Exploring the Association between Sibling Relationship Quality, Parenting Styles, and Theory-of-Mind Development in Chinese Young Adolescents: A Preliminary Analysis

Yining Shi

MSc Education (Child Development and Education), 2023

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Yining Shi
Date:	06/10/2023

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Prof Sonali Nag. During every meeting with her, the invaluable and detailed advice provided was immensely beneficial to me. This has been enlightening for both my academic endeavours over the past year and for my future academic path. My supervisor not only provided guidance in terms of knowledge but also deepened my understanding of the rigorous academic attitude required. Secondly, I would like to extend my appreciation to Principal Liang and Principal Shu for their willingness to allow their schools to participate in my research and for actively responding to my various requests. I also owe a debt of gratitude to my research assistant Jiayu yobo who dedicated a significant amount of time to the data coding for this paper. Without her help, this thesis would not be complete. I would like to thank all the researchers who provided valuable suggestions for this thesis. I am deeply grateful to my parents Yongqin and Jingmin for their unwavering trust in me. Lastly, I wish to express my profound appreciation and love to my boyfriend Kevin, who created a loving and joyful environment for me to complete my thesis. Not only did he continually encourage me, but he also inspired my research ideas. I am delighted to have met all the peers in the Department of Education and St Hugh's college. I had a very amazing year in Oxford.

Abstract

Siblings' influence has often been overshadowed globally by other family factors. This research gap is especially pronounced in China, largely due to the four-decade-long One-Child policy. To date, only one study by Hou et al. (2022) has examined the role of sibling relationship quality (SRQ) in the theory of mind (ToM) development among Chinese children. Since this study, along with previous research predominantly from Western cultures, has focused on preschoolers, the present study sought to employ a cross-sectional, multimethod, multi-informant correlational design to verify the relationship between Chinese young adolescents' SRQ and ToM performance. Furthermore, this study incorporated parenting styles as a parental factor, which has been demonstrated to influence both SRQ and ToM. Consequently, the second aim of this research was to explore the parenting-ToM and parenting-SRQ correlations. Lastly, the study also assessed the associations between sibling structures (number of siblings, birth order, age gap, gender composition) and ToM.

Thirty families (comprising 30 young adolescents aged 11-12 years, their 30 siblings, and 30 parents) participated in individual online research sessions, with one family per session. In each session, the SRQ was gauged through questionnaires completed by both children in the sibling pair, as well as by researcher observations during a cooperative drawing game called Etch-a-Sketch Online. The ToM skills of the young adolescents were evaluated using the performance-based task, Strange Stories. Parenting styles were explored through interviews.

This study found that young adolescents' self-rated sibling intimacy was negatively associated with cognitive ToM. In contrast, the parents' strategy of referring to their own feelings when addressing social incidents with children, termed as Parent Emotions, was positively correlated with cognitive ToM scores. Moreover, when accounting for the Parent Emotions strategy, the self-rated sibling intimacy remained a significant predictor of cognitive ToM. Young adolescents with parents who often employed the Active Non-interference strategy (i.e., non-involvement based on trust in children) scored higher on emotional ToM tasks, while those with parents favouring Practical Solutions (i.e., taking direct action to solve conflicts without discussing with children) scored lower. Further, the study revealed that parents who frequently used

Active Non-interference had children with more negatively observed SRQ by researchers. In terms of demographic information, all sibling structure variables showed no correlation with ToM abilities. However, the parents' sibling status (with versus without siblings) was found to be related to their parenting styles. This study not only bridges the research gap but also offers insightful avenues for future research.

List of Abbreviations

ToM	Theory of mind
FBU	False belief understanding
SRQ	Sibling relationship quality
WEIRD	Western educated industrialised rich and democratic
MST	Mental state talk
EU	Emotion understanding
EFs	Executive functions
PIS	Participant information sheet
ESO	Etch-a-sketch online
ICC	Intraclass correlation
PAC	Positive Affective Climate
NAC	Negative Affective Climate
SES	Socioeconomic status

Tables

Figures

Table of Contents

Chapter 1: Introduction

1.1 Background and Research Gaps

1.1.1 Theory of Mind (ToM)

ToM, the ability to understand others' mental states such as desires, beliefs, and emotions, is essential for daily life as it enables reasoning and predicting behaviours (Premack & Woodruff, 1978; Wellman & Woolley, 1990). It is considered a key skill in social information processing, influencing various developmental and academic outcomes (Yeates et al., 2007). In the early years, children struggle to differentiate their own mental states from others' (Keysar et al., 2003). Around the age of four, children develop first-order false belief understanding (FBU), recognising that someone may hold a belief that does not align with reality from their own perspective (Wimmer & Perner, 1983). By around the age of eight, children grasp second-order FBU, acknowledging that someone can have a false belief about another person's beliefs (Perner & Wimmer, 1985).

According to Wellman and Liu (2004), there are additional ToM skills apart from FBU. These include understanding diverse desires (i.e., people can have preferences for certain objects), diverse beliefs (i.e., people can have different beliefs about the same thing), knowledge ignorance or knowledge access (i.e., someone may not know what the reality is), and hidden emotion (i.e., someone can feel one way but display a different feeling). Among these five types of skills, diverse desires were found to be the easiest for children to grasp, while hidden emotion was the most challenging (Wellman & Liu, 2004).

Due to the interest in these developmental milestones of ToM, research has primarily focused on preschoolers (3-5 years old) and young school-aged children (6-9 years old; Hughes & Devine, 2015; Milligan et al., 2007). While ToM is vital for young children to start to navigate the social world, it continues to play a crucial role in adjusting to a broader social environment and developing advanced social skills in young adolescence (10-12 years old) and beyond. However, despite the significance of this stage characterised by cognitive and socioemotional changes, ToM development

during young adolescence has been relatively neglected in the field (Eccles et al., 2003; Hughes & Devine, 2015; Valle et al., 2015).

1.1.2 The Influences of Siblings

Numerous studies have emphasised the significance of family influences on children's ToM development, considering the family environment as the primary social context, rather than viewing that ToM development is primarily driven by innate maturation and minimally influenced by environmental factors (Devine & Hughes, 2018; Hughes & Leekam, 2004; Perner et al., 1994). Within the realm of family-related factors, sibling-related variables including sibling structures (e.g., birth order), sibling individual characteristics (e.g., personality), and sibling relationship quality (SRQ, e.g., sibling warmth), have received less attention compared with extensively studied factors such as parental factors (e.g., parent-child relationship quality; Buist et al., 2013; Foley & Hughes, 2021; Noller, 2005).

Like other family relationships, siblings provide not only support and warmth but also conflict and stress that greatly shape one's development across lifespan (Jensen et al., 2018; McHale et al., 2012). Unlike other family relationships, sibling interactions are characterised by a combination of hierarchical (e.g., teaching) and egalitarian (e.g., competition) activities. On the one hand, due to the developmental gaps between siblings in physical (e.g., height), cognitive (e.g., language and self-regulation), and social capacities, older or faster-developed siblings often take the role of teachers, caretakers, and role models (Dirks et al., 2015; Hou et al., 2022). On the other hand, due to the similarity in age, interests, and growth environment, siblings are also peers and therefore sibling relationships contain elements of equality, reciprocity, and sharing (Hou et al., 2022; Lam et al., 2021). Despite there has been a growing body of literature recognising the significance and distinctiveness of siblings in the past two decades, most of them focused on families from Western Educated Industrialised Rich and Democratic (WEIRD) countries like the United Kingdom (Henrich et al., 2010; Shahaeian, 2015).

1.1.3 The Context of China

In China, the study of siblings has been largely overlooked due to the implementation of the four-decade-long one-child policy. However, prior to the implementation of the two-child and three-child policies in 2016 and 2021 respectively, there were relaxations to the one-child policy. In 2002, couples were permitted to have two children if both parents were only children at home, and in 2013, the policy was further revised to allow couples to have two children if either of them was an only child. As a result, many of today's young adolescents in China, who are in the final years of primary school, may have both older and younger siblings of varying ages. This offers an excellent opportunity to explore these young adolescents' sibling relationships, as well as the impact these relationships have on their personal growth.

Examining the associations between sibling factors and ToM development in China is crucial for several reasons. Firstly, previous studies have shown differences in ToM abilities between children from Eastern collectivist societies and those from Western individualistic cultures. For example, Chinese and Iranian children indicate an understanding of knowledge access before diverse beliefs, which contrasts with the developmental sequence observed in American and Australian children's ToM performance (Fang et al., 2009; Peterson et al., 2005; Shahaeian et al., 2011; Wellman et al., 2006). Furthermore, the Chinese context is unique compared with other Eastern countries due to its adherence to the Confucian system, which highlights prioritising family needs over individual desires (Chao & Tseng, 2002; Chen et al., 1997). This cultural framework potentially results in unique family contexts and impacts the development of social cognition and social relationships in children. Therefore, conducting Chinese research on the sibling-ToM association is vital for comparisons with Western and other Eastern cultural backgrounds, allowing us to gain a comprehensive understanding of this topic. It is possible that sibling influences on ToM development in Chinese children differ compared to other cultures.

To the student researcher's[1] knowledge, Hou and colleagues (2022) have conducted the only empirical research investigating the effects of sibling structures and SRQ on Chinese children's ToM development. However, there are three major limitations to this study. Firstly, it focused solely on preschoolers. Secondly, the SRQ

[1] After this point, "the researcher" will always refer to "the student researcher".

was assessed using a mono measure reported by mono informant, namely a questionnaire completed by mothers. Thirdly, the study did not control for any parental factors known to be closely related to both SRQ and children's ToM capacities.

1.2 Research Aims

Given the significant research gaps highlighted previously, this study aimed to be the first to investigate the relationships between sibling factors (prioritising SRQ and considering sibling structures as a secondary focus) and ToM development in Chinese young adolescents. Additionally, this study sought to advance the field by (1) employing a multi-informant and multimethod approach to assess SRQ, and (2) including a parental factor, parenting style, which was previously found to be related to both SRQ and ToM, and verifying the associations among SRQ, ToM development, and parenting styles. As a preliminary study, the primary objective was to analyse the correlations among the key variables.

