critical moment; (c) 3-year-olds showed very similar pattern as the 4-year-olds: they looked at the incorrect answer more than the correct answer during 3600-3900ms, and showed no predictions in rest of time over the critical moment

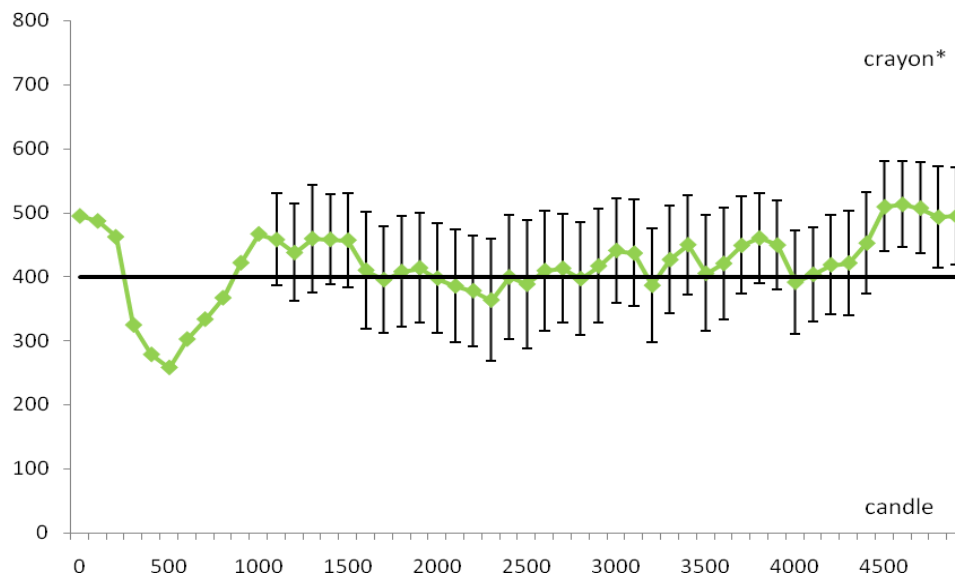


Figure 27: Eye movements of adults per 100 ms of the critical moment in the familiarization trial of study 1

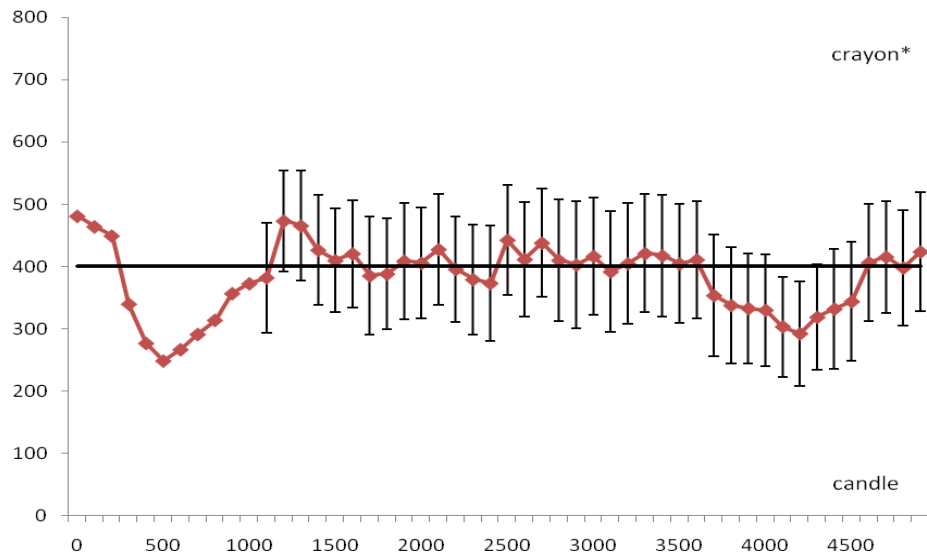


Figure 28: Eye movements of 4-year-olds per 100 ms of the critical moment in the familiarization trial of study 1

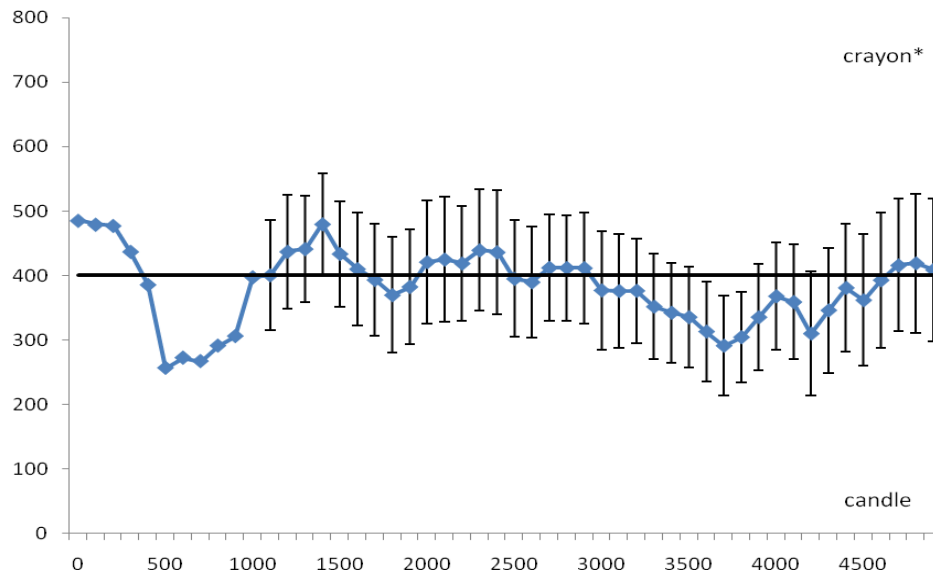


Figure 29: Eye movements of 3-year-olds per 100 ms of the critical moment in the familiarization trial of study 1

Figure 30 shows the total looking time to the left or right during the second part of the FB-FB condition. Results of a paired t-test showed that adults looked at the correct answer more than the incorrect answer at marginal significance $t(26)=1.94, p=.06$; 4-year-olds also looked at the correct answer more than the incorrect answer at marginal significance $t(21)=1.91, p=.069$; 3-year-olds showed no predictions to the correct or the incorrect answers $t(26)=-1.20, p=.24$.

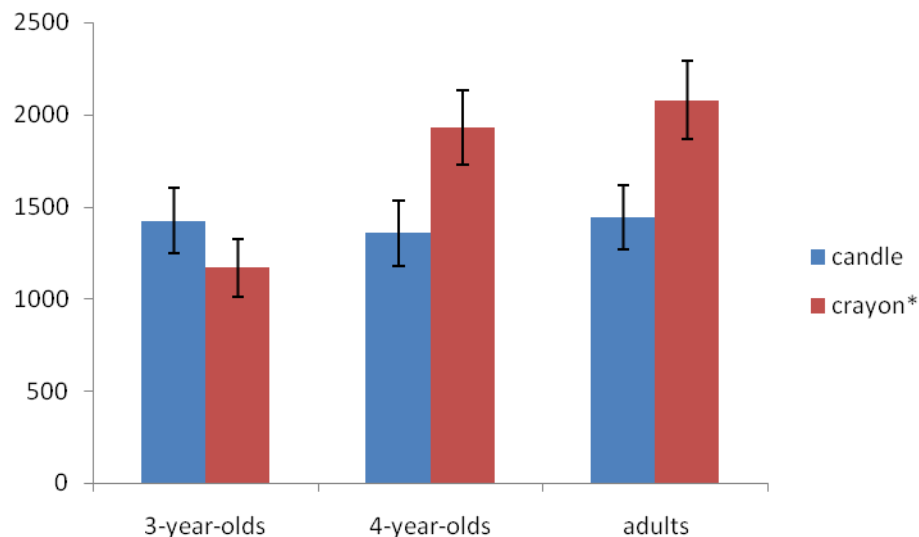


Figure 30: Mean looking time at the crayon and candle in the second part of the FB-FB condition of the unexpected-content task of study 1

The average x positions in the critical moment were shown in the Figure 31.

During the critical moment of the second part of the FB-FB trial, adults looked slightly

longer to the correct answer (on the right) $mean=442.69$, $sd=117.37$, $t(26)=1.89$, $p=.07$; 3- and 4-year-olds showed no predictions for the correct or incorrect answers: 3-year-olds: $mean\ x=371.57$, $sd=141.99$, $t(26)=-1.04$ $p=.30$; 4-year-olds: $mean\ x=439.70$, $sd=125.84$, $t(21)=1.48$ $p=.15$).

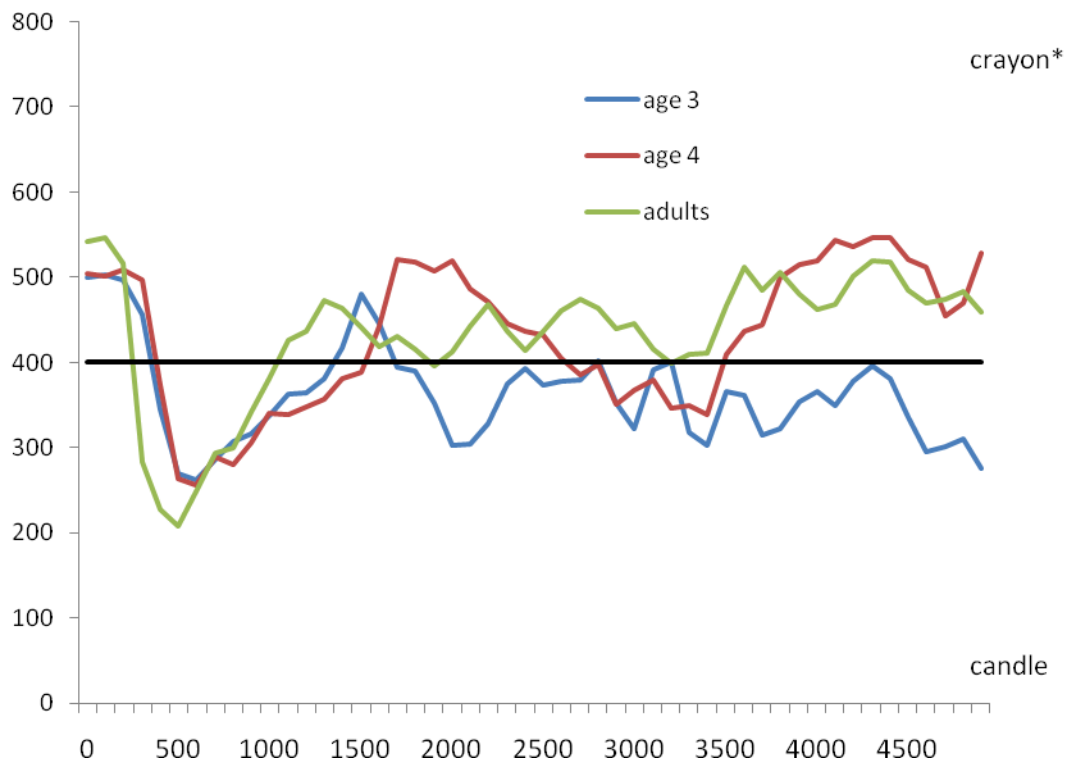


Figure 31: Eye movements of all age groups across time of the critical moment in the second part of the FB-FB condition of study 1

The tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 32-34). The results showed that (a) adults did not show predictions

for the correct or incorrect answers until 3600ms; they looked at the correct answer more than the incorrect answer during 3600-3900ms and 4200-4900ms; (b) 4-year-olds looked at the correct answer more than the incorrect answer during 1700-2100ms, and then went back to the middle for a while, then looked at the correct answer again during 3900-4500ms; (c) 3-year-olds showed no preference of the correct answer or the incorrect answer across the 4 seconds of the critical moment.

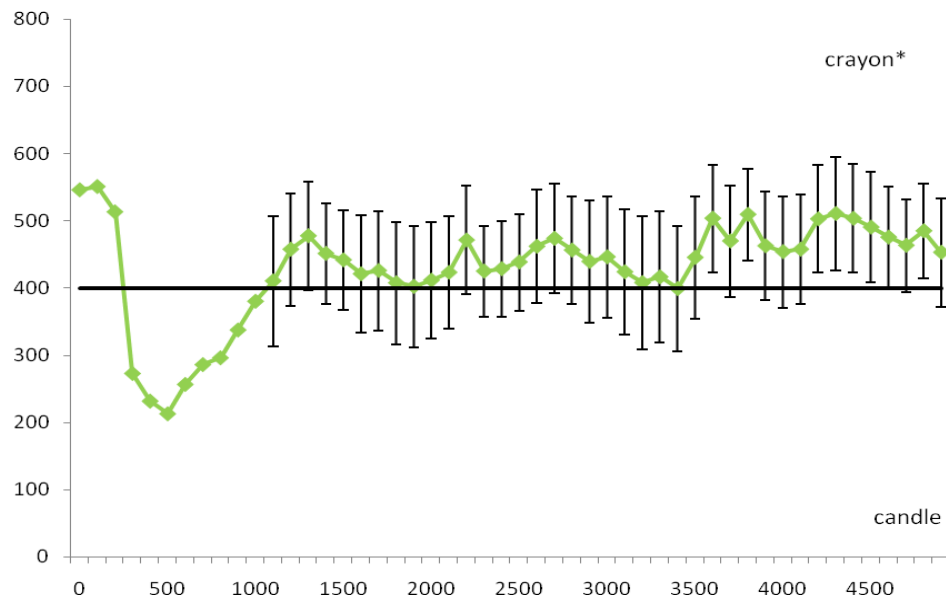


Figure 32: Eye movements of adults per 100 ms of the critical moment in the second part of the FB-FB condition of study 1

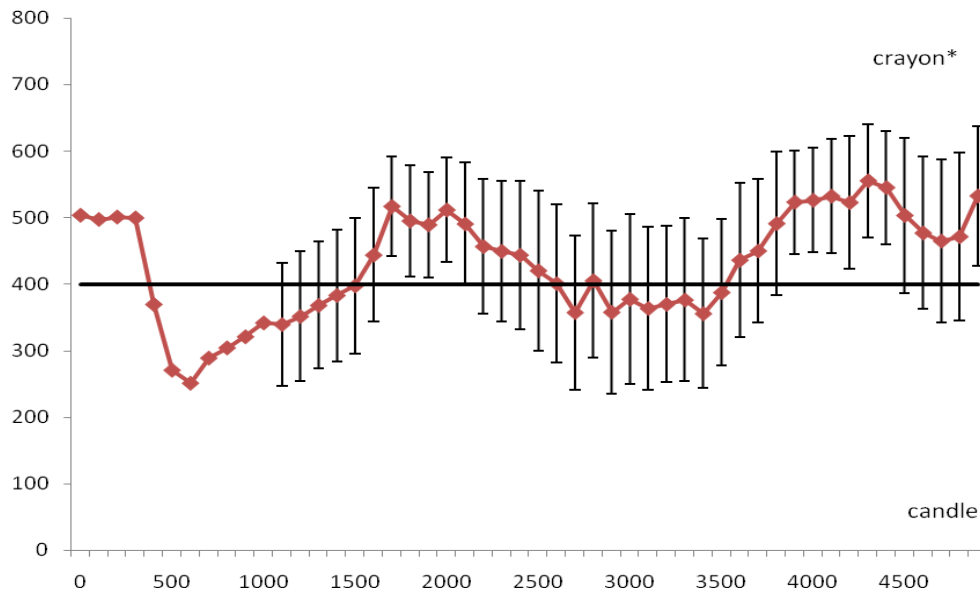


Figure 33 Eye movements of 4-year-olds per 100 ms of the critical moment in the second part of the FB-FB condition of study 1

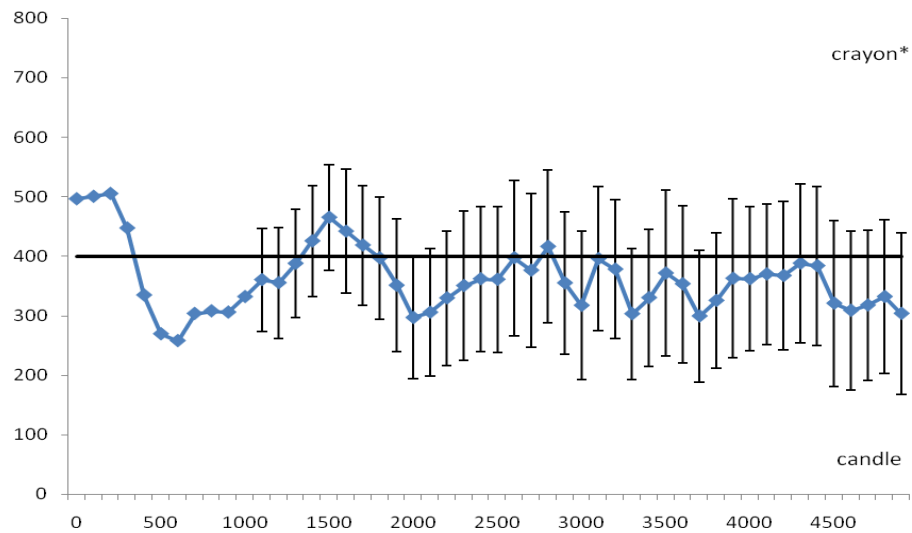


Figure 34: Eye movements of 3-year-olds per 100 ms of the critical moment in the second part of the FB-FB condition of study 1

Figure 35 shows the total looking time to the left or right during the first part of the FB-TB trial. Results of a paired t-test showed that adults looked at the correct answer on the right more than the incorrect answer on the left $t(26)=2.34, p<.05$; 3-year-olds showed no preference to the correct answer or the incorrect answer $t(24)=-1.42, p=.17$; 4-year-olds also showed no preference to the correct or the incorrect answers $t(22)=.55, p=.59$.

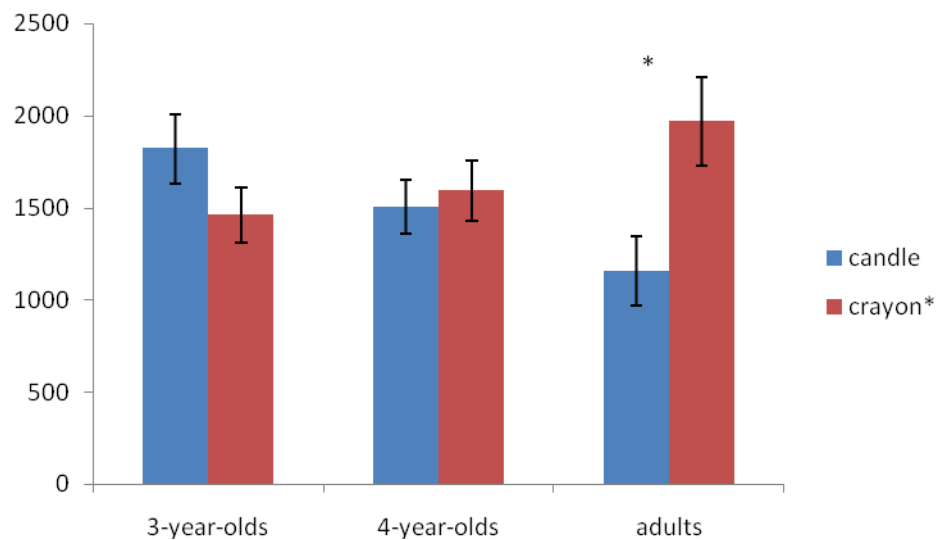


Figure 35: Mean looking time at the crayon and candle in the first part of the FB-TB condition of the unexpected-content task of study 1

Notes: * $p<.05$

The averaging looking points in the critical moment were shown in the Figure 36. In the first part of the FB-TB condition), adults looked at the incorrect answer

significantly more than the correct answer $mean=461.34, sd=155.32, t(25)=2.78, p<.05$; 3- and 4-year-olds showed no predictions for the correct answer or the correct answer: 3-year-olds: $mean x=400.53, sd=107.22, t(24)=.02, p=.98$; 4-year-olds: $mean x=425.28, sd=81.60, t(22)=1.49, p=.15$.

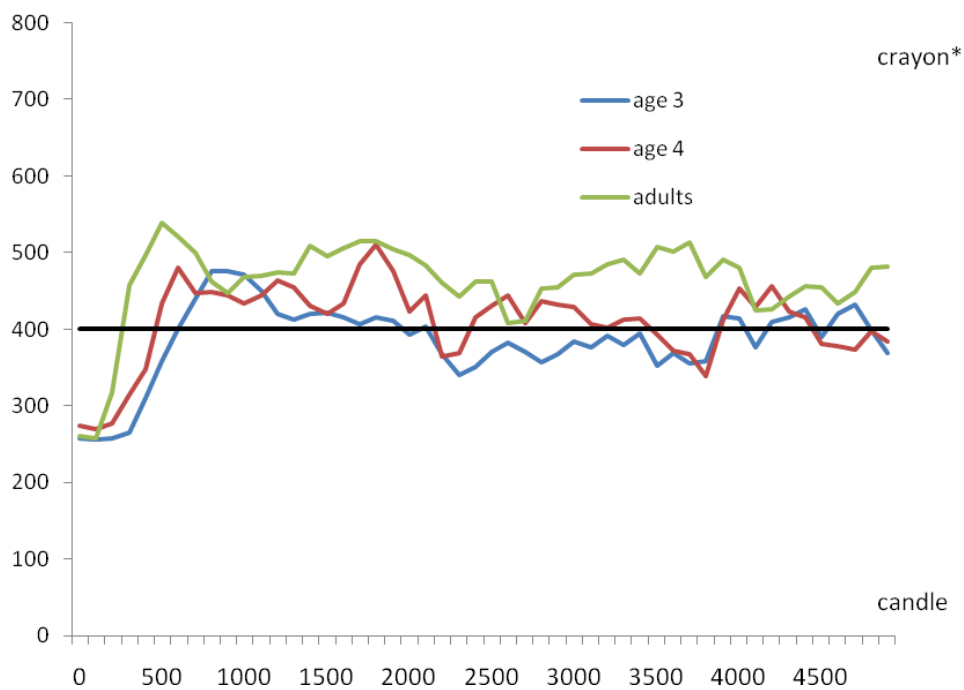


Figure 36: Eye movements of all age groups across time of the critical moment in the first part of the FB-TB condition of study 1

The tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 37-39). The results showed that (a) adults looked at the correct

answer more than the incorrect answer during 1400-2200ms, then went back to the middle for a while, then looked at the correct answer again during 3200-4000ms; (b) 4-year-olds looked at the correct answer more than the incorrect answer during 1800-2000ms; (c) 3-year-olds showed no preference of the correct answer or the incorrect answer across the 4 seconds of the critical moment.