1.3 Dissertation Structure

The next chapter presents a literature review covering three research topics: the role of siblings in children's ToM development, the role of parenting styles in ToM, and the role of parenting styles in SRQ. This chapter first introduces the theoretical perspectives and then presents empirical evidence with a focus on research methods, outcomes, and cultural differences.

The third chapter provides details of the research design, recruitment process, participant information, measures for variables, ethical considerations, and the data analysis plan. For the measures, this chapter explains each one's origins, psychometric properties, adaptations made for the present study, and scoring system.

The fourth chapter presents the descriptive statistics of the measures and variables and the results of testing the hypotheses, represented through graphs and tables. All the results appearing in the third and fourth chapters were calculated using Microsoft Excel 16.75.2 and IBM SPSS 28.

The final chapter offers an interpretation of the findings presented in the fourth chapter and summarizes the limitations and contributions of the current study. It also discusses implications for future research design and research directions, followed by a brief conclusion.

Chapter 2: Literature Review

2.1 The Role of Siblings in ToM

2.1.1 Theoretical Perspectives

Research on the association between siblings and ToM has originated from a variety of theoretical perspectives, examining the potential effects of different sibling factors (Whiteman et al., 2011). This generally prompts researchers to explore two main research questions: does the presence of siblings benefit children's ToM development, and which sibling factor has the most significant influence on this development.

2.1.1.1 Resource Dilution Theory

An early yet influential perspective, based on the Resource Dilution Theory, suggests that finite parental resources such as time, attention, care, and money become diluted among children as the number of siblings increases (Blake, 1981; Downey, 1995). This perspective views siblings as competitors and thus asserts that there are few benefits to having siblings (Yucel & Yuan, 2015). This theory is supported by studies indicating that larger sibship sizes lead to fewer parental resources, resulting in lower levels of educational, cognitive, and social achievements in children (Downey, 1995, 2001). Moreover, Downey (1995) found that even though children can still benefit from certain interpersonal (e.g., the frequency with which their parents talk to them) and financial (e.g., the number of educational objects like books) parental resources, these benefits tend to diminish as the number of siblings increases. Therefore, this perspective predicts that the presence of siblings, or a larger sibship size, may negatively impact child development.

However, some researchers hold a less pessimistic view about competition for resources among siblings. Grounded in Adler's theory of individual psychology, which suggests that siblings play a pivotal role in one's personality development, Sulloway (1996) posited that children themselves can develop sibling differentiation (e.g., personality) to carve out unique niches within the family. This strategy serves to minimize sibling competition and maximize their opportunities to access resources.

2.1.1.2 Social Learning Theory

In contrast to the Resource Dilution Theory, many other theoretical perspectives propose that siblings can serve as valuable resources to each other (Downey & Condron, 2004). According to Social Learning Theory, individuals acquire new behaviours, attitudes, and beliefs through two mechanisms: reinforcement and observation of others' behaviours (Bandura, 1977). For children, family environments provide numerous opportunities for these social learning processes to occur. Children primarily develop their social competencies by reciprocally reinforcing positive or negative behaviours with their family members and by observing and imitating interactions between other family members (Kramer, 2014).

According to Bandura (1977), family members who hold higher status, show more nurturance and warmth can serve as powerful role models for children. This suggests that children tend to engage in more social learning with parents and older siblings, especially those with whom they have a significant age gap and a strong, harmonious relationship. However, the theory also indicates that individuals are more likely to imitate models who share greater similarity with them, suggesting that children may be more influenced by siblings of the same gender and those close in age (Whiteman et al., 2011). Therefore, within the framework of Social Learning Theory, there exists a debate over whether older siblings with a significant age gap or siblings close in age have a greater impact on child development.

2.1.1.3 Apprenticeship Model and Age Threshold Model

Addressing this debate, the Apprenticeship Model was proposed, suggesting that children with older siblings, rather than younger siblings or siblings close in age, have an advantage in ToM development (Lewis et al., 1996; Perner et al., 1994; Ruffman et al., 1998). This model is based on Vygotsky's theory of sociocultural cognitive development, emphasizing the unilateral effect where children enhance their social and cognitive development through interactions with more knowledgeable individuals (Vygotsky, 1978). In this scenario, older siblings, who have more social experiences and higher levels of receptive and expressive language abilities, act as "social mentors" and provide more social opportunities for their younger siblings (the

"apprentices"). This mentoring allows the younger siblings to improve their understanding of others' mental states—a benefit the less socially experienced younger siblings cannot reciprocate.

The Age Threshold Model refutes the unidirectional effect and proposes a bidirectional benefit for ToM abilities within sibling dyads (Kennedy et al., 2015). This model suggests that younger siblings can enhance their older siblings' ToM development if they reach an age where they have sufficient language abilities and social experiences. This enables them to generate quality social interactions (e.g., pretend play and arguments), further contributing to their older siblings' ToM progress. For instance, some researchers propose that for preschoolers, interactions with infant siblings younger than 12 months and older siblings over 12 years old may not help them reflect on mental states (Peterson, 2000). This aligns with Piaget's theory of cognitive development, which views the developing child (e.g., younger siblings) as an active participant in their own development rather than a passive recipient of either maturation or environmental interactions (Piaget, 1957).

2.1.1.4 Reciprocal and Dynamic Development Model

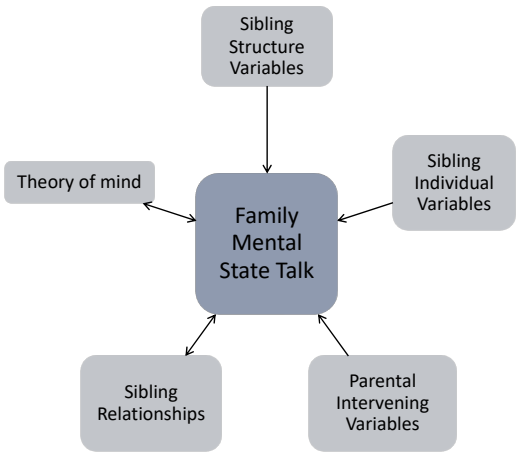
The theories and models discussed earlier place emphasis on structural sibling variables such as birth order (i.e., older versus younger siblings). However, following investigations into the links between these structural variables and sibling relationships, Buhrmester and Furman (1990) suggested that structural characteristics alone cannot fully account for sibling effects and relationship dynamics. Indeed, the processes through which sibling structures influence ToM abilities have never been directly measured and have been inferred only based on sibling outcomes (McHale et al., 2012). As such, Dunn (1983) advocated moving beyond structural variables to focus on sibling relationships within a larger family system and on influence processes that can be directly measured. These proposed areas of research focus remain unresolved to this day.

Aligned with Dunn's ideas, Hou et al. (2020) proposed a Reciprocal and Dynamic Development Model to further elaborate on the sibling-ToM associations. As illustrated in Figure 1, this model acknowledges sibling structure variables and individual characteristics, yet it recognizes their potential impact on ToM development

mainly through siblings' mental state talk (MST), which is considered one type of family MST. Like the Age Threshold Model, this model posits that sibling influences on ToM should be reciprocal and for these reciprocal effects to occur, both siblings must be within a certain age range. However, unlike the Age Threshold Model, Hou et al.'s (2020) model emphasizes that sibling effects on each other can be both positive and negative, depending on the quality of their interactions—specifically, the quantity and quality of MSTs. The authors posit that only those sibling dyads who have higher intimacy and closeness can engage in quality daily interactions that are likely to involve MST, such as cooperative play (Hou et al., 2020). In other words, this model underscores the importance of SRQ, as it impacts siblings' ToM development through sibling interactions.

Figure 1

Reciprocal and Dynamic Development Model



Note. This is a simplified version adapted from the original model by Hou et al. (2020).

Moreover, this model suggests that ToM can also influence children's SRQ through sibling interactions. The authors propose that higher levels of ToM abilities can facilitate the occurrence and frequency of positive MST, subsequently promoting affection between siblings (Hou et al., 2020). The proposed effect of ToM on SRQ aligns with Heider's (1958) Attribution Theory. This theory emphasizes attributions for individuals' reactions to interpersonal events and suggests that harmony and rivalry in social relationships (e.g., sibling relationships) result from one person's understanding

of the mental states (e.g., motivations) underlying another's behaviors (Whiteman et al., 2011).

2.1.2 Empirical Evidence

2.1.2.1 Number of Siblings

In a pioneering study, Perner et al. (1994) determined that the number of siblings significantly influenced the ToM performance in 80 children, aged three and four, during a false-belief task. Their findings indicated that children with two siblings achieved about a year's advancement in ToM compared to those without siblings. Jenkins and Astington (1996) corroborated the influence of sibship size on FBU after controlling for the age and language abilities of 68 preschoolers. Their research also suggests that the presence of siblings could compensate for slower verbal skill development in preschoolers as it pertains to FBU. However, these two studies have small sample sizes. Foley and Hughes (2021) pointed out that studies with sample sizes below 100 are likely underpowered. Furthermore, these studies only utilized FBU as a measure of ToM. Yet, Bloom and German (2000) contended that false belief tasks might not be the most accurate measure for ToM, citing reasons such as the tasks requiring abilities beyond ToM, like working memory.

Furthermore, several studies have been unable to replicate the significant effect of sibling presence on ToM (Calero et al., 2013; Cole & Mitchell, 2000; Hou et al., 2022; Peterson, 2000). For instance, Cutting and Dunn (1999), with a sample size of 128, found that the number of siblings was not related to preschoolers' FBU and emotion understanding (EU). Shahaeian (2015), in a study with 142 Iranian participants, found no significant correlations between the number of siblings and preschoolers' ToM abilities, as assessed by Wellman et al.'s (2006) five ToM tasks (i.e., diverse desires, knowledge access, diverse beliefs, false beliefs, and hidden emotion). For primary-school children using ToM assessments such as Strange Stories and second-order false-belief tasks, studies also report no correlations between ToM and the number of siblings (Lecce et al., 2017; Miller, 2013). However, it is important to note a recent meta-analysis of 45 studies, which showed that for children aged three to seven, the number of siblings has a small ($r = .14$) but significant relationship with FBU (Devine & Hughes, 2018).

2.1.2.2 Siblings' Age Range, Birth Order, Gender

McAlister and Peterson (2006) tested the Age Threshold Model by examining whether there's a ToM advantage for having a child-aged sibling (1-12 years), irrespective of birth order, in a sample of 124 preschoolers. They discovered that preschoolers with at least one child-aged sibling outperformed both only children and those with siblings who were infants or adults in a ToM battery (e.g., false belief task and appearance-reality task). Furthermore, the number of child-aged siblings positively influenced preschoolers' ToM scores, even after adjusting for their language abilities. Importantly, the effect of having a specific number of child-aged siblings was also confirmed longitudinally (McAlister & Peterson, 2007, 2013), offering robust evidence for the direction of causality (Gass et al., 2007).

To test the Age Threshold Model from a different angle, Paine et al. (2018) conducted a study involving 229 first-born seven-year-olds. They discovered that children with younger siblings performed better on the second-order false-belief task compared to those without siblings, even after controlling for children's verbal intelligence. This result supports the Age Threshold Model as it suggests that older siblings can benefit from having younger ones. To delve deeper into this younger-sibling effect, the authors categorized the seven-year-olds into four groups based on the age gaps with their siblings: (1) no sibling, (2) early arrival ($\leq$ 24 months), (3) average arrival ($>$ 24 months and $<$ 43 months), and (4) late arrival ($\geq$ 43 months). However, only the average- and late-arrival group, not the early-arrival group, showed benefits from having a younger sibling in terms of ToM. This outcome contradicts the Age Threshold Model because it indicates that even if siblings close in age to the target children possess adequate language and social skills, they don't enhance the target children's ToM.

Several studies from both Eastern and Western contexts have contested the Age Threshold Model (Farhadian et al., 2010; Ruffman et al., 1998; Wright & Mahfoud, 2012). For example, using a sample of 192 children, Kennedy et al. (2015) examined the effects of sibling structure variables on the ToM abilities of children aged four to 11 using the interpretive ToM task. This task evaluates a child's capability to understand that a single situation can have multiple legitimate interpretations (Chandler & Helm,

1984). The researchers found that, after adjusting for age and executive functions (EFs), only the number of older siblings, not younger ones, influenced children's ToM scores. This held true even when all the siblings were within the same age range (i.e., 6-10 years). This result strongly supports the Apprenticeship Model. However, it contrasts with a meta-analysis which found that the presence of child-aged siblings, rather than birth order, influences children's FBU (Devine & Hughes, 2018). This discrepancy may arise from evaluating different facets of ToM (e.g., FBU versus other aspects).

Importantly, Prime et al. (2016) found in a study involving 385 preschoolers using Wellman et al.'s (2006) five ToM tasks that having more older siblings reduced ToM abilities when those siblings had low cognitive sensitivity, yet siblings' high cognitive sensitivity negated this effect. Furthermore, Wright and Mahfoud (2012) reported that having younger siblings adversely affected three to six-year-olds' ToM. These results support the Reciprocal and Dynamic Development Model, as they illustrate that the effects of siblings can be negative, contingent upon interaction quality influenced by siblings' cognitive sensitivity (Hou et al., 2020).

Kennedy et al. (2015) found that neither the number of female nor male siblings influenced ToM scores. This aligns with Ruffman et al. (1998), who observed no impact of siblings' gender in both English and Japanese children. Yet, Kennedy et al. noted a positive correlation between the number of same-sex siblings and ToM performance, irrespective of age. In contrast, Cassidy et al. (2005) discovered that preschoolers with at least one opposite-sex sibling outperformed in false belief tasks compared to those with only same-sex siblings. As with studies on sibling count, birth order, and age range, results on sibling gender composition are inconsistent. Notably, research focusing on sibling gender composition is scarcer than that on other sibling structures. Beyond the gender composition of sibling pairs, evidence suggests that the gender of the target child can also influence ToM across ages, especially in affective ToM (Calero et al., 2013; Ibanez et al., 2013). However, just as the findings regarding previously mentioned factors are mixed, some studies report no gender differences in children's ToM (Hughes & Dunn, 1998).

2.1.2.3 Sibling Relationship Quality

Failures to replicate the effects of sibling structural variables have shifted research attention to sibling interactions and SRQ (Cole & Mitchell, 2000). Some studies with notable findings on sibling structure also interpret their results through the lens of SRQ. For instance, the aforementioned study by Kennedy et al. (2015) highlighted the positive influence of same-gender sibling composition, attributing this to a higher likelihood of intimacy between same-gender siblings. Yet, only a handful of studies have directly examined the link between SRQ and ToM development, with a predominant focus on toddlers and preschoolers, and limited attention to school-aged children and adolescents.

Howe et al. (2002) suggested that children with higher sibling warmth often display more advanced internal state language, aiding them in understanding the mental state differences between themselves and others. For example, when siblings face challenges, a positive SRQ can lessen destructive resolutions like blaming or hitting and promote constructive conflict resolutions. These constructive resolutions include clarifying personal intentions and comforting based on the other's emotions (Recchia & Howe, 2009). Such interactions offer children opportunities to hone both cognitive and affective perspective-taking skills (Hughes et al., 2006; Lagattuta & Wellman, 2002). Further, these refined perspective-taking skills can then foster more effective problem-solving methods, such as discussion and sharing, reinforcing a close and nurturing SRQ.

One study indicated that positive SRQ, marked by actions like frequent helping and high levels of affection, was a significant predictor of two-year-olds' ToM outcomes after adjusting for factors such as EFs and parent-child relationship quality (Hughes & Ensor, 2005). Another study showed that for preschoolers, the influence of younger siblings on affective perspective-taking was amplified when there were high levels of warmth and affection between the siblings (Jambon et al., 2019). Song and Volling's (2018) longitudinal study highlighted that negative SRQ, such as antagonism, correlated with poorer ToM for 31-month-olds when mothers employed minimal child-centred discipline methods like listening to the child's explanations. Interestingly, while negative SRQ showed a correlation, positive SRQ did not. This study also underscored that early ToM capabilities could forecast later positive SRQ, pointing to a reciprocal relationship between ToM and SRQ.

In the only Chinese research investigating the link between siblings and ToM, Hou et al. (2022) explored the effects of both sibling structures and SRQ on preschoolers' ToM using Wellman et al.'s (2006) five ToM tasks. They utilized a large sample, comprising 113 children without siblings and 150 children with siblings. The study found that factors such as the presence of siblings, having an older sibling, and having a child-aged sibling did not enhance preschoolers' ToM development. Only positive SRQ (i.e., positive involvement) was positively correlated with children's ToM scores, while negative SRQ (i.e., sibling conflict and avoidance) showed no relation to ToM.

2.1.2.4 Strength and Limitation

Most research on the sibling-ToM relationship has focused on preschoolers, frequently using Wellman et al.'s (2006) five ToM tasks designed for this age group. This measure is beneficial as it assesses both cognitive (e.g., false belief) and emotional (e.g., hidden emotion) aspects of ToM skills. The current study contends that it is necessary to distinguish these two aspects of ToM. A potential reason for inconsistent results concerning sibling-ToM correlations might be that sibling effects differ between cognitive and emotional ToM. While some studies have employed single measures (e.g., false belief tasks) that assess only one facet of ToM, others have utilised assessments capturing both aspects, such as Wellman et al.'s (2006) tasks, but they have treated ToM holistically. The varied emphasis on cognitive and emotional ToM in each study's measure might contribute to mixed findings. In a study with 209 children aged 5.5 to 12, Wang et al. (2022) demonstrated that the number of older siblings positively predicted cognitive ToM in Chinese children but negatively for Australian children. However, for both Chinese and Australian children, the number of older siblings did not influence emotional ToM. This research substantiates that sibling effects vary between cognitive and affective ToM.

Regarding the studies on SRQ, most have only utilised parent-reported questionnaires or interviews to assess SRQ. It is reasonable to ask parents to report on SRQ for young children, given these children might lack the cognitive abilities, such as language comprehension, needed to do so themselves. However, researchers could employ observational methods to obtain rich, objective, ecologically valid data on

sibling interactions (Foley & Hughes, 2021). Moreover, the present study believes that older school-aged children (over 8 years) possess adequate abilities to report their own SRQ, and it is important for this field to hear directly from the children themselves.

2.2 The Role of Parenting Styles in ToM

2.2.1 Theoretical Background

According to Bronfenbrenner's (1979) Ecological Systems Theory, the family serves as the microsystem for children, creating immediate contexts of everyday life that have the greatest impact on individual development. Within the family context, multiple relationships, including parent-child and sibling interactions, collectively affect child development, as proposed by Family System Theory (Cox & Paley, 2003). These theories highlight the significant role of parenting styles, which represent relatively stable parental behavioural patterns that determine daily parent-child interactions, in exploring sibling-ToM associations. Crucially, the Reciprocal and Dynamic Development Model predicts the impact of how parents intervene in siblings' interactions (i.e., a concept that can be understood as parenting styles) on ToM development (Hou et al., 2020).