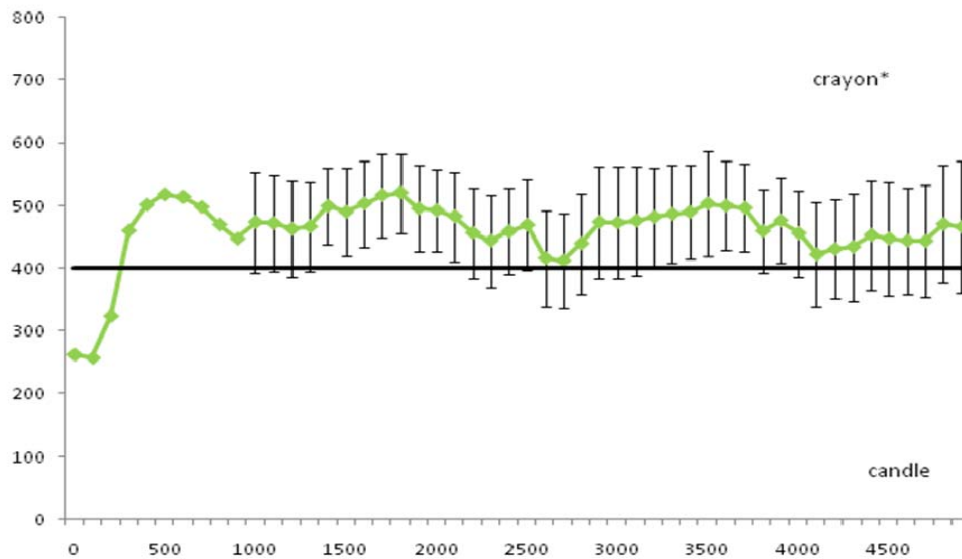


Figure 37: Eye movements of adults per 100 ms of the critical moment in the first part of the FB-TB condition of study 1

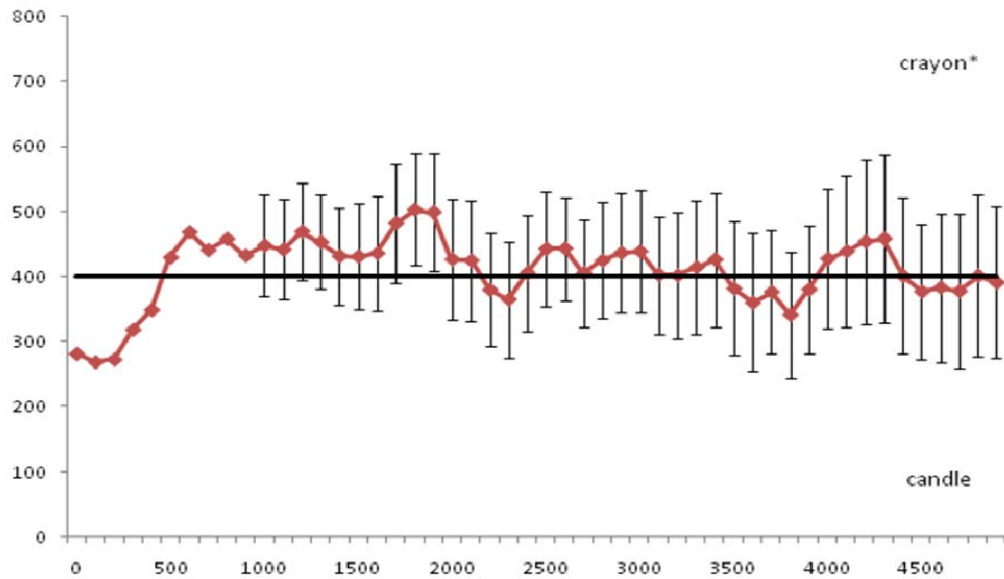


Figure 38: Eye movements of 4-year-olds per 100 ms of the critical moment in the first part of the FB-TB condition of study 1

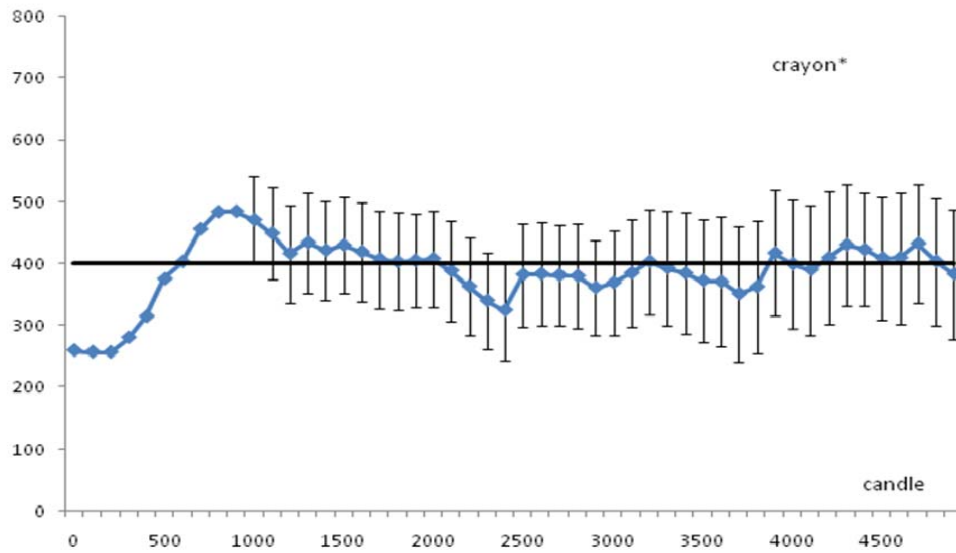


Figure 39: Eye movements of 3-year-olds per 100 ms of the critical moment in the first part of the FB-TB condition of study 1

Figure 40 shows the total looking time to the left or right during the second part of FB-TB trial, in which B looked into the box and discovered the content of it, so that B's FB had been updated into to a TB. Results of a paired t-test showed that adults showed no preference to the correct answer and the incorrect answer $t(27)=-1.29, p=.21$; 3-year-olds showed no preference to the correct answer or the incorrect answer $t(13)=-.06, p=.95$; 4-year-olds also showed no preference to the correct answer or the incorrect answer $t(8)=-1.89, p=.096$.

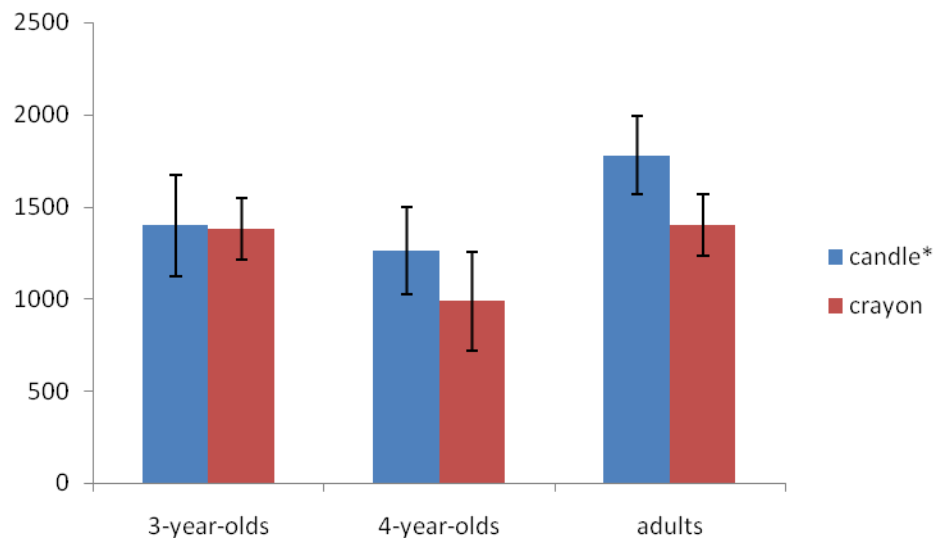


Figure 40: Mean looking time at the crayon and candle in the second part of the FB-TB condition of the unexpected-content task of study 1

The averaging looking points in the critical moment were shown in the Figure 41. Results showed that in the second part of the FB-TB condition, adults and 3-year-olds did not show any predictions of pointing to the correct or the incorrect answers: *adults: mean x=412.52, sd=115.78, t(27)=.57, p=.57; 3-year-olds: mean x=428.68, sd=103.85, t(13)=1.03, p=.32*. Only 4-year-olds showed looked marginally longer at the correct answer than the incorrect answer *mean=335.53, sd=93.29; t(8)=-2.07, p=.07*.

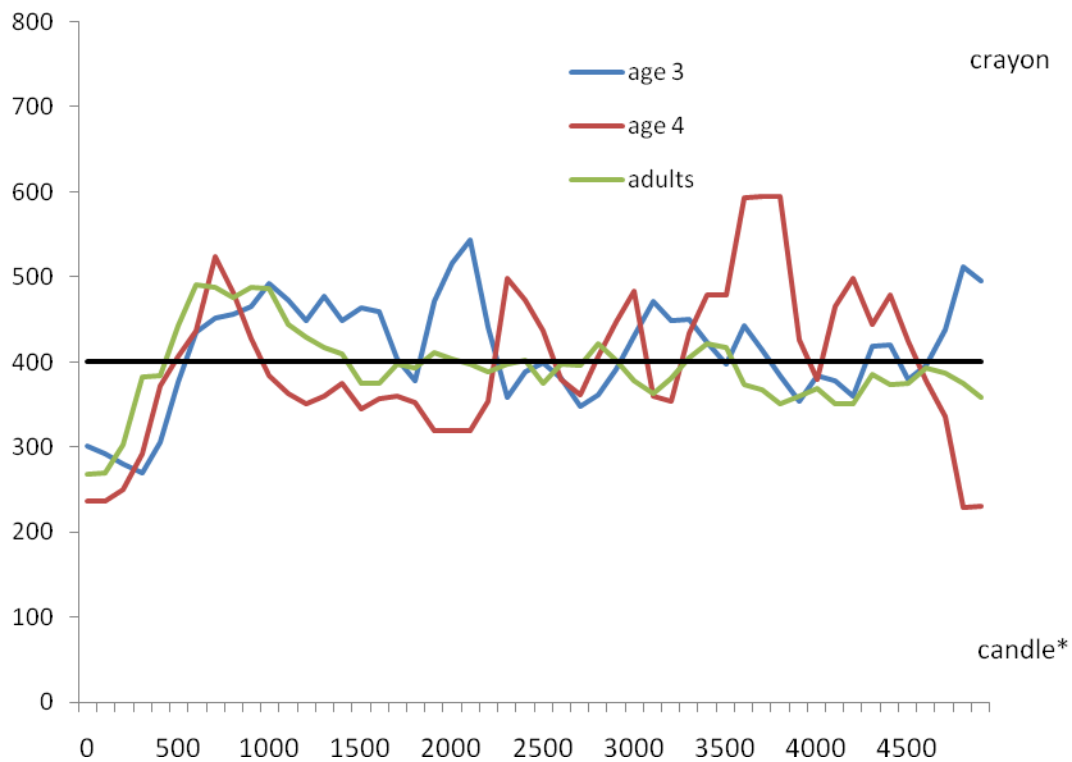


Figure 41: Eye movements of all age groups across time of the critical moment in the second part of the FB-TB condition of study 1

T-tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 42-44). The results showed that: adults and 4-year-olds did not show predictions of pointing to the correct or the incorrect answer across the critical moment; (c) 3-year-olds looked at the incorrect answer more than the correct answer during 2000-2200ms.

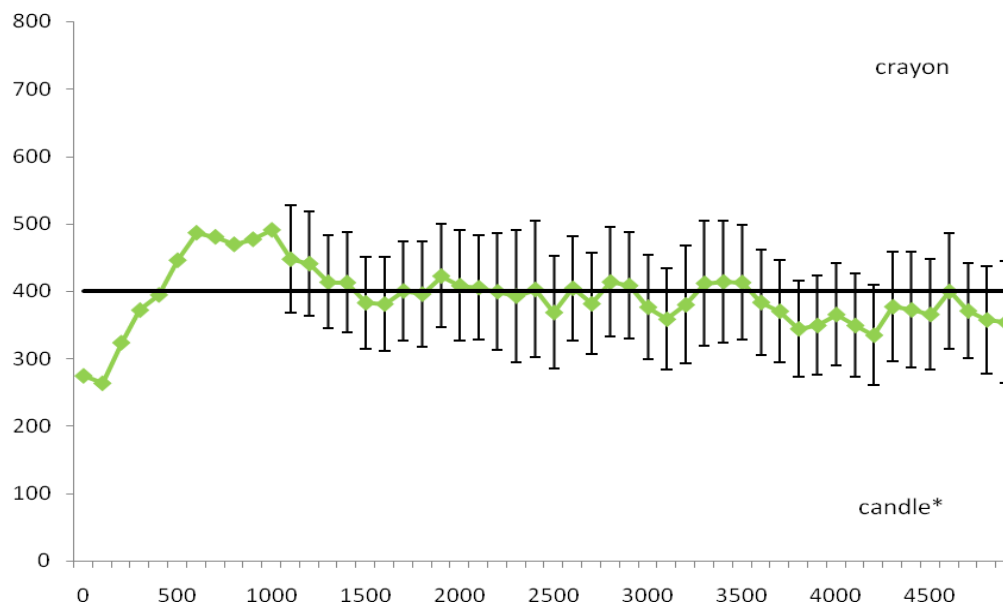


Figure 42: Eye movements of adults per 100 ms of the critical moment in the second part of the FB-TB condition of study 1

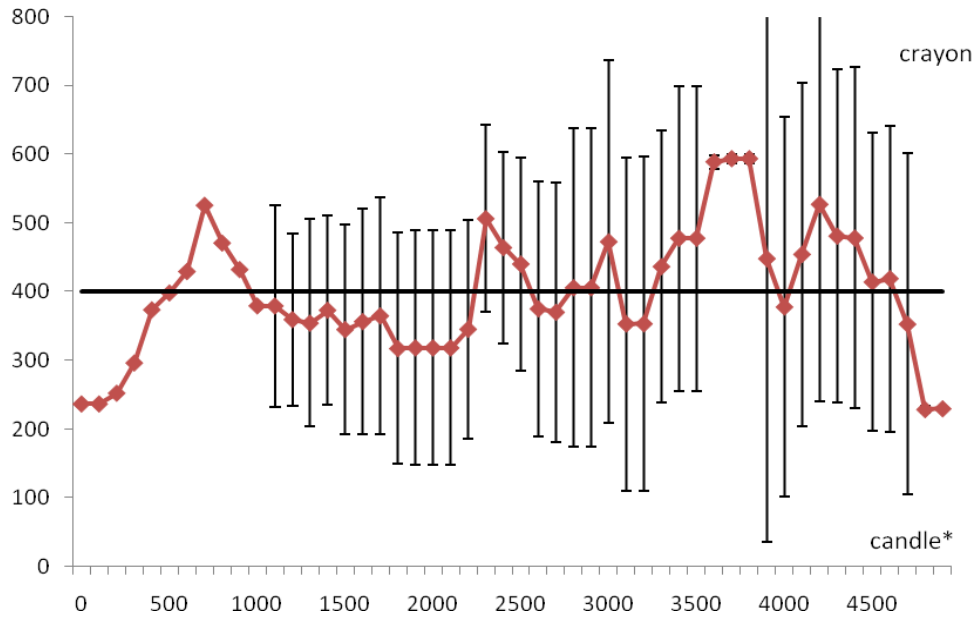


Figure 43: Eye movements of 4-year-olds per 100 ms of the critical moment in the second part of the FB-TB condition of study 1

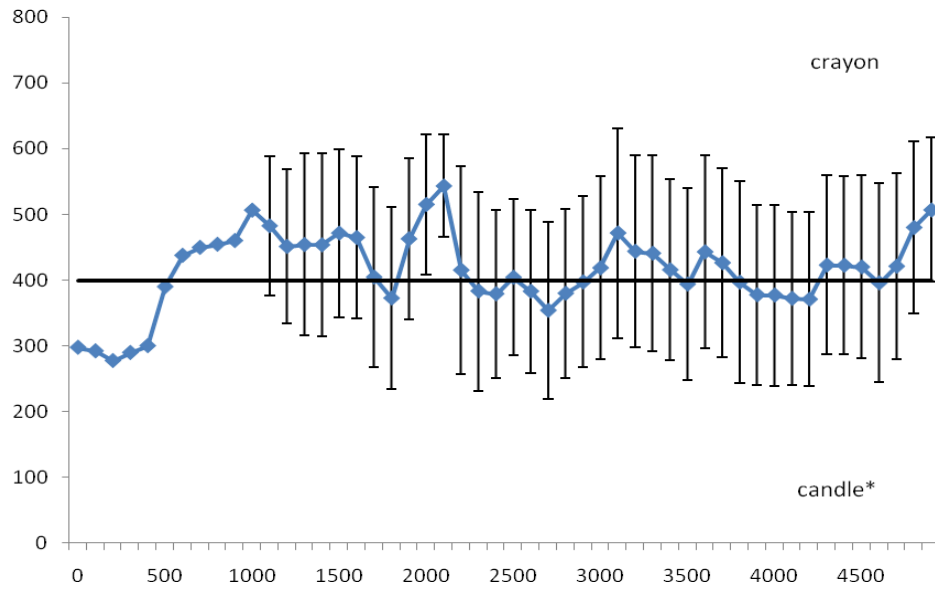


Figure 44: Eye movements of 3-year-olds per 100 ms of the critical moment in the second part of the FB-TB condition of study 1

Relations between performance in the explicit and implicit ToM tasks

The same hypothesis of relations between the traditional ToM performance and the non-verbal ToM reasoning were tested in the unexpected-content task. The eye movement data in Figure 33 were divided in to two groups: trials in which the belief question (“what does she think is in the box?”, the key question asked in the traditional unexpected-content task) was correctly answered versus those that were incorrectly answered. Nineteen out of twenty-five 4-year-olds gave correct answers to this question and six failed to give the correct answer. Although there was no difference in the overall correct-incorrect looking time between the pass group and the fail group $F(1,20)=.73$, $p=.40$, time-based results showed that these 4-year-olds who passed the belief question showed an adult-like pattern of looking time: they looked longer at the correct answer than the incorrect answer at the time 1700-1800ms and after 3700ms (see Figure 45); while these who failed in the belief questions showed a oscillation pattern by switching back and forth between the correct answer and the correct answer (see Figure 46).

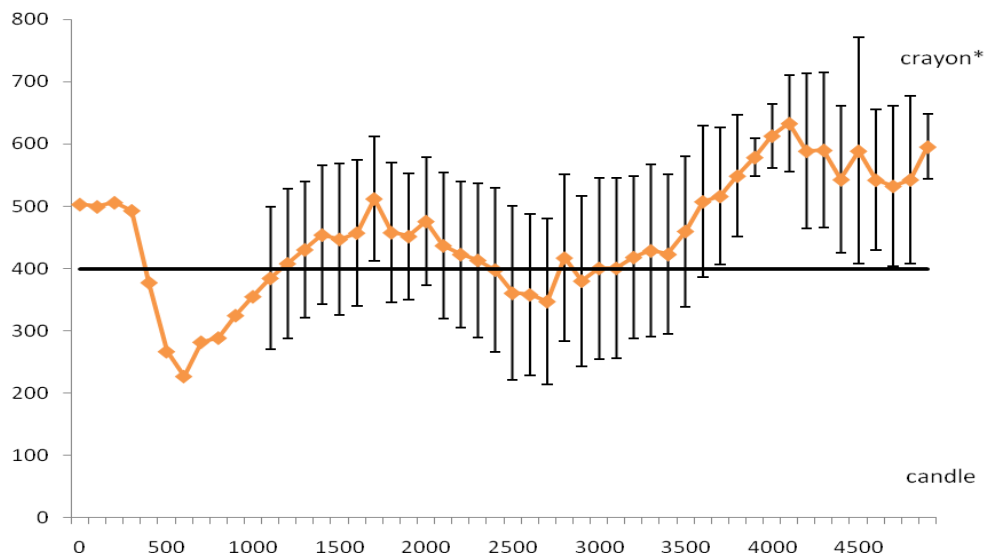


Figure 45: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the belief question

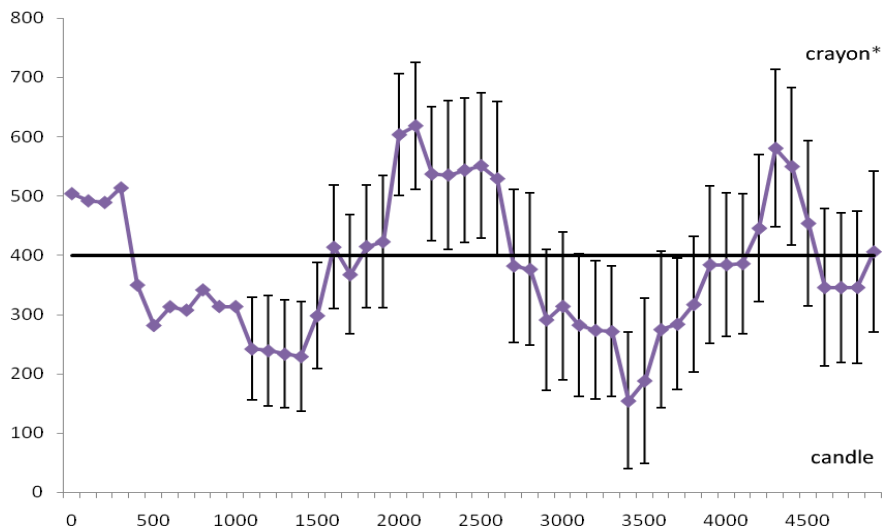


Figure 46: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the belief question

In addition to the item-level analysis, a subject-level analysis was also conducted. Four-year-olds were divided into two groups: the Pass group whose performance of answering all mind-related questions was above average (11 out of 25, average=.65), and Fail group below average (14 out of 25). Results of an ANOVA showed no significant difference between the pass group and the fail group $F(1,20)=3.10, p=.09$. A time-based analysis showed the similar pattern as the dissociation divided by the general ToM performance in the traditional ToM questions.