2.2.2 Empirical Evidence

2.2.2.1 Based on Baumrind's Classification

One line of research on the effects of parenting styles is based on Baumrind's (1978) classification of three main parenting styles: authoritarian, authoritative, and permissive. Authoritarian parenting is generally characterized by high conformity, where parents restrict and control children's behaviours. Authoritative parenting is characterized by high autonomy, as parents encourage children to develop and challenge different perspectives. Permissive parents, on the other hand, are not highly responsive to children's needs and expect them to regulate their own behaviours. Later, researchers extended these parenting styles into four categories based on two dimensions (i.e., demandingness and responsiveness): authoritarian (demanding but not responsive), authoritative (demanding and responsive), indulgent (not demanding

but responsive), and neglectful (neither demanding nor responsive; Baumrind, 1991; Maccoby & Martin, 1983).

Based on mother-rated parenting attitudes via questionnaires, a pioneering study recruited 97 children from ages three to six and found that Korean-American children with mothers who tended to employ an authoritarian style outperformed Anglo-American children with mothers who tended to apply an authoritative style in ToM tasks (Vinden, 2001). However, some argued that this is not strong evidence for parenting styles' impact on children's ToM development because the difference in ToM performance between the two groups of children can be attributed to cultural factors instead of parenting styles (O'Reilly & Peterson, 2014).

However, a later study tested this effect in Anglo-Australian children aged 5-12 years and found significant positive links between ToM and authoritative style, as well as significant negative associations between ToM and authoritarian parenting (O'Reilly & Peterson, 2014). In the Eastern background, one study conducted in Indonesia revealed that children aged four to six years showed a negative correlation between ToM and authoritarian parenting but no significant relationship with authoritative parenting (Kuntoro et al., 2017). There is a lack of Chinese studies directly testing the effect of Baumrind's parenting styles on children's ToM.

2.2.2.2 Based on Parenting Practices

Another line of research assesses parenting styles through specific parenting practices without applying Baumrind's classification. For example, some researchers have focused on the importance of the use of MST by parents (Pavarini et al., 2012). Taumoepeau and Ruffman (2006) found that maternal use of mental-state language (i.e., desire language) predicted 15-month-old children's later EU and development of mental-state language. Additionally, LaBounty et al. (2008) found that mothers' (rather than fathers') references to emotion terms and emotion causal explanations predicted preschoolers' EU, while fathers' (rather than mothers') use of explanatory references to desires and emotions predicted preschoolers' FBU.

However, in Eastern cultures, researchers identified that parents who tend to refer to people's mental states in discussions with children as a strategy to solve conflicts did not show a significant relationship with ToM in both Chinese and Iranian

preschoolers (Lewis et al., 2006; Shahaeian et al., 2014). Furthermore, more recent studies in Western countries (Canada and New Zealand) also confirmed a non-significant relationship between maternal MST and children's ToM across childhood (Carr et al., 2018; Tafreshi & Racine, 2016).

Although the results regarding the use of parental MST are mixed, there are significant findings related to other parenting practices. For example, Olson et al.'s (2011) study demonstrates that preschoolers with parents who use coercive discipline (e.g., physical punishment) experienced delays in ToM development. Furthermore, Shahaeian et al. (2014) discovered that discussing the consequences of certain behaviours with children aged 4-6 years, instead of ignoring the child during social incidents, positively related to their ToM development. Again, there is a dearth of Chinese studies verifying the relationships between parenting styles in the form of parenting practices and ToM, especially in adolescence. Nevertheless, substantial evidence validates the link between parenting styles and children's ToM across cultures.

2.3 The Role of Parenting Styles in SRQ

2.3.1 Based on Baumrind's Classification

There are many significant findings showing the connections between parenting styles and SRQ. Using Baumrind's classification, Milevsky et al. (2011) found that adolescents with authoritative parents reported greater sibling support and sibling closeness than those with authoritarian parents. A study with a culturally diverse sample (Mexican-American, European-American, and Taiwanese) also demonstrates that more authoritative and less authoritarian parenting significantly predicted more positive SRQ of preschoolers, even after controlling for child temperament (Yu & Gamble, 2008). One Chinese study with a sample of 314 two-child families indicates that authoritative parenting promotes positive SRQ, while both indulgent and neglectful parenting are detrimental to SRQ (Sun & Zhang, 2018). Moreover, they also discovered that authoritarian parenting is complex because it is associated with both high scores in positive and negative SRQ.

However, in fact, some researchers have discovered culturally specific parenting styles in China based on Baumrind's classic categories (Kim et al., 2013; Xie

& Li, 2019). For example, they developed "tiger parenting" characterized by high levels of both authoritative and authoritarian styles. Nevertheless, there is a lack of studies testing the role of parenting styles based on Chinese characteristics in SRQ.

2.3.2 Parental Involvement versus Non-Involvement

Kramer et al. (1999) found that mothers' non-involvement in siblings' conflicts was positively related to positive interactions of both the first-born (age 4.5-9) and second-born (age 3-7) children. On the other hand, maternal involvement, such as maternal control based on their superior authority and maternal child-centred problem-solving in children's conflicts, was negatively related to children's positive interactions. This result contradicts a study on adolescents (age 14-16) showing that maternal and paternal coaching involvement styles (i.e., offering advice on how to solve) in children's conflicts lead to greater sibling warmth than non-involvement style (Milevsky et al., 2011).

The contrary results might be because, firstly, parenting styles have different effects on different ages of children; secondly, the former study used home-observational parenting styles, while the latter study applied adolescent-reported parenting styles, and adult researchers and adolescent participants can have different perceptions of parenting. Crucially, a Chinese study on the identical research topic illustrates that mother-reported coaching involvement style was negatively associated, and maternal control was positively associated with sibling conflicts reported by around 12-year-old adolescents (Chen, 2019). To conclude, it seems that for adolescents, parental positive involvement could predict better SRQ regardless of culture.

2.4 The Present Study

Based on prior research findings, the researcher anticipated significant relationships between SRQ and young adolescents' ToM abilities, between parenting styles and ToM, and between parenting styles and SRQ. As secondary objectives, this study also aimed to examine whether young adolescents' age, gender, and sibling structure variables (birth order, number of siblings, sibling age gap, and sibling gender

composition) were related to ToM capacities. Given the inconsistent results across cultures and taking Hou et al.'s (2022) study as the most pertinent Chinese research on sibling-ToM associations, the researcher predicted that, apart from age, all other demographic factors were not related to ToM.

Regarding the philosophical approach, this study adopted a post-positivist paradigm. This paradigm underscores the significance of systematic empirical investigations to derive highly credible explanations or theories, without asserting that something is definitively true (Kivunja & Kuyini, 2017). It posits that a reality exists independent of human perception, yet it acknowledges that our understanding of this reality is constrained (Kivunja & Kuyini, 2017). This limitation arises because the research process and interpretations are influenced by the researcher's personal experiences and perspectives.

Chapter 3: Method

3.1 Research Design

This was a cross-sectional correlational study. Young adolescents aged 11-12 years were the target children for this study. One sibling (at least 8 years old), and one parent, who was most knowledgeable about the target child's studies and daily life, were also invited to participate. In families with more than two children, the sibling closest in age to the target child was selected to participate. The lower age limit for siblings was set to ensure sufficient abilities (e.g., comprehension) for self-reporting the SRQ through questionnaires.

This study employed a multimethod multi-informant design to assess the SRQ variable. SRQ was measured through questionnaire ratings completed by both children in the sibling dyad (section 3.4.2) and through researcher observations during a cooperative drawing game (section 3.4.3). Target children's ToM skills were assessed using the performance-based task Strange Stories (section 3.4.1), and parenting styles were investigated via interviews (section 3.4.4).

At the outset of the research planning, the researcher was unable to travel to China for data collection due to the peak of the COVID pandemic in the country. For this reason, and to ensure convenience for the families and minimize disruption to their daily lives, online research sessions were carried out using the popular Chinese online meeting platform, Tencent Meeting (https://meeting.tencent.com/download/). A separate research session was organised for each family. The procedure for every online session, along with estimated durations for each part, was as follows:

1. The target child completed the Strange Story task (10 minutes).
2. The target child completed the SRQ questionnaire (5 minutes).
3. The sibling dyad completed the cooperating drawing game (10 minutes).
4. The sibling completed the SRQ questionnaire (5 minutes).
5. The parent participated in an interview (10 minutes).

During each part, non-tested family members were asked to leave the room. Notably, if the siblings of the target children were also aged 11-12 years old, their ToM was also tested between parts 3 and 4. This situation only occurred with twins in this study, as detailed in section 3.3. The order of the procedure was designed to ensure that each participant did not need to enter and leave the room multiple times. Audio recordings were made during parts 1 and 5, while video recordings were made during part 3 for data coding purposes. The study aimed to recruit 30 families (see Appendix A for sample-size calculation details). Families were required to have a computer or tablet at home for the cooperating game. There was a preference for computer use because computers can record and allow for operation at the same time, while it is difficult for tablets to record the full picture while operating. However, if families did not have computers, tablets were allowed as an alternative. In this case, families were asked to put an additional recording device (e.g., phone) in front of children.

3.2 Recruitment

Once research ethics approval was obtained in late April, the researcher approached two urban public primary schools in Jinan, Shandong, China, to recruit participants. Shandong province, situated on the eastern coast of China, holds significant cultural importance as it is the birthplace of Confucius and is renowned for its Confucian heritage. Jinan serves as the capital city of this province. One school had 2784 students and 149 teachers, while the other school had 800 students and 45 teachers.

The researcher sent an invitation email and opt-in form to the two headmasters, providing an overview of the research team, topic, design, and participant requirements (see Appendix B). After addressing the headmasters' questions regarding participant compensation and receiving their signed opt-in consent forms, the researcher sought the schools' assistance in disseminating the research details to students and parents. The researcher also sent an invitation email with opt-in form to the headteachers of classes. The headteachers introduced the research in classrooms and shared a text and video created by the researcher through each class's teacher-parent WeChat group, a popular social media platform in China. The text included recruitment requirements and the researcher's contact details, while the video explained the research rationale

and objectives. Families with further inquiries could directly contact the researcher for more information.

The participant information sheet (PIS) for parents, the PIS for children (Appendix C), the assent form for children below 16, the guardian consent form, and the sibling consent form for those over 16 were sent to the headmasters through emails by the researcher. A designated teacher from each school collected the consent forms and returned the digital copies to the researcher. Using the contact information provided on the guardian consent forms, the researcher contacted parents and scheduled appointments for further participation.

A total of 42 families initially expressed their interest in participating in this research, either through the schools or directly to the researcher. However, ultimately only 30 families participated (22 from the first primary school and eight from the second). Twelve families declined participation due to scheduling conflicts with the Gao Kao (i.e., Chinese university entrance examination) or Zhong Kao (i.e., Chinese high school entrance examination) in early June and mid-June, respectively. This experience suggests that future studies involving Chinese siblings should exercise caution in conducting recruitment and data collection during the June exam season.

3.3 Participants

Ninety participants, that is, 30 families, took part in this online study: 30 target children (13 boys, 17 girls, ages 11 or 12, mean age 11.5 years), 30 siblings of target children (nine boys, 21 girls, ages 8 to 21, mean age 11.9 years), and 30 parents (3 fathers, 27 mothers). All of them were Chinese who grew up in China. Among the 30 families, 26 families had only two children, while four families had three children.

Within the 30 sibling dyads, five dyads were twins, having no age gap between them. Nineteen dyads had small age gaps, with the two children being two to four years apart, while six dyads had large age gaps, with the two children being at least five years apart. It is worth noting that there were no dyads with a one-year age gap. Hence, as mentioned earlier, only the twin's siblings were assessed for ToM. The twin's siblings and all the target children were in either grade 5 or 6 in their primary schools. Siblings from the small-gap dyads attended either primary or secondary schools, while those

from the large-gap dyads were either in high schools or universities. Furthermore, among the 30 sibling dyads, 16 dyads consisted of the same gender (i.e., girl-girl: 12 dyads; boy-boy: four dyads), while 14 dyads had different genders (i.e., boy-girl). Among the 30 parents, 19 of them had siblings and 11 of them were the only child at home.

During the online research session, 19 families utilised computers, while 11 families opted for tablets. All participants were blind to the research hypotheses. Although there was no payment for participation, each family had the option to receive a participation certificate featuring the logos of the University of Oxford and the Department of Education, along with the signatures of the principal investigator and the researcher.

3.4 Measures and Tasks

3.4.1 ToM Development

3.4.1.1 Strange Stories: Description

The current study utilised eight mental state stories from the Strange Stories task, which were selected by Fletcher et al. (1995) from the original 24 stories developed by Happe (1994). This study opted to use only a subset of the Strange Stories to minimise the testing burden on child participants, ensuring that their interest and focus were not compromised. This subset has been confirmed to be a reliable measure of individual differences in ToM across a wide age range including 5-13 years old children and even adults (Devine & Hughes, 2013; Wang et al., 2022; White et al., 2009).

The Strange Stories are short vignettes depicting social situations, followed by an open question that requires participants to explain the character's behaviours based on his or her mental states (Happe, 1994). The eight stories of this subset comprise four story types: Double Bluff, Persuasion, White Lie, and Misunderstanding. There are two stories for each type (see the full text in Appendix D). Crucially, Wang et al. (2022) validated a two-factor model of the Strange Stories task, wherein certain stories assess affective ToM while others assess cognitive ToM. They categorized White Lie and Persuasion as affective ToM stories, as these vignettes involve interpreting the

intentions of the utterance based on the characters' emotions; Double Bluff and Misunderstanding were classified as cognitive ToM stories, as they require understanding the thoughts and beliefs of the characters.

3.4.1.2 Adaptations to the Current Study

The researcher made cultural adaptations for the Chinese context by replacing English names with Chinese names (e.g., Peter to Xiao Ming) and adjusting scenarios to be more culturally relevant (e.g., from Christmas Day to one's birthday). A back-translation approach (Brislin, 1970) was then employed: each word was translated into Mandarin by the researcher, who is bilingual in English and Chinese, and subsequently back-translated into English using ChatGPT. The final English version was discussed with the supervisor to ensure equivalence with the original English version. The differences in meaning between the old and new English versions were very minimal and primarily involved synonym substitutions (see Table 1 for example). All present tense in the original version was changed to past tense in the back-translated version, likely due to the use of past tense markers (e.g., Le 了) in Chinese narration. The decision to use ChatGPT was based on its renowned language processing abilities (Jiao et al., 2023), and the researcher believed it would be a valuable endeavour to explore new techniques in the field of research.

Table 1

Synonym substitutions in back-translation in Story 1

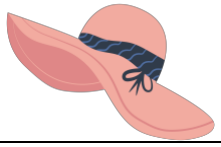
English Original Version	Chinese Translation	Back Translation
Very cross	Hen 3 sheng 1 qi 4 很生气	Very angry
I've looked	Zhao 3 guo 4 le 找过了	I've searched

In contrast to previous studies, this study included a "white question" before the mental-state open question for each story. The white question asked about simple factual details in the story (as illustrated in Table 2) and served two purposes: to assess children's basic comprehension and attention to the story and to boost their confidence and adaptation to the study.

Moreover, this study included an additional cognitive ToM story (Sarcasm) from the original set of Strange Stories as a practice trial. This decision was based on observations from the pilot study (see section 3.7), where it was noted that children tended to rely on the story sentences for their answers instead of offering their own interpretations required for the correct answer. By including the practice trial, the researcher could remind children who faced this challenge to carefully consider the story and provide their answers in their own words, helping them avoid making the same mistakes in the formal trials.

Table 2

Sample of the revised Strange Stories subset

Type of Stories	White Lie	
Type of ToM	Affective	
Content[a]	One day, aunt came to visit Xiao Ming. Xiao Ming loves his aunt very much, but today she was wearing a new hat that Xiao Ming thought was very ugly. Xiao Ming thought that his aunt looked silly wearing the new hat, and that she would look much prettier wearing her old one. But when Auntie asked Xiao Ming, "What do you think of my new hat?", Xiao Ming said, "Oh, it looks very nice".	
Cartoon illustration		
White Question	What did Xiao Ming consider to be ugly? (The new hat)	
ToM Question[b]	Why did he say that?	
ToM Coding Scheme[b]	2 points	Reference to white lie or wanting to spare her feelings; some implication that this is for aunt's benefit rather than just for his, desire to avoid rudeness or insult.
	1 point	Reference to trait (he's a nice boy) or relationship (he likes his aunt); purely motivational (so she won't shout at him) with no reference to aunt's thoughts or feelings; incomplete explanation (he's lying, he's pretending).
	0 point	Reference to irrelevant or incorrect facts/feelings (he likes the hat; he wants to trick her).

Note. [a] Back-translated version. [b] Sourced from the coding scheme by White et al. (2009).

During the online session, nine stories (i.e., one practice and eight formal trials) were presented to children using PowerPoint slides, with one story per page shared on the screen. To ensure consistency in reading time for each child and to control sensory clues received by children (e.g., tone), recordings of a Mandarin native speaker reading each story were included in the slides. The researcher played the recordings, and after each recording, the researcher asked the white question followed by the ToM question. Verbatim questions from a prepared document were used to maintain consistency. To create a relaxed environment and encourage dialogue, the researcher did not use recordings to ask the questions. Cartoon illustrations related to each story were included on the slides to maintain children's interest (see example in Table 2). Answers were provided for the practice question only, and if children asked for answers, the researcher explained them after completing all the stories.

3.4.1.3 Data Coding

This study utilised the coding scheme refined by White et al. (2009) (see Appendix D). The interrater reliability of this scheme has been reported as good (κ = .89) in the literature (White et al., 2009). Both the researcher and a research assistant independently rated each response to the mental-state question on a 0-2 three-point scale as children's ToM scores. A score of 2 indicated a full and accurate mental-state response, a score of 1 indicated a correct factual response or incomplete mental-state response, and a score of 0 indicated an incorrect factual or mental-state response. The research assistant was also a bilingual master's student in developmental psychology and was blind to all the research hypotheses.

During the data collection period, the researcher conducted a preliminary data coding on responses from ten participants. It was observed that for the mental-state question of the eighth formal story, some children responded with the idiom (i.e., Zuo 4 Zei 2 Xin 1 Xu 1 做贼心虚). This idiom metaphorically represents individuals who engage in wrongful acts and are constantly fearful of being discovered and uneasy. Consequently, the only modification made to the coding scheme was to specify that if a child simply answered using this idiom, coders should assign a score of 1, indicating an incomplete mental-state response. This decision was made because the full-mark answer for this story included a reference to the belief that the policeman knew about

the burglary, whereas the idiom only addressed the mental state of the main character (i.e., being afraid of being discovered) without mentioning the mental state of the second character (i.e., the policeman's knowledge). Finally, coders used the revised scheme to code the 30 target children and the other five twin participants.

The answers to the white questions, as they did not assess children's ToM abilities, were not included in their ToM scores. The answers to the practice trial were also not included. Each child received three scores: one cognitive ToM score (0-8), one affective ToM score (0-8), and one total ToM score (0-16). Two coders agreed on 246 out of 280 cases ($280 = 35 \times 8$) and accomplished a good inter-rater reliability, where $\kappa = .81$, $p < .001$ (Landis & Koch, 1977). Disagreements were resolved by discussion.

3.4.2 Self-Rated Sibling Relationship Quality

3.4.2.1 Questionnaire: Description

Target children and their siblings were asked to complete a new Chinese version of the Sibling Relationship Intimacy and Conflict Questionnaire (see Appendix E) adapted from Feng et al. (2019). The original Chinese version, created by Feng et al. (2019), combined and translated the Intimacy Scale of the Social Relations Questionnaire (Blyth et al., 1982; Blyth & Foster-Clark, 1987) and the Conflict and Antagonism Scale of the Network of Relationships Inventory: Behavioural Systems Version (Furman & Buhrmester, 2009) to evaluate sibling intimacy and conflict. Both the English and Chinese versions of the questionnaires have been commonly used to assess SRQ based on adolescents' perceptions (Chen et al., 2021; Waite et al., 2011). The original Chinese scales for sibling intimacy and conflict have demonstrated good internal consistency reliabilities, with Cronbach's alpha values of .76 and .78, respectively (Feng et al., 2019).