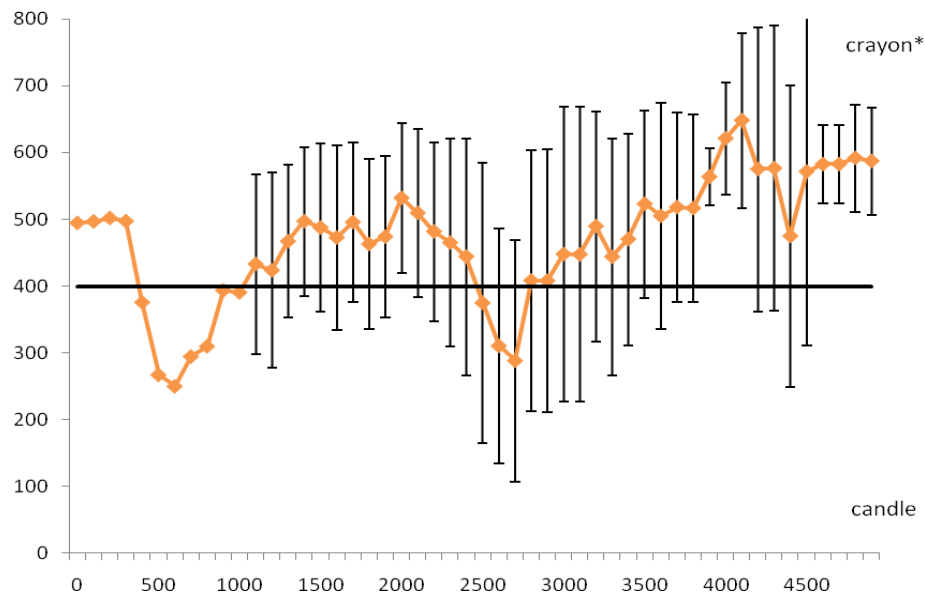


Figure 47: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the overall ToM questions

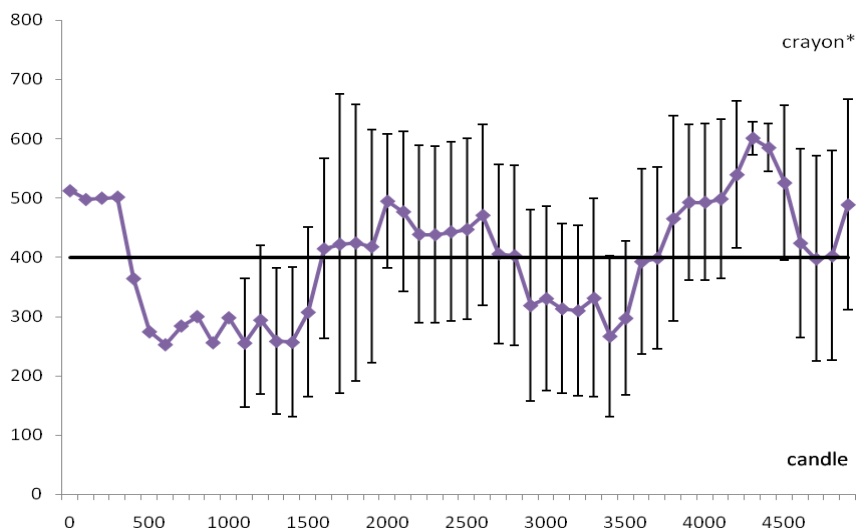


Figure 48: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the overall ToM questions

Then the hypothesis was further tested by separating the graph with general abilities to represent ToM markers in the elicited imitation task. Four-year-olds were divided into two groups: the pass group whose performance in the elicited imitation task was above average (11 out of 25; average=.46), and fail group below average (14 out of 25). Results of an ANOVA showed no difference of correct-incorrect looking pattern between the pass group and the fail group $F(1,17)=2.04, p=.17$. A time-based analysis showed the similar pattern as the dissociation divided by the verbal answers (as shown in the Figure 49 & 50).

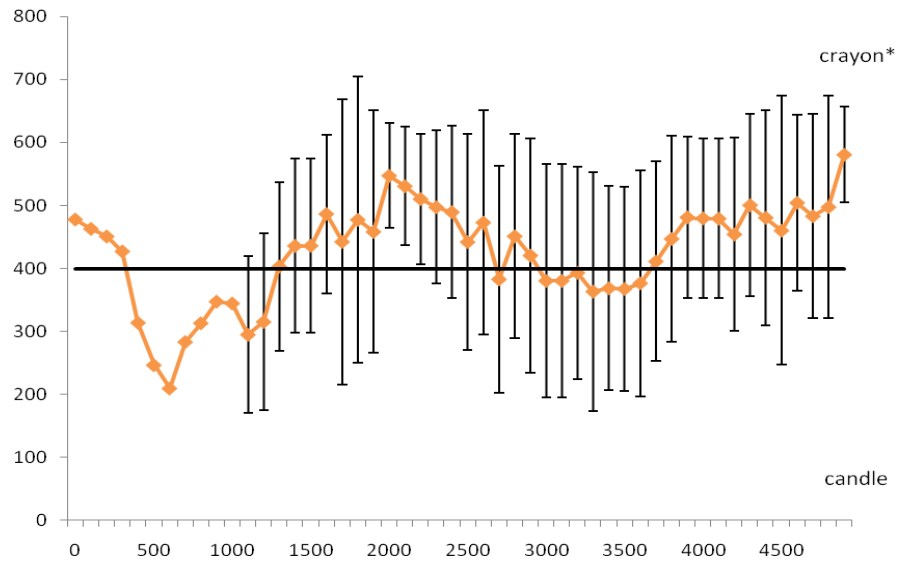


Figure 49: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who passed the elicited imitation task

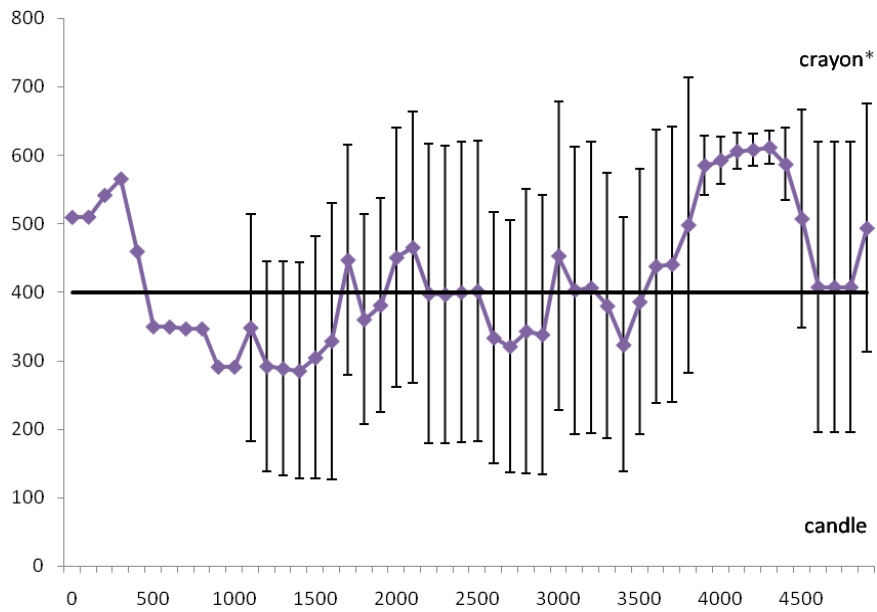


Figure 50: Eye movements per 100 ms of the critical moment in the second part of the FB-FB condition of 4-year-olds who failed the elicited imitation task

Interim discussion for the unexpected-content task

This task provides another case of non-verbal ToM reasoning with implicit measurements. Eye movements of unexpected-content task showed similar anticipatory eye movement patterns as in the change-of-location task: adults, but not 3- and 4-year-old children, showed non-verbal ToM reasoning in the unexpected-content task. The non-verbal ToM reasoning shown in this task is associated with the ToM performance in answering ToM questions. These findings, together with the findings from the nonverbal change-of-location task, imply a consistent pattern of ToM reasoning development across tasks: similar ToM performance was found in both verbal and non-verbal tasks, and in both the change-of-location and the unexpected-content task.

2.3 General Discussion for study 1

Study 1 investigates why children fail in the traditional ToM tasks by hypothesizing that children may have difficulty in two processes. First, an automatic process of ToM reasoning is necessary to encode social information in order to infer others' mental states. Second, linguistic ToM representations are required to correctly interpret traditional ToM questions.

The first process was measured in study 2 using an implicit measurement – the anticipatory eye movements. The anticipatory eye movements were tested in revised

versions of two traditional ToM tasks: the change-of-location and the unexpected-content tasks. Eye movements in both tasks reflect an automatic and real-time process of ToM reasoning in both of the tasks. This process was only found in adults but not in 3- and 4-year-old children, suggesting a similar developmental pattern as the verbal responses in the traditional ToM tasks.

The linguistic representations of ToM, the second process, were investigated with the elicited imitation task, the same as in the study 2. Results confirmed the findings of the study 2: 3-year-olds, compared to older ones and adults, consistently omit ToM markers in the ToM questions. These linguistic representations of ToM were found in this study to predict children's traditional ToM performance, even with their ages controlled. This suggested a unique contribution made by the linguistic representations of the ToM markers to the ToM performance.

Combining the findings with these two processes suggests the sources of children's difficulty in the traditional ToM tasks. In order to pass traditional ToM tasks, children at least need to be able to represent ToM markers linguistically, and to possess an automatic and real-time ToM reasoning. Both of these two processes contribute to the traditional ToM performance independently. Lack of either process will make children fail in the traditional ToM tasks.

3. Study 2: Children's Interpretations of ToM Questions

Study 1 found that young children tend to omit ToM markers in the questions that were beyond their memory load. Study 2 replicated study 1 by using the questions used in previous studies. All questions were chosen and adapted from previous published studies (Clements & Perner, 1994; Gopnik & Astington, 1988; Hogrefe, Wimmer, & Perner, 1986; Sullivan, et al., 1994; Wellman, et al., 2001; Wimmer & Perner, 1983; Zaitchik, 1991). The hypothesis is young children would show similar pattern as in study 1: they will drop the ToM markers in these questions and interpret mind questions as reality questions.

Similar as in study 1, children from age 3 to 6 were asked to repeat verbatim the questions the experimenter asked them. A similar procedure was also used to test adult participants. Adults are very likely to make no error in the elicited imitation task. In order to increase the difficulty level of the task, adult participants were asked to perform a demanding distraction task (e.g. count down by 7 from 100) before they are asked to repeat the questions. By greatly reducing the working memory devoted to remembering the sentences, adult participants are forced to make compromises in what aspect of the linguistic information to encode. I hypothesize that, unlike young children, adults make a clear distinction between factual and belief questions and will therefore recall ToM

markers at the expense of other linguistic information even when they are cognitively overloaded.

3.1 Methods

3.1.1 Participants

Thirty-two 3-year-old (average age=41 months, 15 females, 17 males), twenty-six 4-year-old (average age=54 months; 14 females, 12 males), Eighteen 5-year-old (average age=66 months; 7 females, 11 males), twenty-one 6-year-old and older (average age=80 months; 13 females, 8 males) English-speaking children, and twenty-nine English-speaking adults (average age=19 years; 15 females, 14 males) were recruited from preschools in Durham, NC, and Duke University, respectively. The adults also participated in the studies in studies 1 and 3; 3- and 4-year-olds children also participated in study 1; and 5- and 6-year-olds also participated in study 3.

3.1.2 Material

A set of 9 questions were used in this study listed below. These questions were key questions asked in classic ToM studies. These questions all ask about someone's mental states using ToM markers. Only one question include two markers and the rest 8 questions include only one marker. The results of the questions with one marker and two markers will be analyzed separately. The participants were asked to repeat questions and statements which do not include ToM markers in the training before they

were asked to repeat test questions (see details in the Section 1.2.1). Only the test questions with ToM markers were included in the data analysis. Children and adults were tested with these questions to see if they make errors in any part of the questions, particularly the ToM markers.

The list of test questions in study 2:

Where does he think it is?

When you first saw the house, before we opened it, what did you think was inside it?

Do you remember where Jack put the toy airplane in the beginning?

Why do you think he will open that box first?

Does she know the toy airplane is in the closet?

Does mom know that Jack saw the toy airplane in the bathroom?

Does Jack know that his mom got him a toy airplane for his birthday?

Does Jack know that Mary knows where the ice cream man is now?

Does Jack know that the ice cream man told Mary that he was going to the school?

3.1.3 Procedure

This task was designed to see which part of the questions, if any, is omitted when the length and complexity of the sentence is beyond the memory/cognitive capacity of the participants.

Three- and four-year-old children received the same training as in study 1, asking them to repeat verbatim the questions back but not to answer the questions. The puppet and cover story were not necessary for children over 5 year old and adults. They were asked to repeat verbatim what the experimenter said to them. Adult participants were asked to do a distraction task (e.g. subtraction from 100 by 7) after they heard the questions. After the distraction task they were asked to recall the questions they heard. Participants' answers were recorded verbatim on a recording sheet and also in audio recording tapes.

The recordings of participants' responses were coded in the following categories:

(1) "Invalid": if the participants try to answer the question instead of repeating it, the response will be excluded.

(2) "no error": if the participant can repeated exactly what he/she hears, every word and in the right order.

(3) "refusals": if the participant cannot repeated any single word of the question.

(4) "errors in ToM markers" if the participant omitted or made errors in the ToM markers ("do you think/remember/know"). This is analyzed separately in the questions involving only one ToM marker, and the ones involving two ToM markers.

(5) "other errors": if the participants make other types of errors than in ToM markers.

3.2 Results

All 6-year-olds and adults passed the training and were able to do the elicited imitation task. Twenty-two out of twenty-six 4-year-olds and seventeen out of eighteen 5-year-olds passed the training. Some 3-year-olds had difficulties repeating with the task, as they tried to answer the questions, or simply refused to respond. Only 19 out of 32 3-year-olds passed the training and were able to complete the elicited imitation task. Only the data from these who passed the training will be included in the following analysis.

Participants' performance in the elicited imitation task was coded by two coders independently. The inter-coder reliability between the two coders was 95.94%. There was still a high inter-coder reliability after taking into account the chance levels (Scott's $\pi = 91.48\%$).

Results (see Figure 51) showed that there were age differences of probability of omitting ToM markers in the questions with only one marker $F(3,74)=19.29, p<.0001$. Particularly, 3-year-olds omitted ToM markers more often than 4 year-olds $F(1,74)=7.47, p<.01$; 4-year-olds omitted ToM markers more often than 5-year-olds $F(1,74)=11.87, p<.001$. Participants were more likely to make errors in questions including two ToM markers, although the age difference was not significant, $F(3,60)=0.70, p=0.55$, because only one question involving more than one ToM marker were asked for each participant.

Although there was an overall difference of making other errors across age among the children $F(3,74)=2.73, p=0.0498$, the age differences in individual pair comparisons did not reach significance levels. Particularly, there was no difference of frequencies of making other errors between 3-year-olds and 4-year-olds $F(1,74)=0.01, p=0.93$, 4-year-olds and 5 year-olds $F(1,74)=1.58, p=0.21$, 5-year-olds and 6-year-olds $F(1,74)=1.17, p=0.28$.

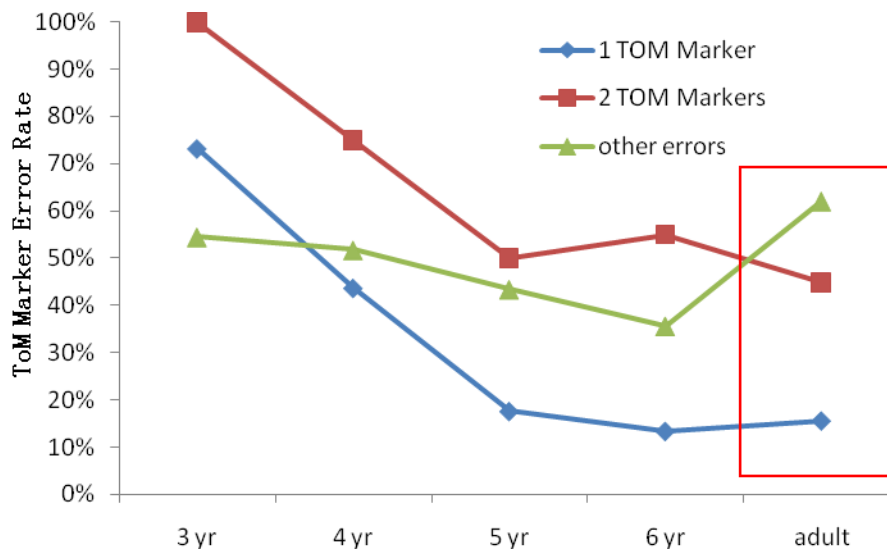


Figure 51: Percentages of making errors in ToM markers and other errors in the elicited imitation task in study 2

3.3 Discussion

Study 2 was designed to address the issue whether children understand the significance of linguistic ToM markers in questions often asked in traditional ToM tasks. This was assessed using an elicited imitation task, where participants were asked to repeat ToM questions under heavy working memory load. It was hypothesized that children will preserve the ToM markers in the question at the expense of other information if they represent the special nature of these markers. On the other hand, children who do not appreciate the importance of ToM markers will likely see them as redundant and therefore drop them in the elicited imitation task.

Results from this study showed that children younger than 4 year old consistently omitted ToM markers when they were asked to repeat the questions. There was a significant improvement from 3- to 4-year-olds, and from 4- to 5-year-olds in representing ToM markers. This developmental change is not driven by a general increase of working memory – while 5- and 6-year-old children make fewer errors overall compared to younger children, there is a dramatic decrease in the relative proportion of errors with ToM markers compared to other types of errors. In particular, 3-year-olds are likely to drop ToM markers more than any other words, whereas 5- and 6-year-olds tend to keep ToM markers and err on other parts of the questions. They may make other errors: for example, they may interpret the question “why do you think he

will open that box first?" as "which one do you think he will open first?". This strongly suggests that by 5 years of age children understand the difference between a factual question and a question about mental states that is marked linguistically, and that they represent the linguistic marking even under working memory overload. It is impressive that the strength of ToM marking representations among the older preschoolers was as strong as adults (under working memory load).

In contrast, 3-year-olds, and to some extent 4-year-olds, do not seem to attach any special significance to linguistic ToM markers in traditional ToM questions in the elicited imitation task. This is unlikely a result of production difficulties, as most ToM markers in the questions are highly familiar words such as "think" or "thought." The fact that 3-year-olds consistently drop these markers strongly suggests that they failed to encode the ToM markers with sufficient strength and significance in the first place. This interpretation is consistent with prior findings on difficulties in acquiring mental verbs (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Johnson & Maratsos, 1977) and other mental state-related linguistic structures.

An important implication of this finding is that 3-year-olds may also fail to represent the ToM markers in questions asked in traditional ToM tasks. Although the elicited imitation task differs in many ways from a ToM task, children are overloaded with information in both cases. Data from this study suggest that under working

memory load, young children tend not to represent the ToM marker in questions about mental states. This, by definition, turns a mind-question into a reality-question. Thus, the fact that 3-year-olds typically fail traditional ToM tasks does not necessarily imply that they do not have the ability to perceive and reason social actions in mental terms. A simple explanation may be that they did not understand the question.

4. Study 3: Understanding of Second-Order Beliefs

The current study will exam whether children linguistically represent ToM markers for second-order beliefs, and whether they attribute the second-order beliefs in a non-verbal setting. Second-order beliefs are represented in English involving two ToM markers (i.e. does A know that B knows...). Similar to the first order belief understanding, children need to acquire this double-embedded structure to correctly interpret the ToM questions about second-order beliefs. The hypothesis is that lack of linguistic representations to the second-order ToM would lead to the failure in the classic second-order ToM questions. In order to test this hypothesis, 5- and 6-year-olds were tested in this study in the elicited imitation task to repeat verbatim the questions with two ToM markers.