In the current questionnaire, sibling intimacy was assessed using 11 statement items, including "I often ask him/her for advice and help". Sibling conflict was assessed using seven statement items, including "We often quarrel". Participants rated each item on a five-point Likert-type scale (1 = strongly disagree to 5 = strongly agree) regarding the sibling who also participated in this research.

3.4.2.2 Adaptations to the Current Version

The original Chinese version of the questionnaire by Feng et al. (2019) included eight items for sibling intimacy and five items for sibling conflict. In this study, the questionnaire was improved in two ways. Firstly, statements (three items) such as "We often share feelings or secrets with each other" were separated into two statements: "He/she often shares his/her feelings or secrets with me" and "I often share my feelings or secrets with him/her" to remove ambiguity. Secondly, statements (two items) such as "He/she knows my interests and hobbies" were expanded to include the reciprocal perspective with the addition of the statement "I know his/her interests and hobbies". It is noteworthy that the English statements reported in this thesis are the final selected versions resulting from translations by the researcher and another MSc Education student.

The current questionnaire was created on Qualtrics Survey, providing a link for participants to access. Participants had the opportunity to ask any questions regarding the questionnaire. Based on the insights gained from the pilot study, the researcher observed that children under the age of 10 may have limited proficiency in operating digital devices. Therefore, for participants in this age group, the researcher shared the screen and assisted them in operating the questionnaire, while the children verbally provided their answers.

3.4.2.3 Scores

Each participant had self-rated scores for sibling intimacy and sibling conflict, calculated by averaging the ratings for the respective items (Chen et al., 2021). Additionally, each participant had an overall self-rated SRQ score, which involved reversing the scores of sibling conflict (i.e., 1 to 5, 2 to 4, 3 remained unchanged, 4 to 2, 5 to 1), and then adding intimacy scores and reversed conflict scores together (Kim & Hong, 2004). A higher overall score indicates a more positive SRQ.

3.4.3 Observational Sibling Relationship Quality

3.4.3.1 Etch-a-Sketch Online (ESO)

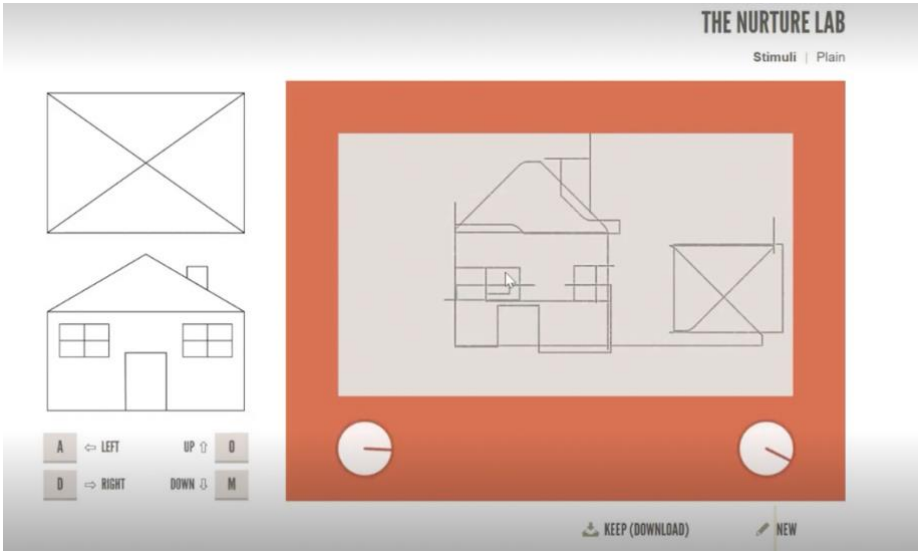
To assess SRQ through direct observation of sibling interactions, this study utilized the ESO (https://thenurturelab.itchaskitch.com/ESO), an online direct-observation tool generated by Oliver and Pike (2021). The ESO is a digital adaptation

of the traditional in-home Etch-a-Sketch drawing toy, where participants use keyboard keys (horizontal lines: A for left, D for right; vertical lines: O for up, M for down) or virtual arrows on a tablet screen to draw lines. Coding of ESO has demonstrated excellent interrater reliability (intraclass correlation (ICC) ranging from .83 to .95) and good convergent validity with the traditional Etch-a-Sketch (e.g., $r = .63$; Oliver & Pike, 2021).

The traditional Etch-a-Sketch task involved a mother playing with her child, but in this study, two siblings were asked to complete the drawing. Before explaining the game rules, the researcher asked the children to open the ESO link to check its functionality in their devices. If it did not work, they were asked to try using different devices. If it worked, the researcher introduced the game. One child was assigned to draw horizontal lines and the other to draw vertical lines. They were instructed not to touch each other's keys or arrows and to cooperate in copying two line-graphic pictures (see Figure 2) provided by the researcher. The task had a time limit of eight minutes, with the researcher informing the children when to start and stop. One challenge in the game was drawing diagonal lines, requiring the children to press the corresponding keys simultaneously. If children asked the researcher about this challenge, they were encouraged to figure it out on their own.

Figure 2

Example of a drawing work by participants



For data coding purposes, participants were instructed to keep their faces within the frame and speak at a normal volume. The video recording did not capture the children's drawing process or hand movements on the keyboard, as these were not relevant to the assessments. Once participants began drawing, the researcher turned off the camera and muted herself to minimize distractions.

3.4.3.2 Coding and Scores

This study applied the Positive Affective Climate (PAC) and Negative Affective Climate (NAC) coding scales developed by Waddell et al. (2001). The interrater reliability was reported as good (κ = .83 for PAC and .81 for NAC; Waddell et al., 2001). To the researcher's knowledge, Waddell et al.'s (2001) study is the only one that has used the Etch-a-Sketch task to assess the SRQ. The PAC and NAC coding scales include 11 and eight categories of behavioural indicators, respectively (see Appendix F). For instance, an indicator of PAC could be one sibling giving verbal compliments to the other, falling under the category of *Approval, Compliments, Praise*. An example of an NAC indicator could be one sibling giving orders to the other, belonging to the *Mild Limits* category.

Based on the preliminary coding of ten video recordings, the researcher introduced a new category called *Listening* to the PAC scale. This category encompasses behaviours that indicate attentive listening to the other person's ideas and even seeking their opinions. It also includes instances where one attempts to confirm their understanding of the other person's ideas. The researcher and the same research assistant independently used this modified PAC and the original NAC to code the final dataset of 28 recordings. One recording was missing because that family's devices were unable to open the link. Another recording was excluded from the coding due to low volume, making it challenging for the coders to make decisions, despite the researcher's request for louder voices in the online session.

Coders watched each recording before the actual coding to familiarise themselves with the content. During coding, each occurrence of an indicator was recorded, regardless of whether it had appeared before. After coding, coders counted the total number of the occurrence of the PAC and NAC indicators separately. The

recordings varied in length, with 20 families using the full eight minutes and eight families completing the task in less than eight minutes (ranging from 4.5 to 7.5 minutes). Therefore, this study assessed the observational SRQ by calculating the average number of PAC indicators per minute for positive SRQ and the average number of NAC indicators per minute for negative SRQ. Based on the frequency of the indicators, coders used a five-point scale from 1 to 5 to score both positive and negative SRQ for each sibling dyad, as shown in Table 3.

Table 3

The five-point scoring scale

Score	1	2	3	4	5
Criteria [a]	≤ 0.75	> 0.75 and ≤ 1.5	> 1.5 and ≤ 2.25	> 2.25 and ≤ 3	> 3

Note. [a] Unit of how many indicators per minute.

Based on the review during the preliminary coding, new coding rules were introduced. First, if an indicator (e.g., laughter) lasted for up to 5 seconds, it was coded once; if it exceeded 5 seconds, it was coded twice. Second, if both children simultaneously exhibited indicators, they were coded separately, even if their indicators were the same. Third, if the same indicator (e.g., giving orders) was used multiple times within one event (e.g., instructing the other to draw a specific line), such as "draw the line to the left..., continue..., now stop," it was coded once. However, if different indicators were used within the same event, they were coded separately.

Due to time constraints, the assistant only coded recordings for 16 randomly selected families. These families were chosen using 16 participant IDs generated by R studio. Before formal coding, the assistant received training from the researcher and practiced by coding two recordings not included in the 16. For the formal coding, coders reached a very good agreement for PAC scores, with an *ICC* of .86 ($p < .001$), and an acceptable agreement for NAC scores, with an *ICC* of .62 ($p < .01$). The researcher's codes were used for data analysis.

3.4.4 Parenting Styles

3.4.4.1 Interview

The researcher conducted interviews with parents based on six scenarios of daily social incidents, as outlined in Table 4, and posed three questions to the parents: (1) “Does this scenario apply to (the name of the target child)?”, (2) “Can you briefly describe what happened?”, and (3) “What did you say or do to the child to resolve this?”. Parents were asked to report one case for each scenario. They were also informed that if a scenario did not apply to their child, they could inform the researcher and proceed to the next one. Confidentiality was emphasized, and parents were encouraged to provide truthful responses without concerns about access to the information by the school or other parties.

Table 4

Scenarios presented to parents

1	A recent incident in which your child quarrelled or fought with another peer, for example, with a classmate.
2	A recent situation when your child shouted at, talked back to, made fun of an elder family member.
3	A recent case when your child damaged something that did not belong to him/her.
4	An instance where you took your child to a party, and he/she did not behave as you expected.
5	A recent incident in which your child disagreed with you.
6	A recent situation when your child quarrelled or fought with his/her sibling.

The first five scenarios were adapted from the six scenarios of Shahaeian et al.’s (2014) study (see Appendix G), which were originally derived from Ruffman et al. (1999). The sixth scenario was specifically created by the researcher with reference to Kramer et al.’s (1999) study. The reason for utilizing only five scenarios from Shahaeian et al.’s (2014) paper was due to considerations in Chinese culture, where children shouting at parents (the second scenario in Shahaeian et al.’s study) is encompassed within the concept of children acting impolitely (the sixth scenario in Shahaeian et al.’s study). To avoid redundancy in responses, this study modified the sixth scenario to one that is highly relevant to the research topic, focusing on sibling interactions. Some adaptations were made to the first five scenarios (see Appendix G). For instance, given that many Chinese families include grandparents in addition to

parents and children, this study expanded the scope of shouting incidents in the second scenario to include any elder family members, not just parents.

3.4.4.2 Coding scheme

Initially, Ruffman et al. (1999) used four categories of parenting strategies for coding responses: (1) *General Discussion*: when parents reported explaining and exploring the situation to the child without referencing others' feelings; (2) *Reprimand*: parents disciplined the child; (3) *How Feel*: parents discussed the situation while mentioning others' feelings; and (4) *Ambiguous*: parents exhibited a combination of both Reprimand and General Discussion approaches.

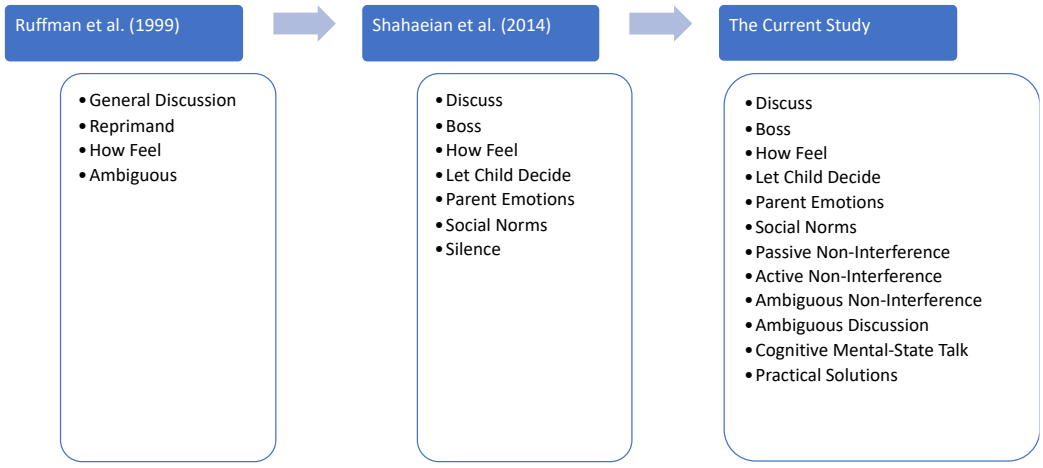
Shahaeian et al. (2014) then developed a new coding scheme (Appendix H) based on the work of Ruffman et al. (1999) and reported good interrater reliability (κ = .84). Firstly, Shahaeian et al. (2014) provided more detailed descriptions for the category previously known as *General Discussion*, which was renamed *Discuss*. It encompassed parents explaining to the child why a particular behaviour is inappropriate or informing them about the consequences of certain actions. Additionally, the *Reprimand* category was renamed *Boss* and expanded to include situations in which parents made decisions for the child, imposed punishments, or exerted control. The key distinction between *Discuss* and *Boss* is that *Boss* parents do not allow children the opportunity to defend themselves or make decisions, whereas *Discuss* parents listen to children's explanations and respect their decisions.

The *How Feel* category remained unchanged. However, they abandoned the category of *Ambiguous* and introduced four new categories: *Let Child Decide, Parent Emotions, Silence*, and *Social Norms*. *Let Child Decide* referred to situations where parents, without any discussion, directly left the decision-making to the child. *Parent Emotions* encompassed responses involving any expression of parental feelings towards the child. *Silence* covered instances where parents adopted a passive approach, refraining from challenging the child's misbehaviour and ignoring the situation. Lastly, *Social Norms* described situations in which parents referred to the importance of adhering to socially accepted norms and the judgments of others regarding certain behaviours.

In the current study, the categories of *Discuss, Boss, How Feel, Let Child Decide, Parent Emotions*, and *Social Norms* were retained without any modifications. However, as indicated in Figure 3, *Silence* was divided into three separate categories (*Active, Passive, Ambiguous Non-Interference*), and three new categories were added (*Ambiguous Discussion, Cognitive MST, Practical Solutions*).

Figure 3

Variations of parenting categories



Firstly, the researcher noticed that in addition to the passive avoidance and silence described by the *Silence* category, parents may also engage in active avoidance and silence, such as believing that their children possess the abilities to resolve issues on their own. To capture this distinction, a new category called *Active Non-Interference* was introduced. The original *Silence* category was renamed *Passive Non-Interference*. Furthermore, an *Ambiguous Non-Interference* category was created for situations where it was challenging to determine whether the parents' lack of involvement was active or passive based on their responses.

Secondly, the researcher observed that parents sometimes referred to the cognitive mental states of individuals involved in the scenarios (e.g., misunderstandings) rather than their affective mental states (e.g., feelings). The existing category of *How Feel* and *Parent Emotions* were intended to capture the latter

situation. The former type of situations did not fit into any of the original categories, leading to the creation of a new category called *Cognitive MST*. Additionally, instances where parents provided explanations to the child but did not clearly fit into any specific category were classified as *Ambiguous Discussion*. Lastly, the *Practical Solutions* category was introduced to capture situations where parents took direct actions to resolve social incidents instead of engaging in discussions or remaining silent.

One category was assigned to each scenario for each parent. In cases where parents applied multiple parenting styles to a single scenario, such as stating, "I first said/did... and then I said/did..." or "It depends. Normally I said/did..., but sometimes I said/did...", the scenario was coded for more than one category. Nevertheless, if parents mentioned changes in parenting styles as children grow up, for instance, stating "I previously said/did..., but now I said/did...", only the current parenting strategies were considered for coding.

3.4.4.3 New Coding Categories

Passive Non-Interference: This category pertained to situations in which parents acknowledged that their children required assistance with social incidents but chose not to participate or interfere. Instead, they opted to ignore or avoid involvement. This category included responses indicating that parents wished for someone else to handle the problem. It also encompassed responses suggesting that parents believed they lacked the necessary skills, time, patience, or appropriate emotional states (e.g., mood) to address the incidents with children. Instances where parents used silence or indifference as a form of punishment were also included. Examples of such responses included: "I pushed her dad to talk to her", "I preferred not to intervene because this occurred daily, and I had been exhausted by it", and "I chose to give her the cold shoulder so she could reflect".

Active Non-Interference: This category described the situations where parents remained silence due to their beliefs that their involvement was not necessarily required in the incidents. This may have stemmed from their trust in their children's abilities to resolve the issues on their own. It could also have been because parents perceived the situation as not warranting immediate intervention or deemed it not serious enough to warrant their involvement. For example, a mother reported that based on previous

experiences, she refrained from getting involved in her siblings' conflicts, recognising that children could solve problems without parental interference and that parental involvement exacerbated the situation. Notably, if the parent instructed the children to resolve the issue on their own, this would fall under the *Boss* category.

Ambiguous Non-Interference: This category was used when parents expressed non-involvement in the incidents without providing specific explanations for their stance. Examples of this code included statements such as "I just left the room".

Cognitive MST: This category represented instances where parents engaged in discussions with their children, focusing on the cognitive mental states such as thoughts and intentions of people in the scenario. This code could involve comparing the mental states of different individuals or discussing how a person's mental states may change over time. This also involved discussions of whether a person's behaviour was intentional or unintentional. For example, a parent shared, "I explained to her the incorrect thoughts I had, which led to the misunderstanding".

Ambiguous Discussion: This code included responses in which parents indicated communication with their children but provided limited details about the content of the conversation. For instance, a mother simply mentioned saying, "I reasoned with him right away". In such cases, the researcher lacked specific information about the reasoning approach used (e.g., discussing the consequences of certain behaviours, considering others' mental states, or referring to social norms).

Practical Solutions: This category encompassed situations where parents took direct actions to resolve incidents without engaging in conversations with children. Crucially, these actions were not disciplinary in nature. For example, a mother reported resolving her children's conflict over a toy by purchasing an additional toy for each child.

3.4.4.4 Coding procedure

During the data collection period, the researcher initiated preliminary coding of the existing 20 interview audio recordings using the original coding scheme of Shahaeian et al. (2014). New coding categories were created to accommodate instances that did not fit existing categories. The researcher presented the new coding scheme in a dissertation group meeting, incorporating feedback to further improve it.

Subsequently, the revised coding scheme was used to code the entire dataset of 30 recordings.

In line with Ruffman et al. (1999) and Shahaeian et al. (2014), each parent's scores were determined based on the proportion of their responses within each category. For example, if a parent reported four scenarios out of six, including one Discuss, one Parent Emotions, and two Boss responses, the parent would receive 25% for Discuss, 25% for Parent Emotions, 50% for Boss, and 0% for the remaining categories. In cases where a parent reported four scenarios with a total of five responses (e.g., one Discuss and four Boss responses, with one scenario containing both categories), each response carried equal weight. Therefore, the parent would receive a score of 20% for Discuss, 80% for Boss, and 0% for the other categories.

The same research assistant independently coded the parenting. For both coders, the coding process began one week after coding the ToM responses and one week before rating the observational SRQ. The one-week interval was implemented to minimise the potential influence of previously coded data on current coding. Due to time constraints, the assistant only coded interviews for ten parents, randomly selected using ten participant IDs generated by R studio. Prior to formal coding, the assistant received training from the researcher and coded three interviews outside of those ten together with the researcher as practice.