Another focus of the current study is whether children attribute second-order belief understanding in their anticipatory eye movements. This will be tested with videos with minimum language input which tell stories involving someone's understanding of another's second-order beliefs.

4.1 Method

4.1.1 Participants

Eighteen 5-year-old (average age=66 months; 7 females, 11 males), twenty-one 6-year-old and older (average age=80 months; 13 females, 8 males) English-speaking

children, and twenty-nine English-speaking adults (average age=19 years; 15 females, 14 males) were recruited from Durham preschools and Duke University. The adults also participated in study 1.

4.1.2 Material

Like in study 1, the participants watched three videos, each involving a different pair of persons play with a different set of toys. These videos are in three conditions, the first-order FB, FB-TB, and second-order belief. They all follow a similar storyline (see Table 2) but vary in a critical aspect, i.e., the knowledge state of the character B. The first-order FB condition is identical to the FB-FB condition in the unexpected-content of study 1. The FB-TB condition is again similar to the FB-TB condition in study 1. In the second-order belief condition is similar to the FB-TB condition; except that in FB-TB condition B is allowed to look into the box while in the second-order belief condition she is not allowed to look. B has to pretend she did not look to get the reward.

At the end of each video, a still shot of two objects was shown for five seconds. During the five seconds participants' anticipatory eye movements were recorded. Then an ending was shown and some probe, first-order and second-order belief questions were asked after showing each video. The descriptions of each scene and questions are listed in the Table 2.

Table 3: Descriptions of each scene of the videos used in study 3

	second-order Belief condition	first-order FB	FB-TB
training	A puts a box in the table and ask B not to look into the box while A leaves. If B can help looking into the box she will get a sticker as a reward; if B looks into the box, she will not get the sticker.	same as second-order belief condition	no training
	While A leaves B looked into the box.	same as second-order belief condition	no training
	When A comes back and asks B if she looked in the box. B admitted, and A does not give B the sticker.	same as second-order belief condition	no training
main story	A, alone in the scene, shows a crayon box, takes out the crayon in the box, and puts a candle in it.	same as second-order belief condition, but replace crayon with chocolate, and candle with marble	same as second-order belief condition, but replace crayon with cookie, and candle with plate
	A asked B the content of the box, and B says “crayons”. Then A leaves the box on the table and asks B not to look in the box, otherwise she will not get the sticker.	same as second-order belief condition	same as second-order belief condition, but A allows B to look in the box, and B will get a sticker anyway.
	While A leaves, B looks into the box and discovers the candle.	While A leaves, B does not look in the box.	same as second-order belief condition

Table 4 (continued)

	A returns and asks if B looked into the box. B denies. A repeats the rule (“if you looked in the box you will not get a sticker”) and asks B to point to one of two objects (crayon and candle) to indicate what is in the box.	same as second-order belief condition	A returns and asks if B looked into the box. B admits. A then asks B to point to one of two objects to indicate what is in the box.
anticipatory looking scene	scene of two objects for 5 s	scene of two objects for 5 s	scene of two objects for 5 s
prediction	"Which one will she point to?"	"Which one will she point to?"	"Which one will she point to?"
pointing	B points to the crayon	B points to the chocolate	B points to the plate
verbal questions (answer and then repeat the sentence)	Justification question: "Why did she point to the crayon?"	Justification question: "Why did she point to the chocolate?"	Justification question: "Why did she point to the plate?"
	Evaluation question: "Will she get a sticker?"	same as second-order belief condition	same as second-order belief condition
	Probe question: "What is in the box?"	same as second-order belief condition	same as second-order belief condition
	First-order belief question 1: "What does she think is in the box?"	same as second-order belief condition	same as second-order belief condition

Table 5 (continued)

	First-order belief question 2: "What did she think was in the box before she opened it?"	N/A	same as second-order belief condition
	First-order belief question 3: "Does A know that B looked into the box?"	First-order belief question 3: "Does A know that B did not look into the box?"	same as second-order belief condition
	Second-order belief question: "Does A know that B knows it is candle in the box?"	N/A	N/A

4.1.3 Procedure

Participants watched all three videos. Participants' eye movements while watching the videos were recorded by a remote eye tracker, the Tobii X50. The eye tracker was connected to a 17" monitor with 800X600 resolutions. Stimuli were presented on this monitor via a computer running Tobii's Clearview software.

Before watching each video, participants were familiarized with the toys which they would see in the video. There were two orders of watching the videos and these two orders are counterbalanced in the participants. Half of the participants watched the first-order FB condition video first and then FB-TB video, and the second-order belief condition video; the other half watched the FB-TB video first and then first-order FB

video, and the second-order belief condition video. After each video, questions listed in the Table two were asked in order. Participants were asked to repeat verbatim the questions (elicited imitation) after answering the questions.

4.2 Results

4.2.1 Verbal answers to ToM questions

Figures 52 and 53 show the correct rates of answering the prediction question and the belief questions in all ages and in all three conditions. Figure 52 shows the correct rates of answering the prediction question (“which one will she point to?”). 5- and 6-year-olds showed a very similar pattern: they were as good as adults in the FB-TB condition, but made more errors than adults in the 1-st order FB condition *5-year-olds*: $F(1,65)=8.77, p<.0005$; *6-year-olds*: $F(1,65)=62.43, p<.0005$ and the second-order belief conditions *5-year-olds*: $F(1,65)=61.61, p<.0005$; *6-year-olds*: $F(1,65)=15.90, p<.0005$.

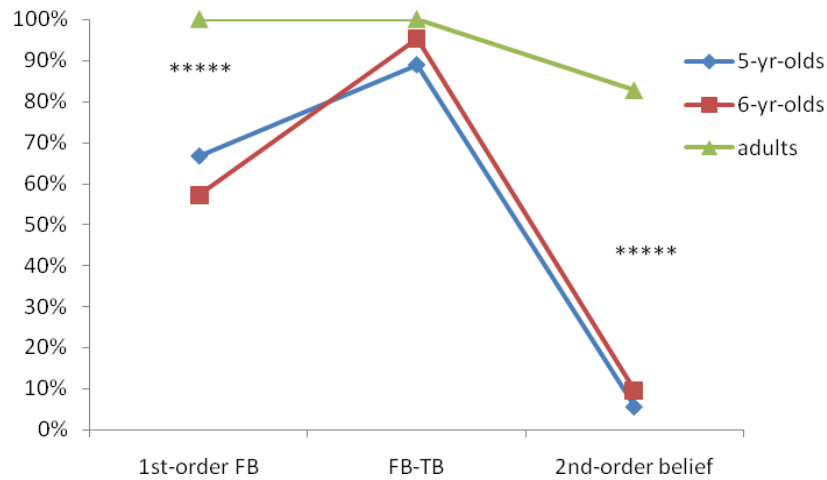


Figure 52: Correct rates of answering the prediction questions in the three conditions of study 3

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; **** * $p < .0001$

Figure 53 shows the correct rates of answering the questions with one ToM marker (“Does A know that B looked into the box?”) in all three condition, and two ToM markers in the second-order belief condition (“Does A know that B knows there is a candle in the box?”). Five- and six-year-olds made more errors than adults when answering this question with one ToM marker in the first-order FB condition *5-year-olds*: $F(1,60)=3.69, p=.59$; *6-year-olds*: $F(1,60)=8.79, p<.005$; in the FB-TB condition, 5-year-olds made more errors than adults when answering this questions with one ToM marker $F(1,58) = 8.78, p<.005$; in the second-order-belief condition, 5-year-olds made more errors than six-year-olds when answering this questions with one ToM marker $F(1,64) = 10.88,$

$p < .005$ and adults $F(1,64)=12.41, p < .001$. Five- and six-year-olds made more errors than adults when answering the ToM questions with two ToM markers *5-year-olds*:

$F(1,63)=12.63, p < .001$; *6-year-olds*: $F(1,60)=24.88, p < .0001$.

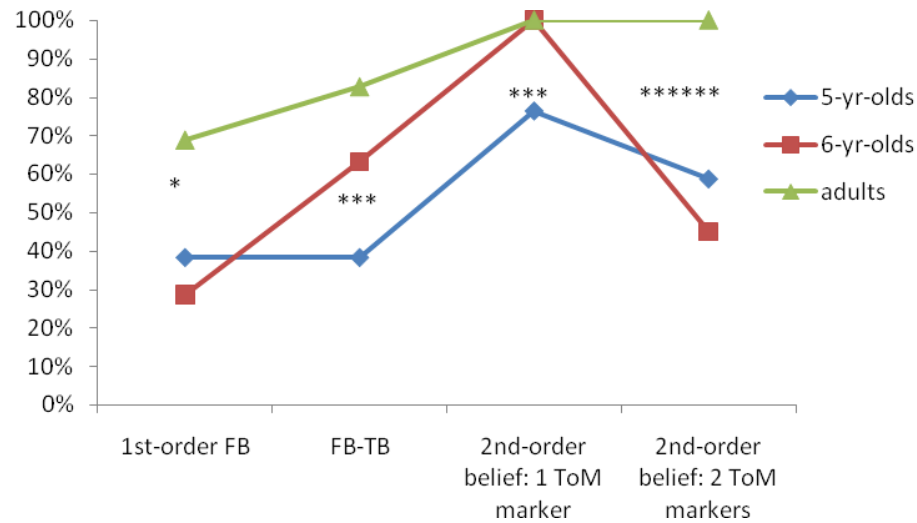


Figure 53: Correct rates of answering the questions with one or two ToM markers in the three conditions

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; **** * $p < .0001$

4.2.2 Elicited imitation

Participants' performance in the elicited imitation task was coded in the same way as in the previous study. The reliability between the two coders was high *inter-coder reliability*=99.17%; *Scott's Pi*=96.42%.

As shown in Figure 54, although adults did not make any errors, there were differences of frequencies of omitting the second-order ToM markers between 5-year-olds and 6-year-olds: 5-year-olds omitted the second-order ToM markers approximately 67% of time and 6-year-olds 29% of time $F(1,36)=5.76, p<.05$. However, there was not such a significant difference of omitting first-order ToM markers and making other errors: 5-year-olds omitted the first-order ToM markers 27% of time and 6-year-olds 16% of time $F(1,37)=3.22, p=0.08$; 5-year-olds made other errors 37% of time and 6-year-olds 26% $F(1,37)=0.29, p=0.59$.

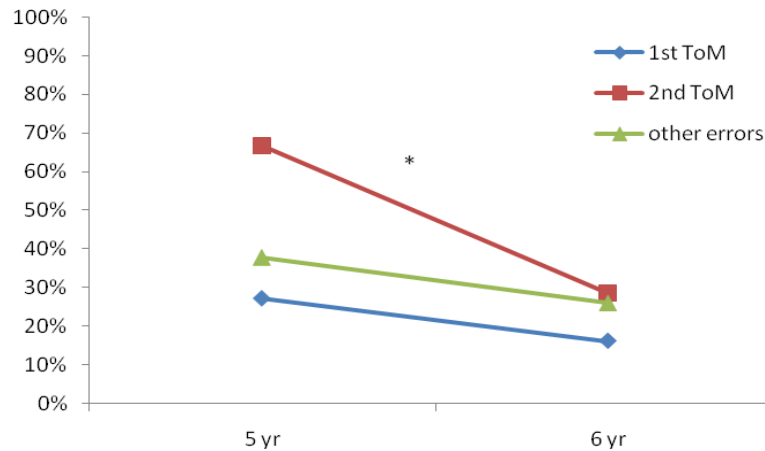


Figure 54: percentages of making errors in ToM markers and other errors in the elicited imitation task of study 3

The relation between children's verbal answers to ToM questions and elicited imitation was examined at the item level. A chi-square test for 5- and 6-year-old children showed that there is no correlation between children's errors in the first order ToM marker in the elicited imitation and their verbal answers $\chi^2=0.23, p=0.62$, also no correlation between children's errors in the second order ToM marker in the elicited imitation and their verbal answers $\chi^2=0.10, p=0.74$, and no correlation between children's other errors in the elicited imitation and their verbal answers $\chi^2=0.19, p=0.66$.

4.2.3 Eye Movements

Participants' anticipatory eye movements during the critical event were analyzed. As in study 1, three analyses were conducted – the total looking time to the left vs. the right side the screen, the mean x points across the critical moment, mean x points at every 100 ms.

Figure 55 showed the total looking time to the left and right in the first-order FB condition. A paired t-test showed that adults showed no predictions of pointing to the correct or the incorrect answers $t(25)=-.92, p=.37$; 5-year-olds $t(17)=-.73, p=.48$ and 6-year-olds did not show any predictions either ($t(18)=-1.00, p=.32$).

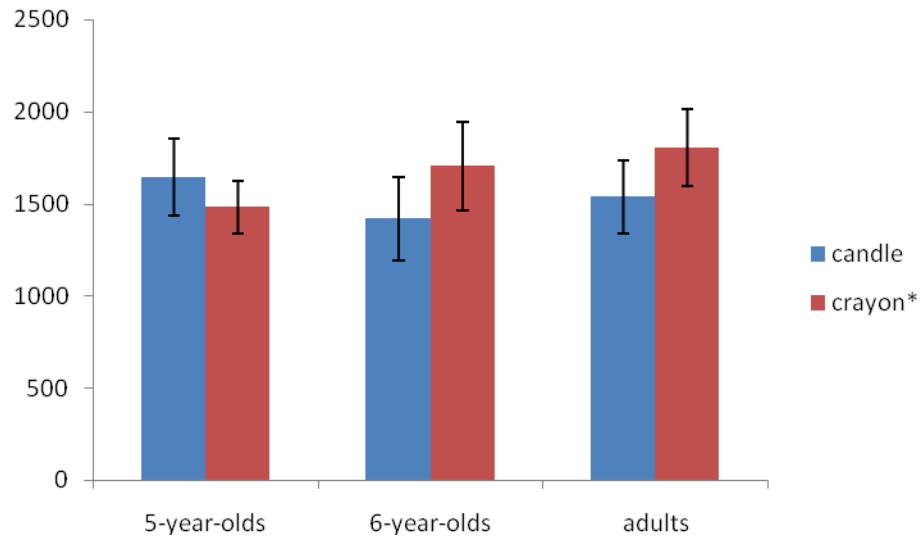


Figure 55: Mean looking time at the crayon and candle in the first-order FB trial of study 3

The average x position during the critical moment was shown in the Figure 56. Results showed that adults $mean=382.59$, $sd=84.32$, $t(24)=-1.03$, $p=.31$, 5-year-olds $mean=388.43$, $sd=74.89$, $t(16)=-.64$, $p=.53$, and 6-year-olds $mean=368.56$, $sd=414.87$, $t(18)=-1.43$, $p=.17$ showed no predictions for pointing to the correct answer or the incorrect one.

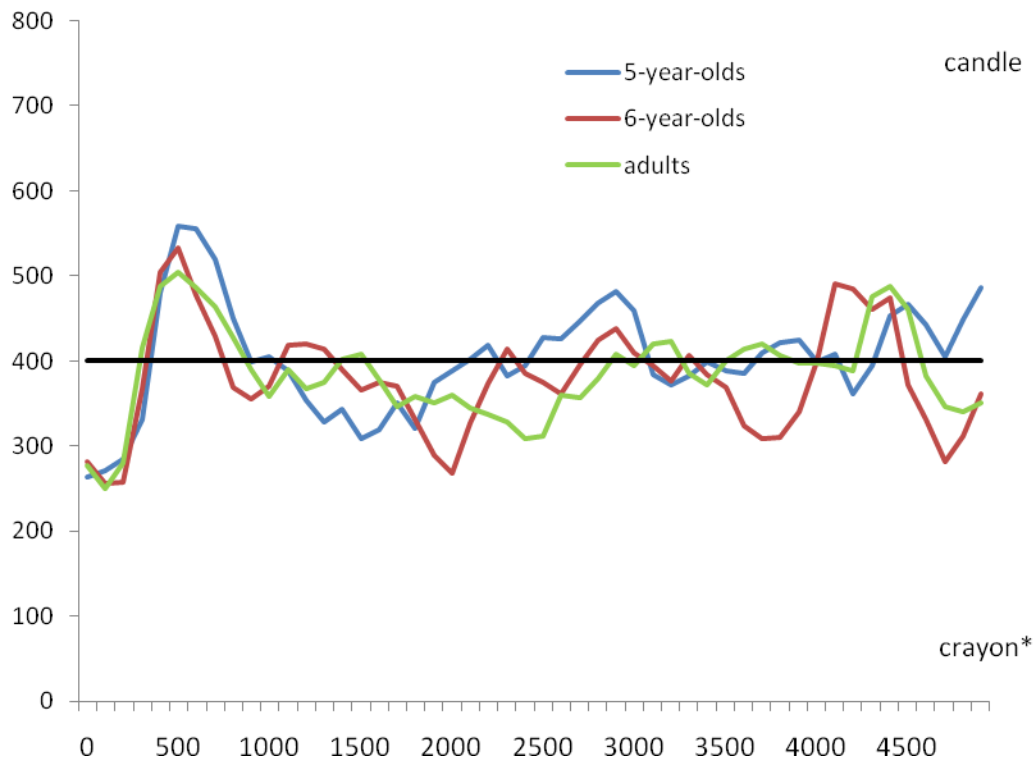


Figure 56: Eye movements of all age groups across time of the critical moment in the first-order FB condition of study 3

The tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 57-59). The results showed that (a) adults looked at the correct answer more than the incorrect answer during 2400-2600ms; (b) 6-year-olds looked at the correct answer during 1900-2100ms, and then went back to the middle for a while, then looked at the correct answer again at 4800ms; (c) 5-year-olds looked at the correct answer at 1500ms.

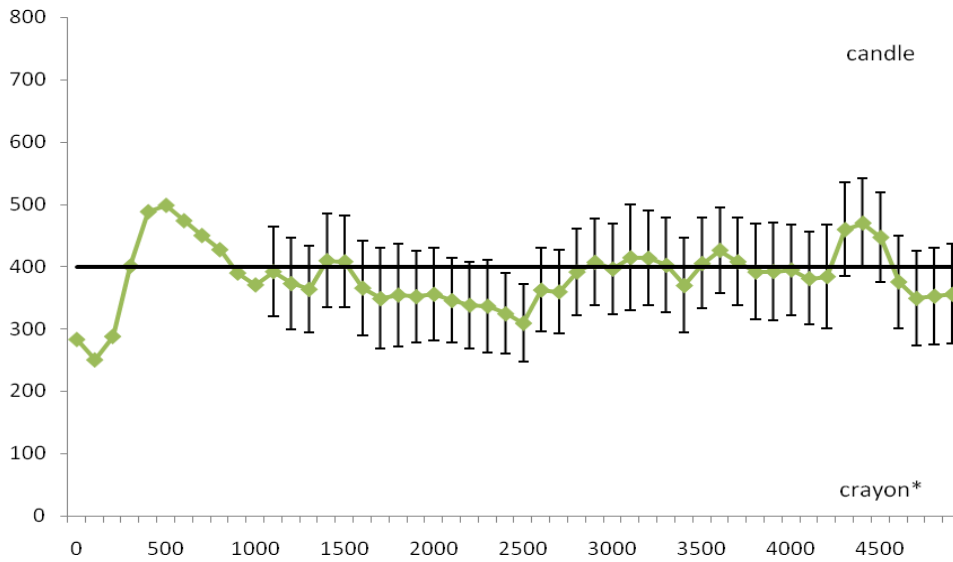


Figure 57: Eye movements of adults per 100 ms of the critical moment in the first-order FB condition of study 3

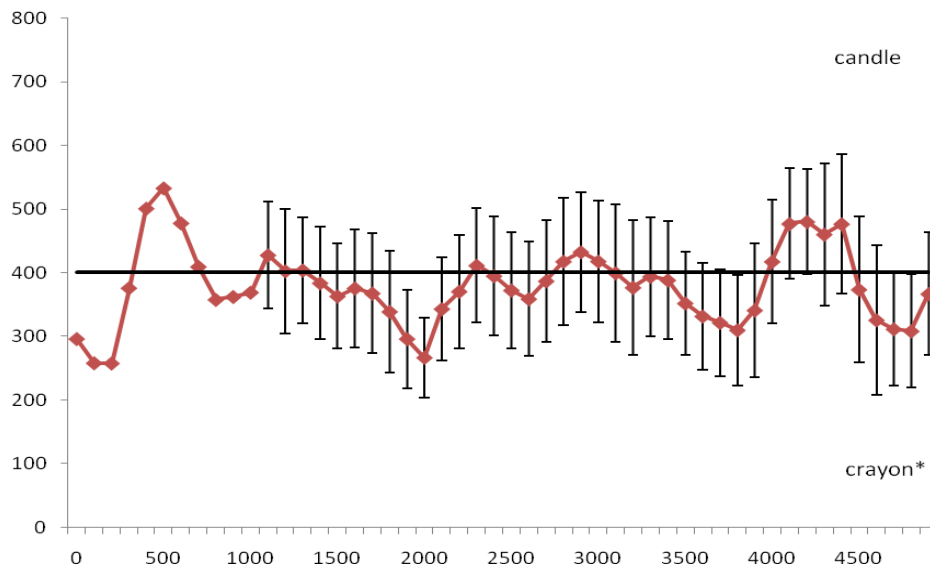


Figure 58: Eye movements of 6-year-olds per 100 ms of the critical moment in the first-order FB condition of study 3

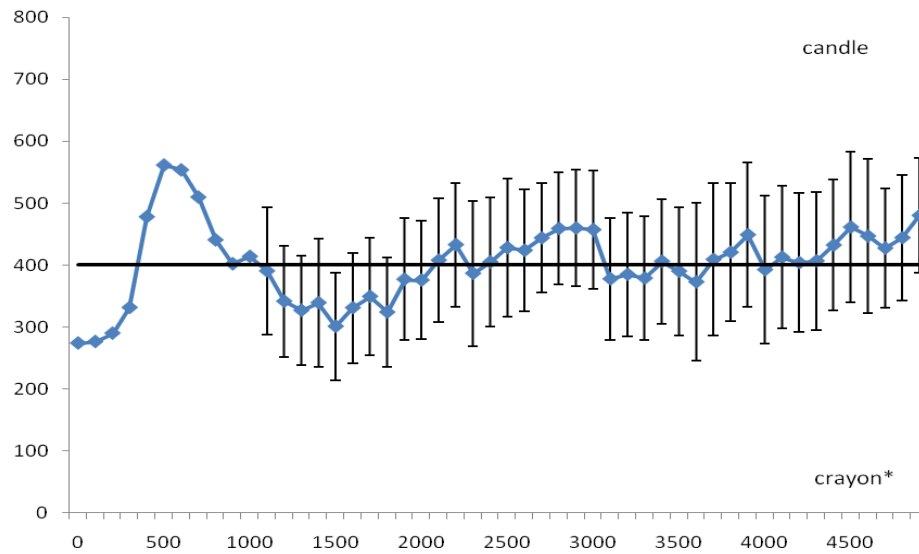


Figure 59: Eye movements of 5-year-olds per 100 ms of the critical moment in the first-order FB condition of study 3

Figure 60 shows the total looking time to the left and right in the FB-TB condition.

Results of a paired t-test showed that adults showed no predictions of pointing to the correct answer or the incorrect one $t(25)=-1.40, p=.17$; 5-year-olds $t(16)=.14, p=.89$ and 6-year-olds also showed no predictions $t(20)=.65, p=.52$.

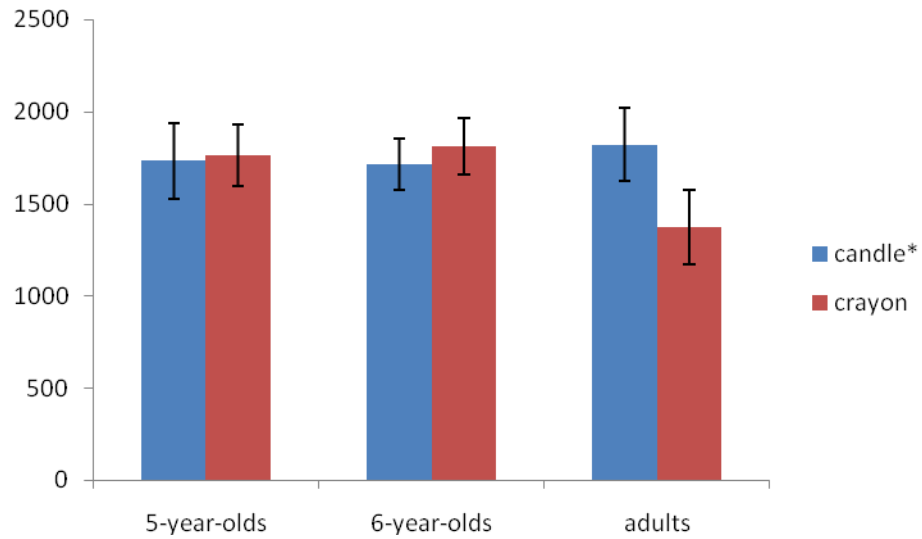


Figure 60: Mean looking time at the crayon and candle in the FB-TB trial of study 3

The average x position in the critical moment was shown in the Figure 61.

Results show that adults looked at the correct answer more than the incorrect answer

*mean=348.14, sd=111.85; $t(25)=-2.36, p<.05$; 5-year-olds *mean=368.46, sd=105.82, $t(16)=-1.23,$**

p=.24 and 6-year-olds *mean=403.41, sd=62.12, $t(16)=.25, p=.80$* did not show any predictions.

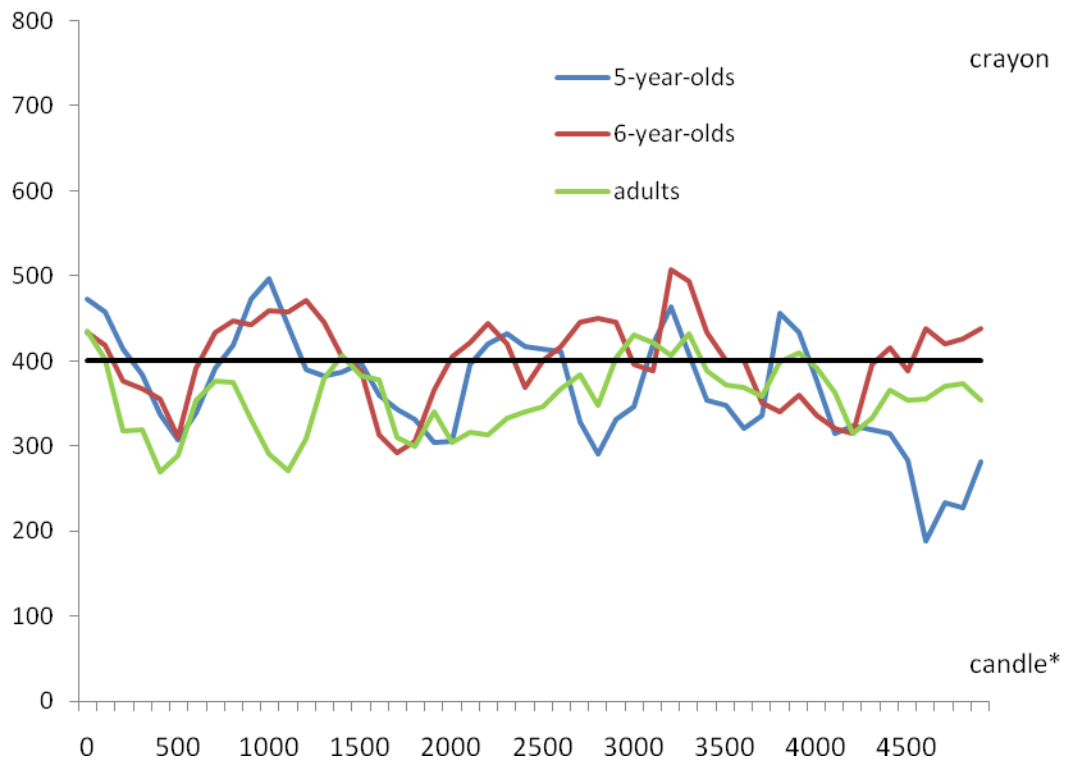


Figure 61: Eye movements of all age groups across time of the critical moment in the first-order FB trial

The tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 62-64). The results showed that (a) adults looked at the correct answer during 1000-1200ms, and 1800-2200ms; (b) 6-year-olds looked at the correct answer at 1700ms, and 4200ms; (c) 5-year-olds looked at the correct answer at 1900ms, and after 4500ms.

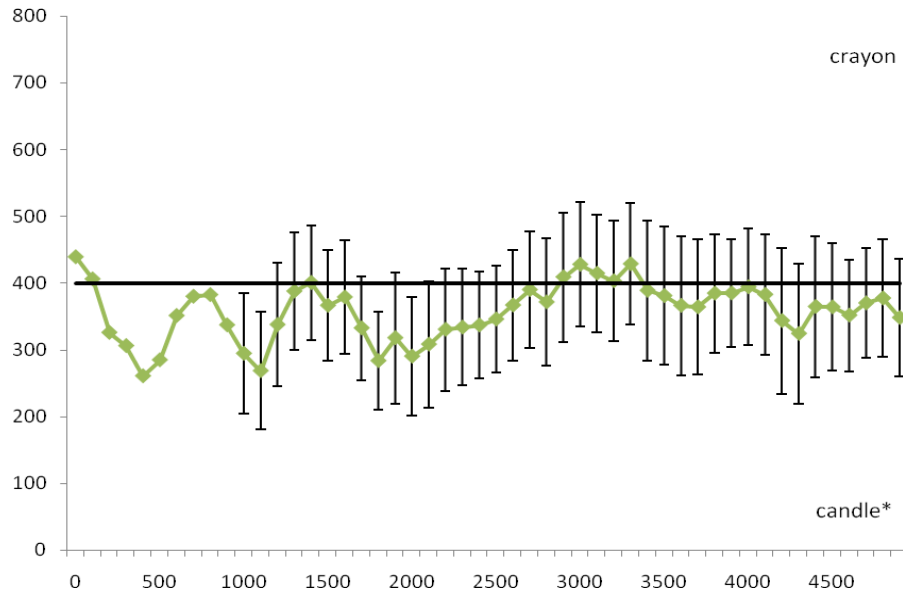


Figure 62: Eye movements of adults per 100 ms of the critical moment in the FB-TB FB condition of study 3

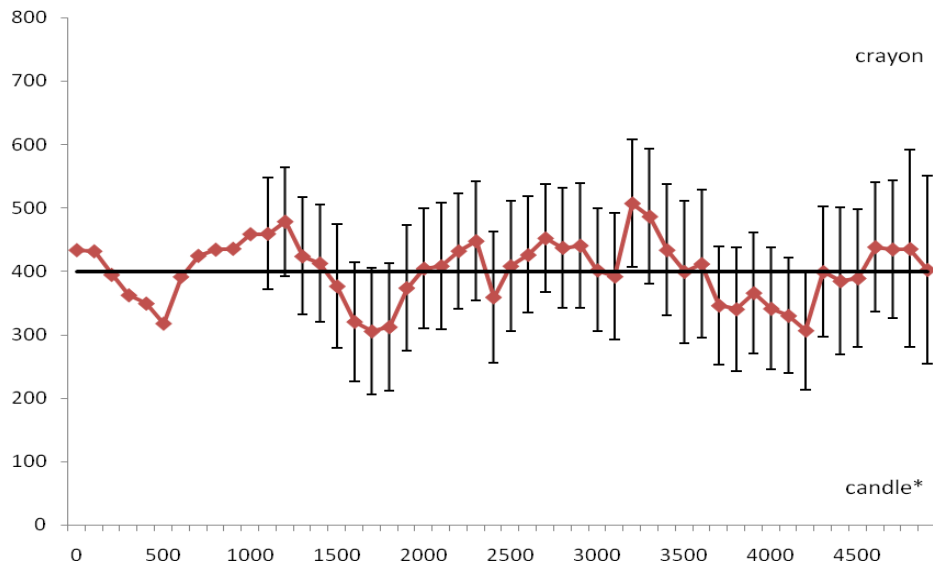


Figure 63: Eye movements of 6-year-olds per 100 ms of the critical moment in the FB-TB FB condition of study 3

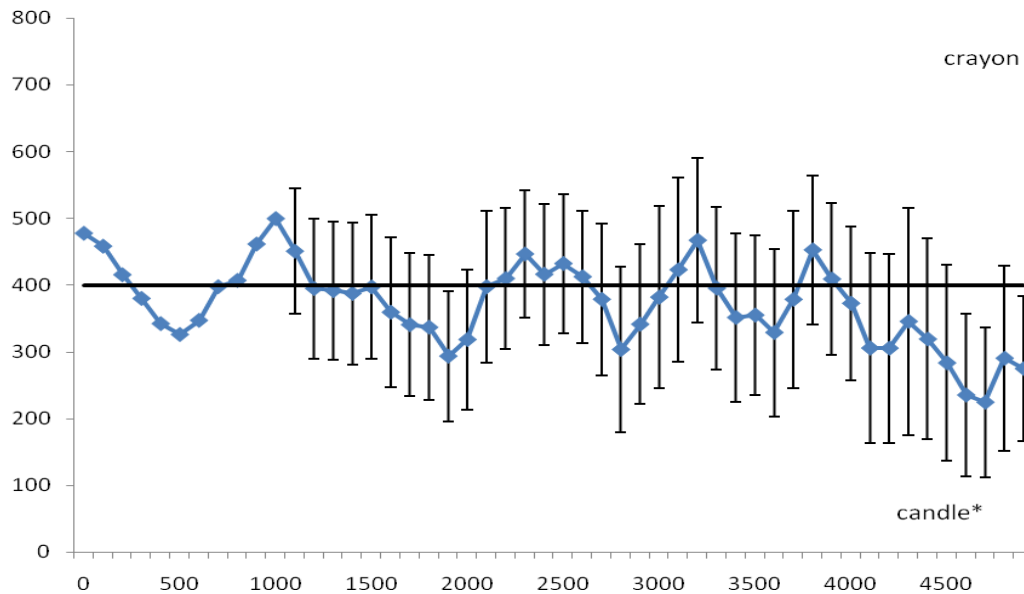


Figure 64: Eye movements of 5-year-olds per 100 ms of the critical moment in the FB-TB condition of study 3

Figure 65 shows the total looking time to the left and right in the second-order false-belief condition. Results of a paired t-test showed that adults looked at the correct answer significantly longer than the incorrect answer $t(24)=3.56, p<.005$; 5-year-olds showed no predictions of pointing to the correct answer or the incorrect answer $t(17)=-1.65, p=.12$; neither did 6-year-olds $t(18)=-.89, p=.38$.

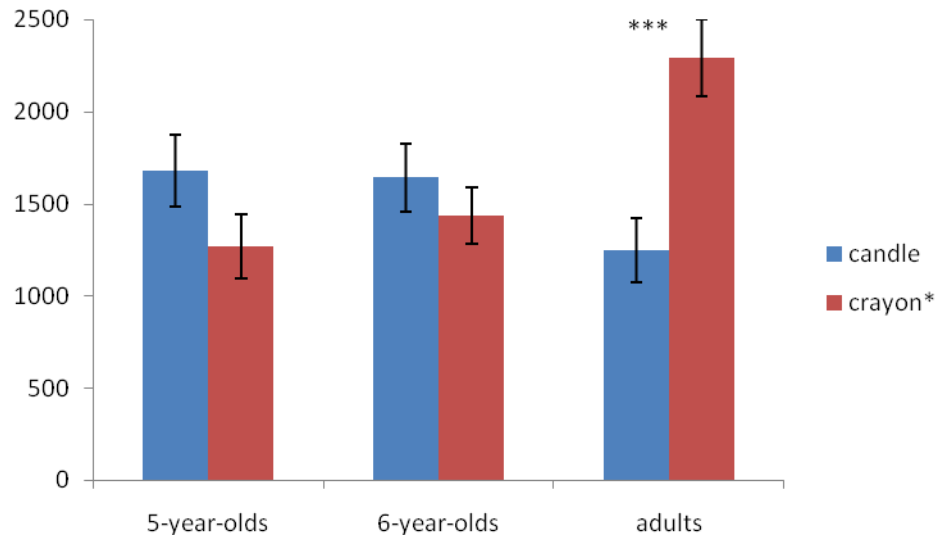


Figure 65: Mean looking time at the crayon and candle in the second-order belief trial of study 3

Notes: *** $p < .005$

The average x positions in the critical moment were shown in the Figure 66.

Results showed that adults looked at the correct answer more than the incorrect answer $mean=478.52$, $sd=91.95$; $t(24)=-4.27$, $p<.0005$; 5-year-olds $mean=376.29$, $sd=100.15$; $t(17)=-1.00$, $p=.32$ and 6-year-olds $mean=389.01$, $sd=70.48$; $t(18)=-.68$, $p=.50$ did not predict the pointing to the correct answer or the incorrect answer.

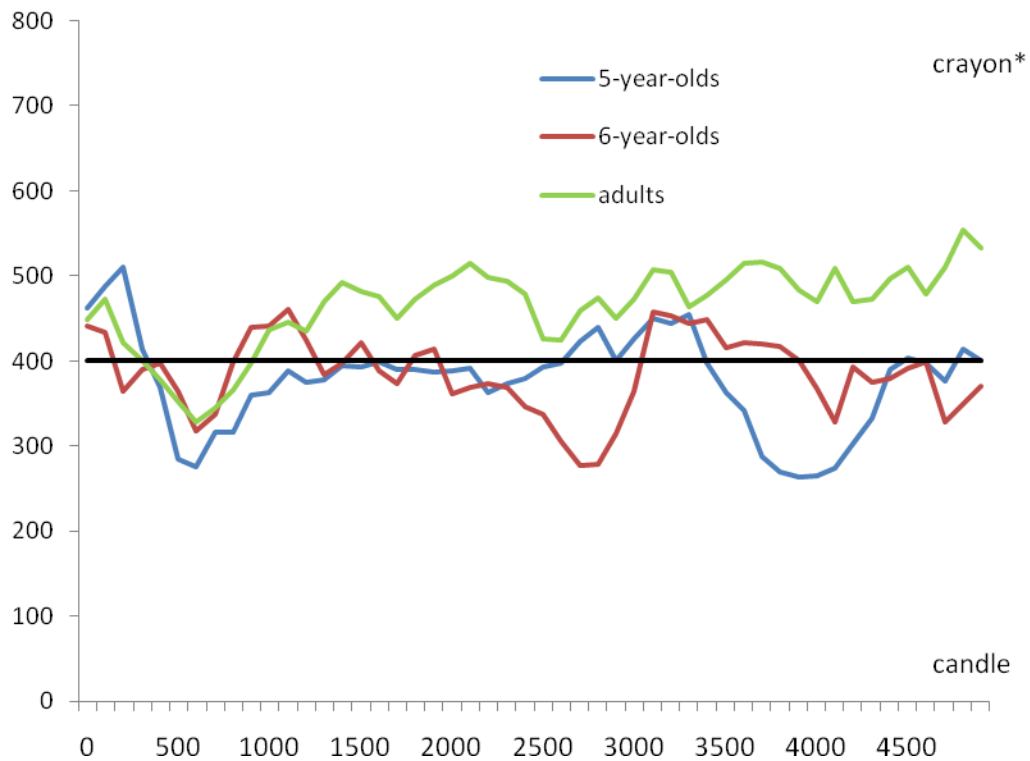


Figure 66: Eye movements of all age groups across time of the critical moment in the second-order belief condition of study 3

T-tests were also conducted at each time point (per 100ms) excluding the first second (see Figure 67-69). The results showed that (a) adults looked at the correct answer more than the incorrect answer during 1400-1600ms, 1800-2500ms, and after 3500ms; (b) 6-year-olds showed no predictions of pointing to the correct answer or the incorrect answer; (c) 5-year-olds did not show any predictions before 3800ms, and then looked at the incorrect answer more than the correct answer during 3800-4200ms.

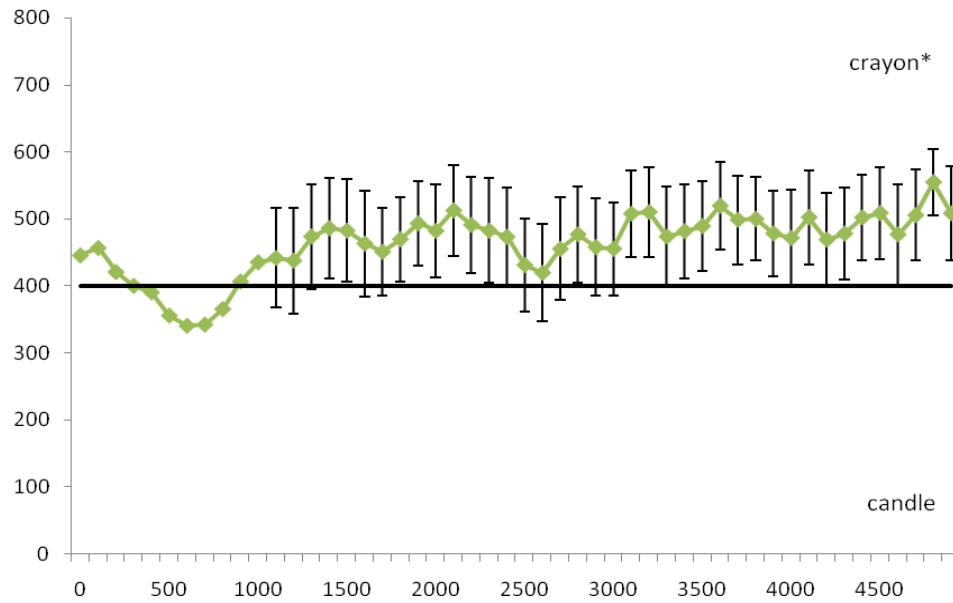


Figure 67: Eye movements of adults per 100 ms of the critical moment in the second-order belief condition of study 3

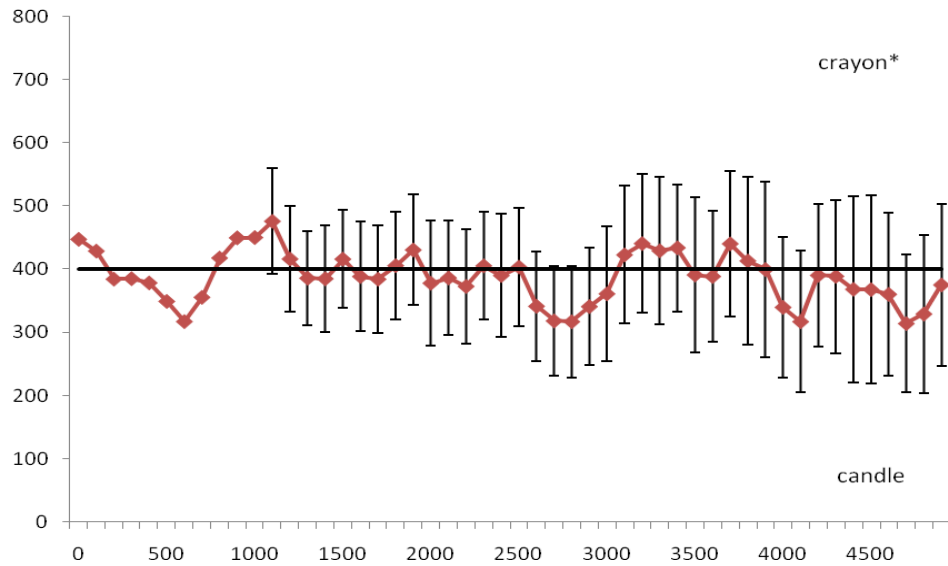


Figure 68: Eye movements of 6-year-olds per 100 ms of the critical moment in the second-order belief condition of study 3

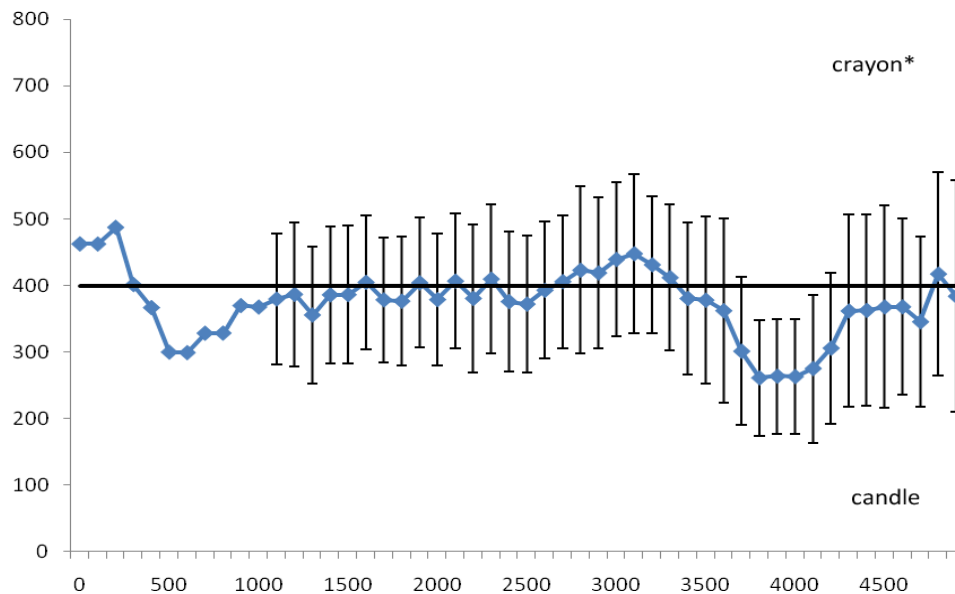


Figure 69: Eye movements of 5-year-olds per 100 ms of the critical moment in the second-order belief condition of study 3

4.3 Discussion

Similar to the study 1, this current study was designed to investigate the source of children’s difficulty in the classic second-order belief understanding tasks. It was hypothesized that there are two potential processes that may be related to children’s failure in the second-order belief understanding. First, children younger than 6 year old may not represent the linguist markers of second-order belief understanding, so that they misinterpret the questions involving two ToM markers. Second, children may not possess the non-verbal ToM reasoning for second-order beliefs.

The first process was assessed by the elicited imitation tasks which asked 5- and 6-year-olds children to repeat verbatim ToM questions involving two markers. Results showed that 5-year-olds, compared to 6-year-olds, consistently omitted at least one ToM markers in the ToM questions with two markers; 5- and 6-year-olds did not differ in representing ToM markers in the questions with only one marker, and in making other types of errors. This finding implies that 5-year-olds may misinterpret the questions with two ToM markers, either as the questions about the first-order beliefs (omitting one of the two markers), or about the reality (omitting both two markers).

In their verbal responses, both 5- and 6-year-olds failed to give correct answers to either the prediction questions or the belief questions involving two markers. Also, the current data did not show any correlations between the performance in the elicited imitation and the classic second-order ToM tasks. This may be due to the limited number of questions involving two markers (only one per participant). The correlations between linguistic representations of the second-order ToM and the classic second-order ToM performance could be tested in future research.

The second process was tested with the implicit measurement, the anticipatory eye movements, which were recorded while participants watched the videos involving second-order ToM. Results showed that the anticipatory eye movements in adults, but

not in 5- and 6-year-olds, showed an automatic and real-time process of others' second-order ToM.

In summary, the current study found that 5-year-olds have difficulty representing the linguistic markers of second-order ToM, while 6-year-olds and adults can linguistically represent second-order ToM markers. Adults, but not children, show an automatic process of second-order ToM reasoning.

5. Study 4: Chinese Children's Interpretations of ToM Questions and Implicit ToM

A meta-analysis conducted by Wellman et al (2001) showed that there is a universal developmental trend of ToM, although children of the same age from different countries may perform differently. In order to examine whether the current findings would be expanded to another culture, this study will be a replication of study 2 using Chinese counterparts.

The same videos were presented and same questions in Chinese were asked to Chinese participants. Chinese language uses similar ToM markers as English. Therefore, similar findings in the Chinese children would be expected.

5.1 Methods

5.1.1 Participants

Sixteen 3-year-old (average age=44 months, 7 females, 9 males), twenty-four 4-to 5-year-old and older (average age=62 months; 10 females, 14 males) Chinese-speaking children, and thirty Chinese-speaking adults (average age=22 years; 22 females, 8 males) participated in the elicited imitation, the change-of-location and the unexpected-content task. Chinese-speaking children were recruited from The May First kindergarten in Beijing, China. Adult participants were recruited in Beijing Normal University in Beijing China.

5.1.2 Material and Procedure

Children and adults participated in the elicited imitation task, the change-of-location task and the unexpected-content task. The same material as study 2 were used in Chinese language and conducted by a native Chinese-speaking experimenter, with the following exceptions: (1) Chinese children were asked to verbatim repeat all questions including those used in the videos first, and then watch the video. At the end of each video, they were asked to answer these questions only (in English studies, they were asked to answer and then repeat the questions); (2) In the change-of-location task, the FB complete condition was presented before the FB non-complete condition; (3) in the unexpected-content task, the order for the FB and the FB-TB conditions were counterbalanced, so that half of the participants watched the FB condition first and the other half watched the FB-TB condition first. This would counterbalanced the orders so that performance in one condition would be not consistently different than the other. However, in this case there was no familiarization trial, which may influence the performance in the first-exposed videos. (4) Since there is no tense marking in Chinese, the belief question in present tense “what does she think is in the box” was clarified by adding a tense reference “after she opened the box”.

Chinese ToM markers “jue2 de” (think) and “zhi1 dao4” (know) were used in the questions in the current study. These two ToM marker are neutral markers which do not imply the false belief situation as “yi3 wei2” and “dang4” in Lee et al (1999)’s study.

5.2 Results

5.2.1 Elicited imitations

Chinese participants’ responses in the elicited imitation task were coded by two coders separately. The inter-coder reliability between the two coders was 97.47%, even after taking chance into account *Scott’s Pi* =94.66%.

As shown in Figure 70, there were age differences in the probability of omitting ToM markers in the questions involving only one ToM marker $F(2,60)=14.39, p<.0001$. Particularly, 3-year-olds omitted the ToM markers more often than 5-year-olds $F(1,60)=4.32, p<.05$ and adults $F(1,60)=23.17, p<.0001$; 5-year-olds omitted ToM markers more often than adults $F(1,60)=9.27, p<.005$. There was no significant difference of making other errors across age between 3- and 5-year-olds $F(1,34)=2.76, p=0.11$.

Results showed that there were also age differences in the probability of omitting ToM markers in the questions involving two markers $F(2,49)=14.95, p<.0001$. Particularly, 3-year-olds omitted ToM markers more often than adults $F(1,49)=11.70, p<.005$; 5-year-olds omitted ToM markers more often than adults $F(1,49)=25.69, p<.0001$; 3-year-olds did

not differ from 5-year-olds in the probability making errors on ToM markers in these questions $F(1,49)=.04, p=.84$.

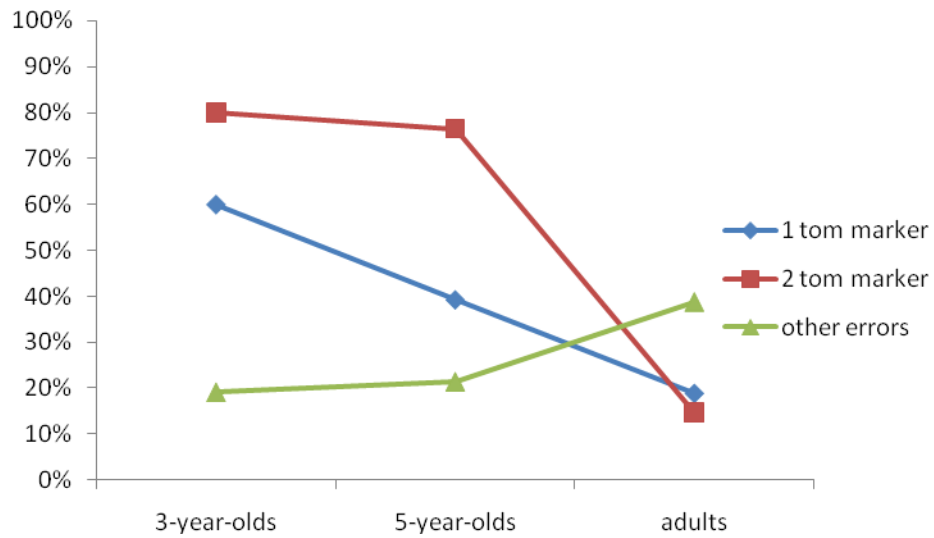


Figure 70: Percentages of making errors in ToM markers and other errors in the elicited imitation task in study 4

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; ***** $p<.0001$

5.2.2 Verbal answers to ToM questions

Figure 71 shows the correct rates of answers to ToM questions in the change-of-location task. Results showed that (1) 3- and 5-year-olds made more errors than adults in answering prediction questions (“Where will Lisa go to look for the car?”) in false belief trials *age 3*: $F(1,58)=32.77, p<.0001$; *age 5*: $F(1,58)=33.51, p<.0001$; (2) 3-year-olds made more

errors than adults in answering prediction questions (“Where should Lisa go to look for the car?”) in false belief trials $F(1,63)=6.83, p<.05$; (3) 3-year-olds made more errors than 5-year-olds $F(1,59)=4.39, p<.05$ and adults $F(1,59)=24.75, p<.0001$ in answering the belief question “where did Lisa think the car was before she put on the mask?”; 3- and 5-year-olds made more errors than adults *age 3*: $F(1,58)=42.72, p<.0001$; *age 5*: $F(1,58)=39.54, p<.0001$ when answering the belief question “where did Lisa think the car was when she opened her eyes?”

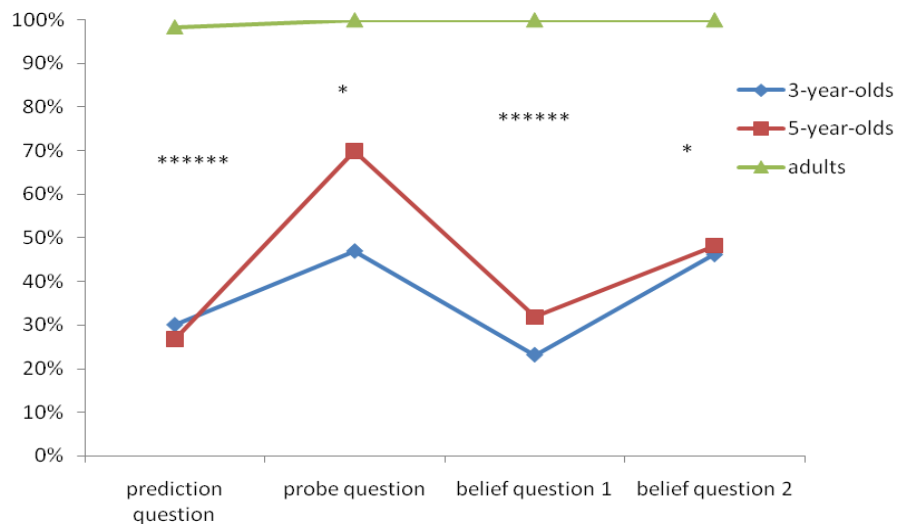


Figure 71: Correct rates of all ages answering different types of questions in the change-of-location task of study 4

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; * * * * * $p<.0001$

In the unexpected-content task, as shown in Figure 72, results showed that (1) 3-year-olds made more errors than 5-year-olds $F(1,67)=7.48, p<.01$ and adults $F(1,67)=26.64, p<.0001$ in answering the prediction question “which one will she point to?”; (2) 3-year-olds made more errors than 5-year-olds $F(1,67)=18.36 p<.0001$ and adults $F(1,67)=70.35, p<.0001$ when answering the probe question “what is in the box?”; (3) 3-year-olds made more errors than 5-year-olds $F(1,64)=15.56, p<.0005$ and adults $F(1,64)=37.07, p<.0001$ in answering the prediction question “what does she think is in the box after she opened the box?”; (4) 3-year-olds made more errors than 5-year-olds $F(1,65)=14.81 p<.0005$ and adults $F(1,65)=54.19, p<.0001$ when answering the prediction question “what does she think is in the box before she opened the box?”

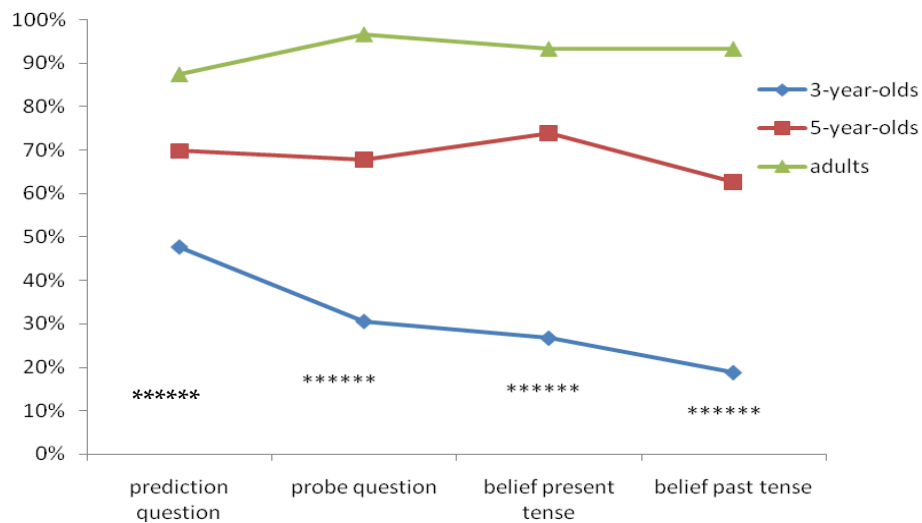


Figure 72: Correct rates of all ages answering different types of questions in the unexpected-content task of study 4

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

5.2.3 Eye movements in the Change-of-location task

True Belief Condition

Total looking time at the left or right exits were compared across subjects using a paired t-test. Figure 73 shows that adults looked at the correct prediction longer than the incorrect prediction $t(27)=5.30, p < .0001$; 5-year-olds also looked at the correct prediction longer than the incorrect prediction $t(20)=2.61, p < .05$; 3-year-olds did not show any predictions $t(13)=1.37, p = .19$.

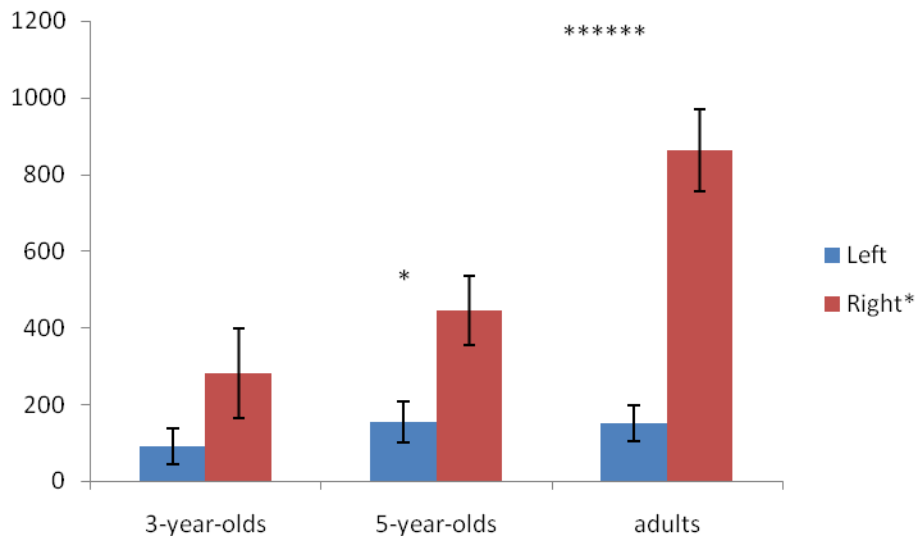


Figure 73: Mean looking time to the left and right in the TB trial of study 4

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

Figure 74 shows the anticipatory eye movements in the true belief trial of the change-of-location task. T-tests were also conducted to test in the critical moment excluding the first second. Results of t-tests showed that adults and 5-year-olds ages looked longer at the correct prediction *adults: mean=472.66 sd=92.12, t(28)=5.83, p<.0001; age 5: mean=471.33, sd=118.53, t(22)=3.89, p<.001*; 3-year-olds showed no predictions *mean=418.05, sd=128.38, t(14)=1.36, p=.20*. This is due to the shift of the average x position to the incorrect direction near the end of the critical moment.

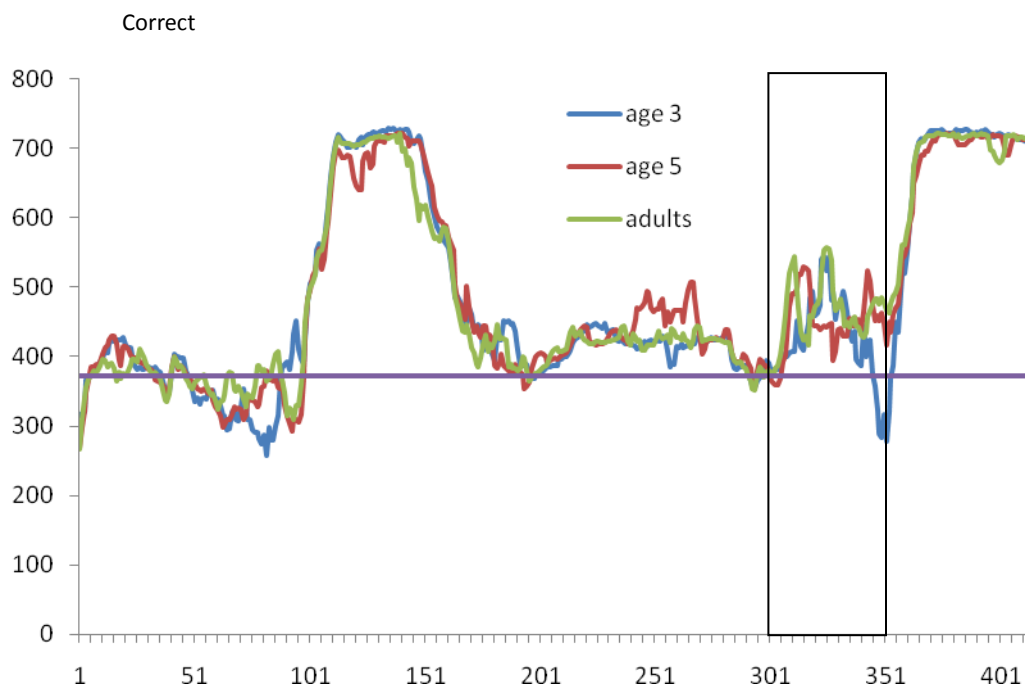


Figure 74: Eye movements of all age groups across time in the true belief trial of study 4

False Belief Complete Condition (first trial of false belief condition)

In the first trial of the false belief condition, the car moves from right to left. A correct prediction for Lisa's movement should be the **right** side of the screen. Total looking time at the left or right exits were compared across subjects using a paired t-test. Figure 75 shows that adults looked at the correct prediction more than the incorrect prediction $t(27)=3.47, p<.005$; 3-year-olds and 5-year-olds showed no predictions *age 3*: $t(12)=-1.72, p=.11$; *age 5*: $t(21)=-.38, p=.71$.

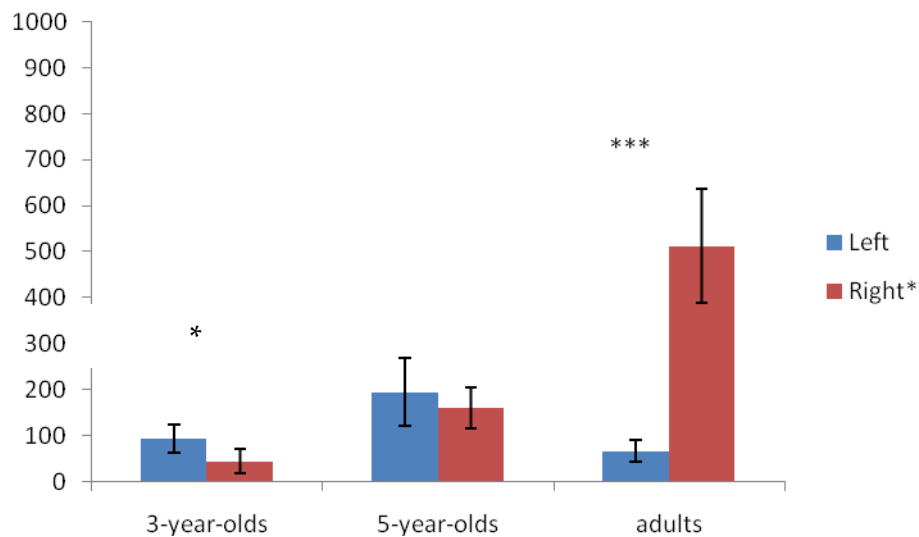


Figure 75: Mean looking time to the left and right in the second FB trial of study 4

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

Eye movements in the third trial across time were shown in the Figure 76. T-test showed that (a) adults looked at the correct prediction more than the incorrect prediction $mean=475.63.14$; $sd=90.27$, $t(27)=6.12$, $p < .0001$; (b) 3- and 5-year-olds showed no predictions to the left or right side of the screen *age 3*: $mean=321.69$, $sd=130.11$, $t(14)=-1.53$, $p=.15$; *age 5*: $mean=352.61$, $sd=103.87$, $t(21)=-.92$, $p=.37$.

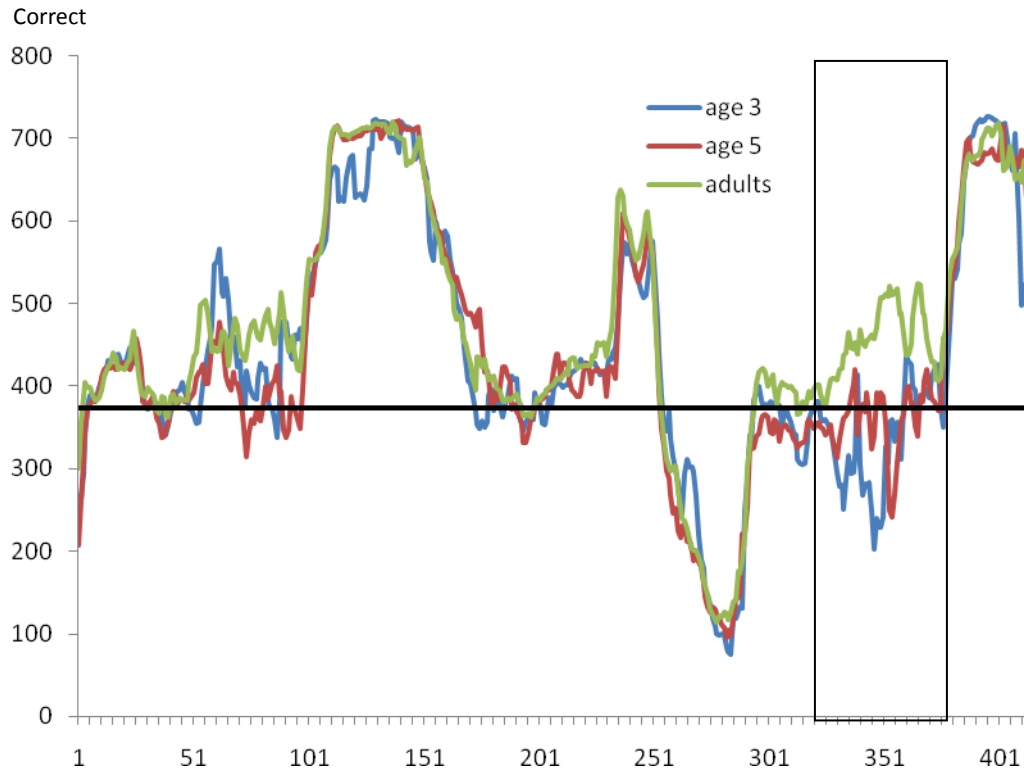


Figure 76: Eye movements of all age groups across time in the false belief complete trial of study 4

False Belief non-complete Condition (second trial of false belief condition)

In the second trial, also the second trial of the false belief condition, the car moves from left to right. A correct prediction for Lisa's movement should be the left side of the screen. Total looking time at the left or right exits were compared across subjects using a paired t-test. Results (Figure 77) showed that adults and 5-year-olds looked at

the correct prediction more than the incorrect prediction *adults*: $t(25)=-3.12, p<.005$; *age 5*: $t(20)=-3.94, p<.001$; 3-year-olds showed no predictions to the left or right paths $t(9)=-.58, p=.57$.

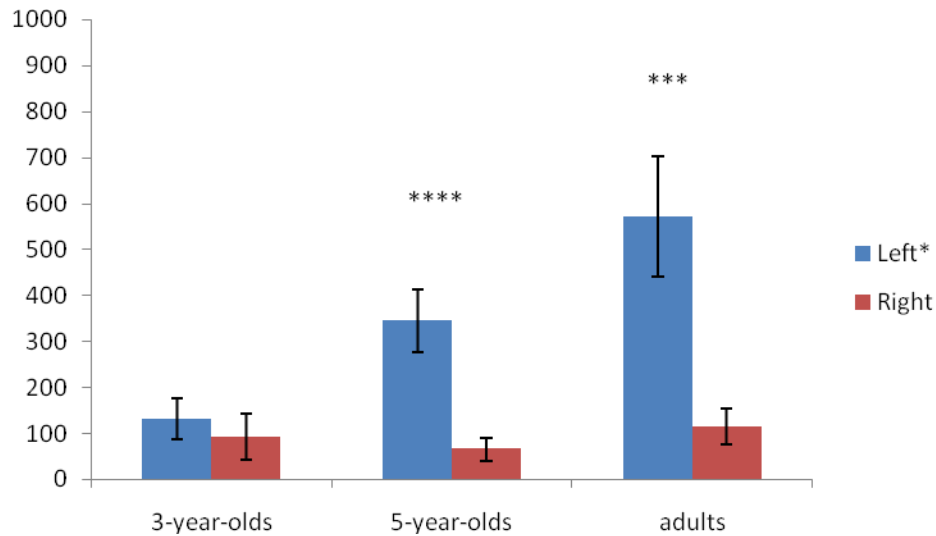


Figure 77: Mean looking time to the left and right in the first FB trial of study 4

*Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; **** * $p<.0001$*

Eye movements across time were shown in the Figure 78. T-test showed adults did not show clear predictions over time $mean=358.26; sd=102.89, t(28)=-.77, p=.45$: 5-year-olds looked at the correct prediction more than the incorrect prediction $mean=321.07; sd=88.28, t(21)=-2.76, p<.05$; 3-year-olds showed no predictions of Lisa's movements $mean=407.76, sd=190.88; t(9)=0.58, p=.58$.

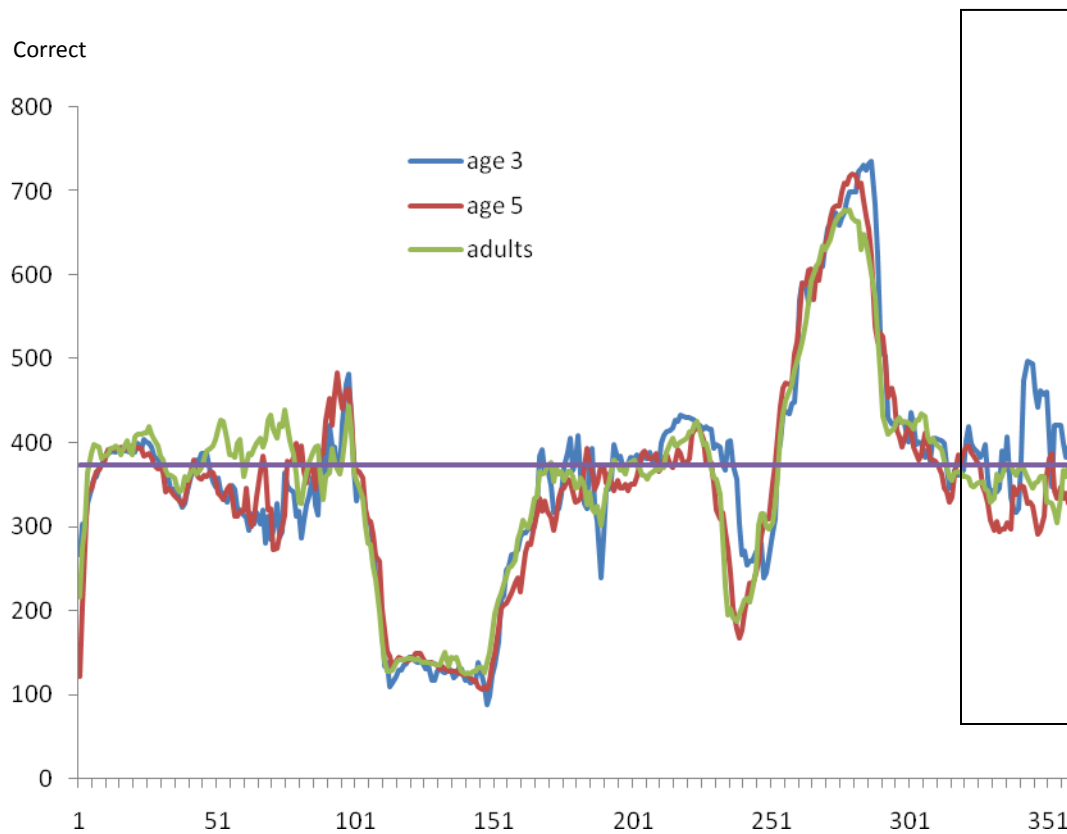


Figure 78: Eye movements of all age groups across time in the false belief non-complete trial of study 4

5.2.4 Eye movements in the unexpected-content task

Participants' anticipatory eye movements while watching the videos of the unexpected-content task were recorded and analyzed. In the FB trial, the total looking time at the left or right sides were compared across subjects using a paired t-test. Figure 79 shows that adults looked at the correct answer on the left significantly more than the

incorrect answer on the right $t(26)=-7.66, p<.0001$; 5-year-olds also looked at the correct answer more than the incorrect answer $t(20)=-3.78, p<.005$; 3-year-olds showed no predictions of the pointing behaviors $t(26)=-.43, p=.68$.

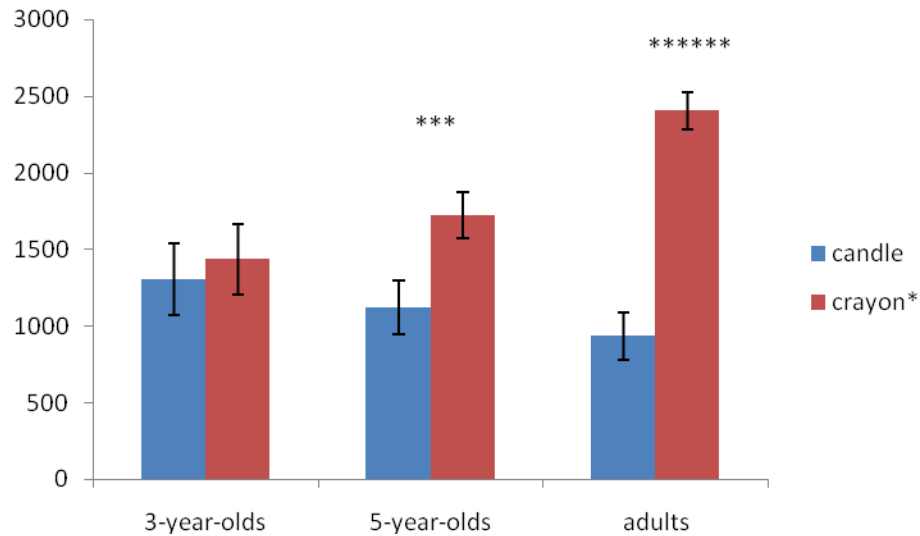


Figure 79: Mean looking time at the crayon and candle in the FB trial of the unexpected-content task of study 4

Notes: * $p<.05$; ** $p<.01$; *** $p<.005$; **** $p<.001$; ***** $p<.0005$; ***** $p<.0001$

The average x position in the critical moment was shown in the Figure 80.

During the critical moment of the second part of the FB-FB condition, adults showed prediction of pointing to the correct answer than the incorrect answer $mean=296.96, sd=53.74, t(26)=-8.96, p<.0001$; 5-year-olds also predicted the pointing to correct answer

$mean=321.78, sd=68.67, t(18)=-4.96, p<.0005$; 3-year-olds showed no predictions of the pointing behaviors $mean=375.16, sd=118.78, t(13)=-.575, p=.47$.

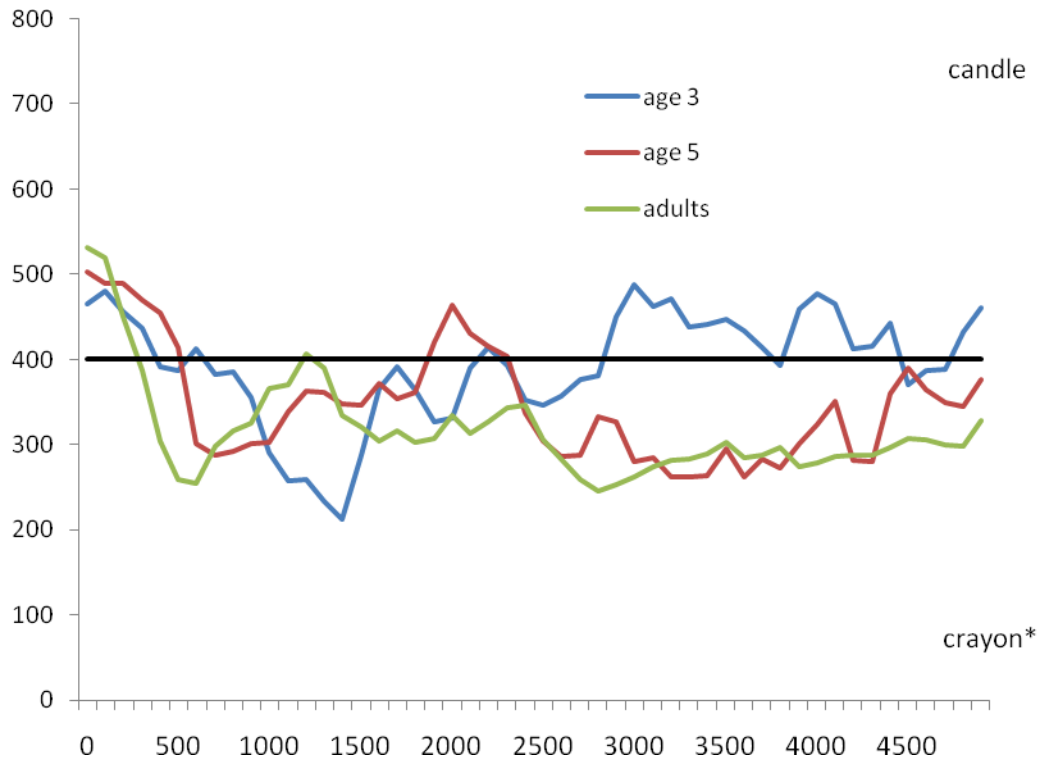


Figure 80: Eye movements of all age groups across time of the critical moment in the FB trial of study 4

In the first part of the FB-TB trial, the total looking time over the correct answer and the incorrect answer were compared across subjects using a paired t-test(as shown in the Figure 81). Results showed that adults $t(26)=-3.43, p<.005$ and 5-year-olds $t(26)=-$

3.15, $p < .005$ looked at the correct answer longer than the incorrect answer ; 3-year-olds showed no predictions of pointing to either of the two objects $t(14) = -.04, p = .97$.

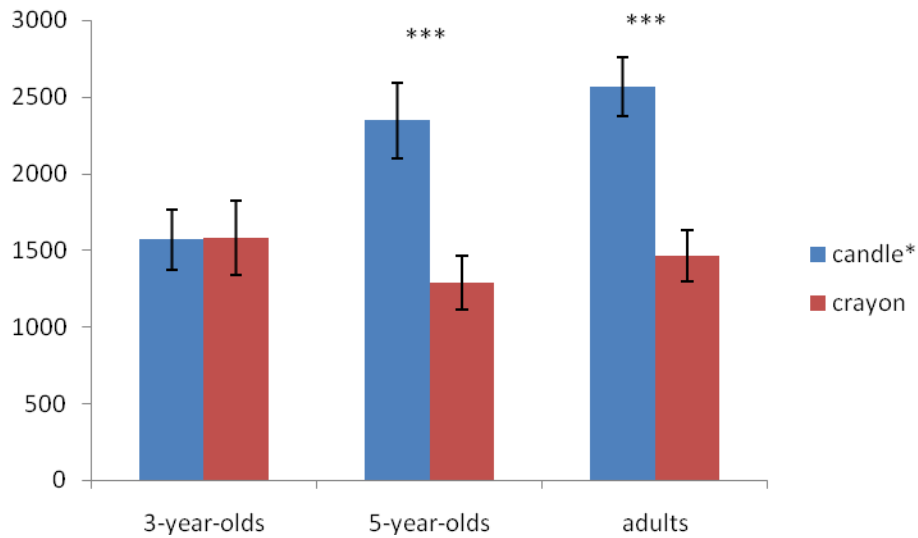


Figure 81: Mean looking time at the crayon and candle in the FB-TB trial of the unexpected-content task of study 4

Notes: * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$; ***** $p < .0005$; ***** $p < .0001$

The average x positions in the critical moment were shown in the Figure 82. Results showed that in the second part of the FB-TB trial, adults $mean = 324.43, sd = 89.69, t(26) = -4.38, p < .0005$ and 5-year-olds $mean = 321.30, sd = 123.97, t(22) = -3.04, p < .01$ looked at the correct answer more than the incorrect answer; 3-year-olds showed no predictions for the pointing to either of these two objects $mean = 379.36, sd = 140.73; t(14) = -.57, p = .58$.

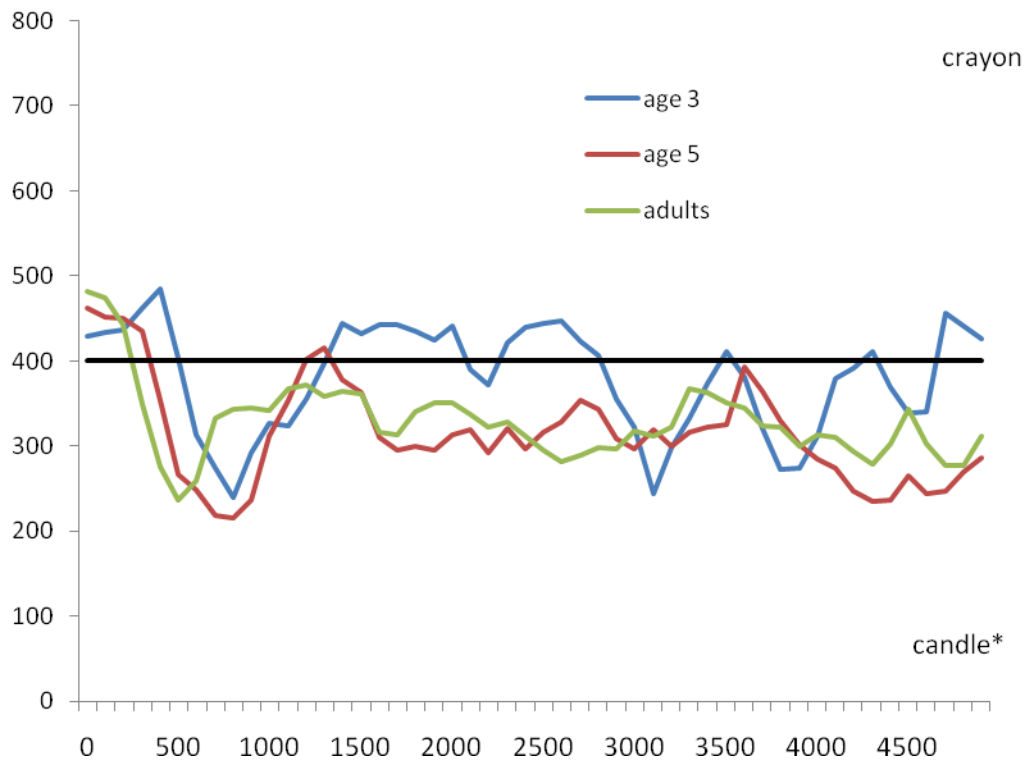


Figure 82: Eye movements of all age groups across time of the critical moment in the FB-TB trial of study 4

5.3 Discussion

The current study tested whether the conclusions made by study 2 can be expanded to another culture. Chinese-speaking children and adults were tested in the same tasks as in study 1: the elicited imitation tasks to test their linguistic representations of ToM, the traditional ToM questions, and anticipatory eye movements

in non-verbal ToM tasks. Results showed by Chinese counterparts replicated main findings from the English studies.

In the elicited imitation task, Chinese-speaking 3-year-olds, like their English-speaking counterparts, consistently omitted ToM markers. Five-year-olds were more likely than 3-year-olds to represent ToM markers linguistically. These two age groups both failed to represent second-order ToM markers. They did not differ in making other types of errors. These findings suggested a similar developmental pattern as shown in the English-speaking children: young children in both cultures failed to represent ToM markers.

In the traditional ToM tasks, 3-year-olds failed to give correct answers to questions in both the change-of-location task and the unexpected-content task. This also confirmed the findings in English-speaking children, and the findings from previous traditional ToM research including different languages (Wellman, et al, 2001).

In both the change-of-location and the unexpected-content tasks, anticipatory eye movements of adults and 5-year-olds showed an automatic and real-time process of the ToM reasoning, 3-year-olds failed to show the non-verbal ToM reasoning as 5-year-olds and adults. These findings are also consistent with the developmental trend found in English-speaking children and adults in study 1.

In conclusion, this current study shows a similar pattern as the English studies. Chinese-speaking 3-year-olds also have difficulty representing ToM markers linguistically in Chinese. Also they failed to show a non-verbal ToM reasoning in their anticipatory eye movements. Five-year-olds and adults show both the linguistic representations of ToM markers and an automatic and real-time process of ToM reasoning.

6. General Discussion

The current studies were designed to investigate the source of children's difficulties in traditional ToM tasks. They are based on the observation that traditional ToM tasks require two processes: the automatic and real-time ToM reasoning, and the linguistic representations of ToM for understanding mental state questions. It was hypothesized that both are necessary to succeed in the traditional ToM tasks; lacking of either one of these two processes can result in a pattern of responses observed among 3-year-olds in traditional ToM tasks.

The ToM reasoning was measured using an implicit measurement, the anticipatory eye movements, while participants watch videos of false (and true) belief events. Given the prior research on implicit ToM, it may be surprising that results from the current studies mostly confirm findings based on traditional ToM tasks. As a group 3-year-olds in the present studies failed to make anticipatory eye movements according to an adult-like ToM. Instead, their eye movements suggest that their eye movements often go to the wrong places or objects, confirming the findings of traditional ToM research. Indeed, children's implicit ToM performance (i.e., anticipatory eye movements) is significantly correlated with their explicit ToM performance (verbal answers to ToM questions). Depending on the task, some 4-year-olds and older children exhibit adult-like anticipatory eye movements, suggesting they are able to reason about social events

in a mentalistic framework and derive predictions in real time. I will discuss the implications of these findings in more details in the next section.

To understand young children's linguistic representations of ToM, an elicited imitation task is adopted as a measure for the strength of encoding linguistic ToM markers. Compared to adults and 5- to 6-year-olds, 3- and some 4-year-old children consistently omit ToM markers in repeating questions used in traditional ToM tasks. This strongly suggests that young children may mistake a mental state question as a factual question in a traditional ToM test, where the identical questions were used and the information processing demand is higher than in the simple elicited imitation task. This is supported by an item-level analysis, in which children's verbal answers to ToM questions are predicted by whether or not they omit the ToM marker in repeating the question in the same trial. This suggests that children need to correctly represent ToM markers in order to answer the ToM questions correctly.

In summary, the current studies suggest that young children have difficulty in performing non-verbal ToM reasoning and representing ToM markers linguistically. Both of these two deficits were associated with their performance in answering ToM questions. In order to succeed in the traditional ToM tasks, children need to possess a non-verbal ToM reasoning and to represent ToM markers linguistically.

6.1 Performance in Traditional ToM Tasks

The current studies replicated main findings of the mainstream ToM research (Gopnik & Astington, 1988; Wellman et al., 2001; Wimmer & Perner, 1983). Not only the verbal responses in the current studies, but also the anticipatory eye movements in the non-verbal tasks show the similar developmental trend as in the traditional ToM tasks. This confirms the findings of the meta-analysis conducted by Wellman, et al. (2001) that the developmental patterns are similar across all types of tasks and questions.

The current studies also suggest that in order to pass the traditional ToM tasks, children at least need to correctly represent linguistic ToM markers and to reason about mental events in real time. Lacking of either of these two processes would lead to the failure in the traditional ToM tasks. In the current studies, 3-year-old children consistently failed to represent linguistic ToM markers, neither do they show a non-verbal ToM reasoning in their anticipatory eye movements. Interestingly, there does not seem to be a direct link between the linguistic representation of ToM marking and real-time ToM reasoning in these data. It seems that the two processes independently contribute to the difficulties 3-year-olds have in answering ToM questions.

6.2 Non-verbal ToM Reasoning

Adults' anticipatory eye movements in the non-verbal tasks showed that they possess an automatic and real-time ToM reasoning. Three-year-olds and some 4-year-

old children did not show this non-verbal ToM reasoning in their eye movements. Children's eye movements of ToM reasoning in this task is associated with their traditional ToM performance. This implies that children need to have a non-verbal ToM reasoning in order to succeed in the traditional ToM questions.

In the current studies, stories in the traditional ToM tasks were presented with cartoons and real-person videos with minimum language involved. The videos were designed by keeping the complexity of the classic ToM stories in order to test the issue why children fail in the traditional ToM tasks. Therefore, the scenario in the videos of current studies are more close to the real social situation with a lot of details compared to the scenario used in the Onishi & Baillargeon (2005)'s study. The difference between the current findings and previous research on implicit understanding of false belief (Clements & Perner, 1994; Onishi & Baillargeon, 2005; Southgate, et al., 2007) may be due to different experimental settings, implying that the early understanding of false belief found in infants may be only applied to narrow and constrained situations.

Anticipatory eye movements in this study provide reliable and sensitive measurements to the automatic and real-time process of ToM reasoning. The non-verbal ToM reasoning revealed by anticipatory eye movements is a different

process from the ToM revealed in the traditional ToM tasks. With the implicit measurements, participants watched a video and do not need to produce extra responses. Thus, their eye movements reflect their automatic and real-time processes to the social information they are exposed to. In contrast, traditional tasks ask ToM questions to assess children's understanding of the protagonist's beliefs in the stories. Children's verbal responses need to be triggered by the ToM questions, which brings up the problem of their language limitations (addressed in the next section). The responses in the violation-of-expectation paradigm used by Onishi & Baillargeon (2005) also need to be triggered by the impossible events: although it also used the implicit measurements (looking time), it does not reflect the automatic responses as in the anticipatory looking paradigm. In addition, since the violation-of-expectation measure the responses after showing the ending, it does not provide a measurement for the real-time process as the anticipatory looking paradigm does.

6.3 Linguistic Representations of ToM

The correlation between linguistic representations of ToM and the explicit ToM performance is consistent with previous findings on children's linguistic representations and their false belief understanding (de Villiers & de Villiers, 2000; de Villiers & Pyers,

2002; Hale & Tager-Flusberg, 2003; Lee, et al., 1999; Lohmann & Tomasello, 2003; Milligan, et al., 2007; Tardiff, et al., 2004). However, instead of focusing on a particular syntactic structure or group of lexical items, the current study takes a functional approach. A central notion of the present research is the linguistic ToM markers, which is defined as any linguistic element that, if eliminated from a sentence, changes the interpretation of the statement or question from about mental states to about the reality. The tensed sentential complement (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002) and mental verbs (e.g., "know", "think") are examples of linguistic ToM markers in English.

These linguistic markers play a critical role in communicating mental states. Without representing the ToM markers, children would misinterpret the mentalistic questions as factual questions. Evidence from the present studies strongly suggests that the degree to which young children represent ToM markers predicts their verbal answers to ToM questions.

Does the inability to encode linguistic ToM markers imply a lack of theory of mind? Not necessarily. In the present studies there is no direct association between children's performance on the elicited imitation task and their implicit ToM reasoning, as measured by anticipatory eye movements. It is possible that children spontaneously understand social events in terms of mental states, but they have not mastered the

linguistic devices to communicate these ideas. Young children may correctly represent ToM markers when questions are shorter and less demanding. However, the current study shows that, when tested with questions identical to those used in classic ToM tasks, 3- and 4-year-old children have tremendous problems encoding the linguistic ToM markers. If they do have a rudimentary ToM, it is probably masked by the failure to understand the questions.

Finally, the focus on the linguistic ToM markers in this study is not to the exclusion of the broader relation between language and mentalizing. The acquisition of linguistic ToM markers always occurs in specific linguistic and cultural interactions. It would not be unexpected if there is a high correlation between the age at which children acquire linguistic ToM markers and their general linguistic competence (Brown et al, 1996; Cheung, Chen, Creed, Ng, Wang, & Mo, 2004; Dunn et al, 1991; Milligan, et al., 2005; Ruffman et al, 2002). For the purpose of understanding why young children fail the traditional false belief tasks, though, the linguistic marking of ToM appears to be a proximal mechanism that deserves scrutiny.

6.4 Development of ToM

The current studies show similar developmental trends of classic ToM responses as shown in the previous traditional ToM research (Gopnik & Astington, 1988; Wellman et al., 2001; Wimmer & Perner, 1983). Three-year-olds' performance in answering ToM

questions is below chance level. The performance of verbal responses is improved in 4- and 5-year-olds. These developmental trends are confirmed in both the change-of-location task and the unexpected-content task and consistent in both English- and Chinese-speaking children.

Anticipatory eye movements of 3-year-olds do not show an adult-like prediction pattern. However, it does not necessarily mean that 3-year-olds do not have a ToM competence. This may be due to their limitations to deal with the complex social situation in real time. In a simplified version of the ToM story shown by Onishi & Baillargeon (2005) even infants show early understanding of beliefs, although this finding is still under controversy (Perner & Ruffman, 2005; Ruffman & Perner, 2005; Surian, et al., 2007). That suggests that if the stories are simplified as in Onishi & Baillargeon (2005)'s study, 3-year-olds can show an understanding of mental states. The stories presented in the current task are designed to replicate the classic ToM stories nonverbally. The stories in the current tasks provide a more complex situation with a lot of details to infer the person's mind, and thus are more close to the real social situation than Onishi & Baillargeon (2005). Also the anticipatory eye movements require real-time and immediate responses, though implicit, to the social situation. Three-year-olds' performance in this task could be impacted by their limited information processing capacities to encode social information in a short time and to make immediate inferences

of others' mental states. The current studies also used different experimental setting with Clements & Perner (1994) and found implicit understanding of beliefs in different ages. This implies that although infants may obtain the competence to reason about other people's false belief as shown in Onishi and Baillargeon (2005), this competence may only be applied to constrained situations. Children may fail to keep track of people's mental states in more complex social situations.

The acquisition of linguistic markers of ToM is relatively late compared to other categories of words (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Limber, 1973; Shatz, et al., 1983). As discussed in the Chapter 1, this may be due to the low frequency and unreliability of the ToM markers in young children's language input. In most of the cases in young children's language input, the ToM markers are used in a situation in which they can be ignored from the utterances without changing the meaning of the utterances. This makes children not always attend to the ToM markers, which explains why they consistently omitted ToM markers in the elicited imitation task in the current studies.

6.5 ToM, Language, and Culture

The developmental trends of ToM shown in the English-speaking participants were also confirmed with Chinese-speaking participants. This is consistent with the conclusion in Wellman et al. (2001)'s meta-analysis that the developmental trend of ToM

is universal. It is reassuring that children in both countries demonstrated similar performances in the implicit anticipatory eye movement task, where one does not predict cultural differences. It is also interesting that, despite lexical and syntactic differences in linguistic ToM markers in the two languages, 3-year-olds consistently fail to repeat the ToM markers in elicited imitation tasks. It suggests that the failure in encoding the markers is rooted in a problem deeper than the surface structure of the questions.

This raises the interesting question of whether certain linguistic features and/or cultural practices may enhance or hinder the acquisition of ToM markers and ultimately the competence of social understanding. This is clearly beyond the scope of the present study, but it calls for well-designed cross-cultural developmental studies in the future.

6.6 Conclusions

In conclusion, the current studies show that the difficulty of young children in the traditional ToM tasks stem from two sources: first, 3- and some 4-year-olds fail to reason about social events in mental terms in real-time; second, these young children may also have problems understanding linguistic markers in traditional ToM questions, without which the questions become factual rather than mentalistic. Findings from these studies support unequivocally the classic finding that 3-year-old young children fail to demonstrate adult-like theory of mind in traditional false belief tasks. However, instead

of concluding that 3-year-olds do not have a theory of mind, the current studies strongly suggest that we need to look into limitations young children have in keeping track of complex social events in real time and in understanding language conventions in real time.

References

- Ambridge, B., & Pine, J. M. (2006) Testing the agreement/tense omission model using an elicited imitation paradigm, *Journal of Child Language*, 33, 879-898.
- Astington, J. (1999). What is theoretical about the child's theory of mind? A Vygotskian view of its development. In *Lev Vygotsky: Critical assessments: Future directions, Vol IV* (pp. 401-418). Florence, KY: Taylor & Francis/Routledge.
- Astington, J. (2003). Sometimes necessary, never sufficient, false belief understanding and social competence. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: implications for typical and atypical development* (pp.13-38). New York: Psychology Press.
- Astington, J. W., & Baird, J. A. (Eds.) (2005). *Why language matters for theory of mind*. New York: Oxford University Press.
- Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion*, 9, 151-165.
- Baird, J. A., & Astington, J. W. (2004). *The role of mental state understanding in the development of moral cognition and moral action*. In J. A. Baird & B. W. Sokol (Eds.) *Connections between theory of mind and sociomoral development. New Directions for Child and Adolescent Development*, 103, 37-49. San Francisco: Jossey-Bass.
- Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. New York, NY: Oxford University Press.
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1).
- Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, 18(6), 906-921.
- Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, 3(3), 253-272.
- Butterworth, G., & Groer, L. (1990). Joint visual attention, manual pointing, and

preverbal communication in human infancy. In *Attention and performance 13: Motor representation and control* (pp. 605-624). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversation with friends, siblings, and mothers. *Child Development, 67*, 836-884.

Cheung, H., Chen, H.-C., Creed, N., Ng, L., Wang, S. P., & Mo, L. (2004). Relative Roles of General and Complementation Language in Theory-of-Mind Development: Evidence From Cantonese and English. *Child Development, Vol, 75(4)*, 1155-1170.

Clements, W., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9(4)*, 377-395.

Csibra, G., & Southgate, V. (2006). Evidence for infants' understanding of false beliefs should not be dismissed. *Trends in Cognitive Sciences, Vol, 10(1)*, 4-5.

Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 1*, 568-570.

de Villiers, J., & Pyers, J. (2002). Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17(1)*, 1037-1060.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind*. Hove, UK: Psychology Press.

Dunn, J., & Brophy, M. (2005). Communication, relationships, and individual differences in children's understanding of mind. In J. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 50-69). Oxford: Oxford University Press.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & et al. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development, 62(6)*, 1352-1366.

- Dunn, J., & Cutting, A., L., (1999) Understanding others, and individual differences in friendship interactions in young children. *Social Development, 8*, 201-219.
- Frith, C., & Frith, U. (1999). Interacting minds – a biological basis, *Science*, 286, 1692-1695.
- Gleason, J. B. (2000). *The development of language*. Boston: Allyn and Bacon.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*, 26-37.
- Gordon, R. M. (1986). Folk psychology as simulation, *Mind and Language, 1*, 156-171.
- Greenwood, J. D. (1991). *The future of folk psychology: Internationality and cognitive science*. New York: Cambridge University Press.
- Hale, C. M., & Tager-Flusberg, H. (2003). The Influence of Language on Theory of Mind: A Training Study. *Dev Sci, 6*(3), 346-359.
- Harris, P. L. (1989). *Children and emotion: the development of psychological understanding*. Oxford: Blackwell.
- Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 57*(3), 567-582.
- Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*(2), 233-253.
- Johnson, C. N., & Maratsos, M. P. (1977). Early Comprehension of Mental Verbs: Think and Know. *Child Development, 48*, 1743-1747.
- Keller-Cohen, D. (1981). Elicited imitation in lexical development: evidence from a study of temporal reference, *Journal of Psycholinguistic Research, 10*, 273-288.
- Lanlonde, C. E., & Chandler, M. J. (1995). False belief understanding goes to school: on the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion, 9*, 167-185.

- Lee, K., Olson, D. R., & Torrance, N. (1999). Chinese children's understanding of false beliefs: the role of language. *J Child Lang*, 26(1), 1-21.
- Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: conceptual deficit or linguistic artifact? *Child Dev*, 61(5), 1514-1519.
- Lillard, A. (1999). Developing a cultural theory of mind: The CIAO approach. *Current Directions in Psychological Science*. Vol, 8(2), 57-61.
- Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.
- Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M., (2004). Decoupling beliefs from reality in the brain: an ERP study of theory of mind, *Cognitive Neuroscience and Neuropsychology*, 15, 991-995.
- Lohmann, H., Tomasello, M., & Meyer, S. (2005). Linguistic communication and social understanding. In J. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 245-265). Oxford, UK: Oxford University Press.
- Lust, B., S Flynn, S., & Foley, C. (1996). *Methods for Assessing Children's Syntax*, MIT Press.
- McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6(2), 203-229.
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2).
- Montgomery, D. E. (2005). The Developmental Origins of Meaning for Mental Terms. In *Why language matters for theory of mind* (pp. 106-122). New York, NY: Oxford University Press.
- Newton, A., & De Villiers, J. G. (2007). Thinking while talking: adults fail nonverbal false-belief reasoning. *Psychological Science*, 18(7), 574-579.

- Olson, D. R. (1988). On the origins of beliefs and other intentional states in children. In *Developing theories of mind* (pp. 414-426). New York, NY: Cambridge University Press.
- Onishi, K., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, 308(5719), 255-258.
- Perner, J., (1991). *Understanding the representational mind*. Cambridge, MA: The MIT Press.
- Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science*, 308, 3.
- 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, 431-471.
- Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. In *Pathologies of belief* (pp. 123-144). Malden, MA: Blackwell Publishing.
- Polak, A., & Harris, P. L. (1999). Deception by young children following noncompliance. *Developmental psychology*, 35(2), 561-568.
- Povinelli, D., Reaux, J., Bierschwale, D., Allain, A., & Simon, B. (1997). Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees. *Cognitive Development*, 12(4), 423-461.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*. Vol, 1(4), 515-526.
- Roth, D., & Leslie, A. M. (1998). Solving belief problems: toward a task analysis. *Cognition*, 1-33.
- Ruffman, T., & Perner, J. (2005). Do infants really understand false belief?: Response to Leslie. *Trends in Cognitive Sciences*, 9(10), 462-463.
- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Dev*, 73(3), 734-751.

- Saltmarsh, R., Mitchell, P., & Robinson, E., (195) Realism and children's early of caretaking casualty. In F. D. Horowitz, M. hetherington, S. Scarr-Salapatek, & G. Siegel (Eds.) *Review of child development research* (4). Chicago, IL: The University of Chicago Press.
- Scholl, B., & Leslie, A. (1999). Modularity, Development and 'Theory of Mind'. *Mind & Language*, 14(1), 131-153.
- Shatz, M., Wellman, H. M., & Siber. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. 14(3), 301-321.
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587-592.
- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental psychology*, 30(3), 395-402.
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 17(8), 580-586.
- Tager-Flusberg, H. (2001). A re-examination of the theory of mind hypothesis of autism. In J. Burack, T. Charman, N. Yirmiya & P. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 173-193). Mahwah, NJ: Erlbaum.
- Tager-Flusberg, H., & Joseph, R. M. (2005). How language facilitates the acquisition of false-belief understanding in children with autism. In J. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 298-378). Oxford: Oxford University Press.
- Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition*. Vol, 76(1), 59-89.
- Tardif, T., Wellman, H. M., & Cheung, K. M. (2004). False belief understanding in Cantonese-speaking children. [Original]. *Journal of Child Language*, 31(4), 779-800.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav Brain Sci*, 28(5).

- Watson, A. C., Nixon, C. L., Wilson, A., & Capagne, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology, 35*, 386-391.
- Wellman, H. M., (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development, 72*(3), 655-684.
- Wimmer, H., & Hartl, M. (1991). Against the Cartesian view on mind: young children's difficulty with own false beliefs. *British Journal of Developmental Psychology 9*, 125-138.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128.

Biography

Birth:

September, 15, 1980, Fujian, China

Education:

Ph.D., Psychology, 2009, Duke University.

M.A., Psychology, 2007, Duke University.

B.S., Psychology, 2003, Peking University, China

Publications and Presented Research:

Feng, G. & Yi, L. (2009, April).. Eye Movements as a Window to the Theory of Mind. Paper Symposium presented on the biennial meeting of Society for Research in Child Development, Denver, US.

Feng, G. & Yi, L. (2009, March). Anticipatory eye movements and Theory of Mind, Talk Presentation on the Annual Meeting of the North Carolina Cognition group, Durham, NC, US.

Feng, G. & Yi, L. (2008, Aug). Peeking into the Theory of Mind: Anticipatory eye movements and Children's Understanding of False Beliefs. Paper presented on the Tobii Eye Movement Conference, Berlin, Germany

Yi, L., & Feng, G. (2007, March). Theory of Mind and executive control in preschools: an eye-movement study. Poster presented at the North Carolina Cognition Conference, Chapel Hill, NC.

Yi, L., & Feng, G. (2007, March). Eye Movements during Card Sorting Games and Correlation with Theory of Mind Tasks. Poster presented at the biennial meeting of Society for Research in Child Development, Boston, MA.

Feng, G., & Yi, L. (2006). What if Chinese had Linguistic Markers for Counterfactual Conditionals? Language and Thought Revisited. Conference paper of the 28th Annual Conference of the Cognitive Science Society

Awards:

2007 Graduate Award for Research & Training, Center for International Studies, Duke University.

2007 Summer Research Fellowship, Graduate School, Duke University

2006 Sixteenth Annual Sally Hughes-Schrader Travel Grant, Sigma Xi, the Scientific Research Society

2007 International Research Travel Award, Graduate School, Duke University

2005 Summer Research Fellowships, Asian Pacific Study Institute APSI, Duke University