Among ten interviews, the coders reached agreement on 48 out of 60 cases ($\kappa = .76$, $p < .001$), and any disagreements were resolved through discussion.

3.4.5 Demographic Information

Parents were asked demographic questions, including: (1) the age, sex, and current education level of their two children (after the researcher collected consent forms and before scheduling appointments with parents), and (2) the number of children in their family and whether the parent has any siblings (during the interviews conducted with parents in the online research session).

3.6 Ethics

The Central University Research Ethics Committee (CUREC) has granted approval for this study: the research ethics reference number is EDUC_C1A_23_173 (see Appendix L).

The present study handles highly identifiable data, including video recordings. To ensure data protection, all data is securely stored in the Nexus 365 OneDrive project folder, accessible only to the researcher and the research assistant. Prior to gaining access to the folder, the assistant signed a confidentiality agreement.

During the recruitment process, the headteachers assisted in mobilising students and parents to participate in the research. However, to ensure the prevention of any form of power abuse, the teachers were explicitly informed not to actively persuade or coerce families into participating. Lastly, to enhance recruitment, the researcher used a personal WeChat account for contacting parents, considering its prevalence in China than emails and mobile phone text messages. To safeguard the researcher's privacy, the WeChat account was adjusted to "chats only" mode, concealing all personal information from the participants.

3.7 Pilot Study

The pilot study, involving three families (see Table I1 for demographic information), aimed to assess the feasibility of the original research design and provide the researcher with practice for the procedures and instructions. Due to the pilot families not meeting the precise age requirements and time limitations for the thesis, no data analysis was conducted on the pilot study data.

3.8 Data Analytic Plan

Table 5 presents all the variables used for data analysis. Notably, birth order contained three categories: the target child being the first born, second born, or part of a set of twins. Age gap also contained three categories: zero, small, and large (as previously mentioned in section 3.3). Gender composition had two categories: same-gender and different-gender. Furthermore, observed positive and negative SRQ were interval-scale variables. This was because there were equal intervals (0.75) between

consecutive data points, and the zero point on their scale hold no inherent meaning (Cohen et al., 2018). Moreover, the age and number of siblings were treated as categorical variables because both variables have only two levels (age: 11 and 12 years; number of siblings: two and three children at home).

Table 5

Lists of variables and their data type

Demographic information				ToM		SRQ		Parenting styles	
Target child	Type	Sibling structure	Type		Type		Type		Type
Age	nominal	Number of siblings	nominal	Cognitive score	ratio	Self-rated[a] SI	ratio	Discuss	ratio
Gender	nominal	Birth order	nominal	Emotional score	ratio	Self-rated[a] SC	ratio	How feel	ratio
		Age gap	nominal	Total score	ratio	Self-rated[a] total SRQ	ratio	Parent emotions	ratio
		Gender composition	nominal			Sibling-rated SI	ratio	Cognitive MST	ratio
						Sibling-rated SC	ratio	Social norms	ratio
						Sibling-rated total SRQ	ratio	Ambiguous discussion	ratio
						Observed positive SRQ	interval	Passive NI	ratio
						Observed negative SRQ	interval	Active NI	ratio
								Ambiguous NI	ratio
								Boss	ratio
								Let child decide	ratio
								Practical solutions	ratio

Note. NI = Non-interference. SI = Sibling intimacy. SC = Sibling conflict. [a]rated by target children.

Based on the data types, the researcher planned to conduct a Pearson's product-moment correlation analysis to test whether the SRQ-ToM, parenting-ToM, and parenting-SRQ associations are significant. The researcher intended to use a point-biserial correlation analysis to examine whether the target children's age, gender, birth order, number of siblings, sibling age gap, and sibling gender composition are significantly correlated with ToM scores. If significant relationships between demographic variables and ToM are found, the researcher planned to conduct a Pearson's partial correlation analysis to control for these demographic variables when testing the sibling-ToM and parenting-ToM associations again.

Additionally, the researcher recoded birth order and age gap, each containing three categories, into dichotomous (dummy) variables, as required for point-biserial

correlation analysis. For age gap, small gap was the reference group, being the largest in the sample (Field, 2018). For the first dummy variable, “large gap”, large was coded as 1 and the rest as 0. For the second dummy variable, “zero gap”, zero was coded as 1 and the others were coded as 0. Similarly, for birth order, the first-born category was used as the reference group for the same reason. For the first dummy variable, “second born”, second born was coded as 1, and 0 otherwise. For the second dummy variable, “twin”, twin was coded as 1, and 0 otherwise.

However, for continuous variables, the emotional ToM scores, sibling-rated sibling conflict, and all parenting variables did not pass the normality checks, while the others did (see Appendix J). Because the distribution of emotional ToM was only moderately negatively skewed, as illustrated in Figure J1, the researcher applied a “reflect and square root” transformation ($\sqrt{(Max + 1) - data}$) to the emotional ToM data (Fox, 2016). This also suggests that the emotional Strange Stories tasks exhibited a ceiling effect. To convert the moderately positively skewed data to normality (Figure J2), the researcher applied a “square root” transformation ($\sqrt{data}$) to the sibling-rated sibling conflict data (Fox, 2016). After these transformations, both variables passed the normality checks. However, one extreme outlier was removed from the transformed emotional ToM scores (Osborne & Overbay, 2008).

All the distributions of the parenting variables were strongly positively skewed (Figure J3), suggesting that these variables should undergo a log transformation (Fox, 2016). However, all the parenting variables contain zero data points, making log transformations impossible. Hence, the researcher planned to conduct a Spearman’s rank-order correlation analysis to examine the parenting-ToM links and parenting-SRQ links.

Notably, although the researcher collected ToM scores from both children in twin families, only the twin who participated in the study first was included in the sample for all the above-mentioned correlation analyses (i.e., the main analysis; see section 4.2). Thus, the sample included 30 participants instead of 35. Following these analyses, due to time constraints, the researcher planned to conduct a separate analysis on twins, focusing solely on SRQ-ToM and parenting-ToM associations (see section 4.4).

Chapter 4: Results

4.1 Descriptive Statistics

This section presents descriptive data for measures and variables.

Firstly, regarding the Strange Stories tasks, the most challenging task for cognitive ToM was the first story of Double Bluff, with the lowest average score of 0.8 out of 2. In contrast, the easiest was the eighth story, with an average score of 1.43. For emotional ToM, the most challenging task was the fourth story (average score 0.89), and the easiest was the sixth story (average score 1.69). As shown in Table 6, the average cognitive ToM score for young adolescents (M = 4.73, SD = .31) was lower than the average emotional ToM score (M = 5.13, SD = .29). This led the researcher to conduct a Wilcoxon signed rank test to examine whether this difference was significant (see section 4.3.1).

In the sibling intimacy questionnaire, the highest-rated item was "He/she is important to me" (average 4.40), and the lowest was "I wish to be like him/her" (average 3.12). In the sibling conflict questionnaire, "I often disagree with him/her" had the highest average rating (2.93), while "I often get impatient with him/her" had the lowest (2.15). As indicated in Table 6, the means of self-rated[2] sibling intimacy, conflict, and total SRQ scores were very close to the means of sibling-rated sibling intimacy, conflict, and total SRQ scores, respectively. This prompted the researcher to investigate the correlations between self-rated and sibling-rated SRQ variables (see section 4.3.2).

Regarding parenting styles, 30 parents reported a total of 116 cases/scenarios, with an average of 3.87 per parent. The sixth scenario, concerning sibling conflicts, was the most frequently reported, with all parents providing cases for this scenario. Conversely, the fourth scenario, about behaviours at parties, was the least reported, with only 8 parents providing cases. As indicated in Table 7, *Boss* was the most frequently reported, with 53.3% of parents using this strategy at least once. *Discuss* was also commonly used by parents (46.7%), while *Let Child Decide* was the least used (10.0%).

[2] From this point forward, "self-rated SRQ" refers to the SRQ as rated by the target children.

Among the strategies used, *Boss* had the highest average proportion per parent (*M* = 18.4%, *SD* = 3.5%), while *Let Child Decide* had the lowest (*M* = 2.3%, *SD* = 1.4%).

Table 6

Descriptive data for ToM variables and SRQ variables

Variable	Mean (SD)	N	Maximum	Minimum	Skewness	Kurtosis
Cognitive ToM	4.73(0.31)	30	8	2	0.41	-0.42
Emotional ToM[a]	5.13(0.29)	30	8	0	-1.17	2.86
Total ToM	9.87(0.47)	30	15	3	-0.21	0.46
Self-rated SI	3.75(0.14)	30	4.91	2	-0.58	-0.37
Self-rated SC	2.39(0.17)	30	5	1	0.50	0.36
Self-rated SRQ	66.57(2.25)	30	88	42	-0.25	-0.81
Sibling-rated SI	3.70(0.12)	30	4.64	1.91	-0.65	0.89
Sibling-rated SC[a]	2.50(0.18)	30	4.86	1.14	0.95	0.21
Sibling-rated SRQ	65.20(1.89)	30	79	41	-0.54	-0.32
O positive SRQ	2.71(1.24)	28	5	1	0.09	-0.95
O negative SRQ	2.82(1.42)	28	5	1	0.34	-1.27

Note. SI = Sibling intimacy. SC = Sibling conflict. O = Observed. [a] Before transformation.

Table 7

Descriptive data for parenting variables

Parenting variables	Percentage[a]	Mean[b]	Standard deviation	Maximum
Discuss	46.7%	16.6%	3.9%	75.0%
How feel	16.7%	4.2%	2.0%	50.0%
Parent emotions	16.7%	3.8%	1.7%	33.3%
Cognitive MST	13.3%	3.1%	1.5%	33.3%
Social norms	20.0%	5.2%	2.1%	40.0%
Ambiguous discussion	36.7%	16.2%	4.5%	66.7%
Passive NI	16.7%	4.1%	1.8%	40.0%
Active NI	30.0%	7.8%	2.5%	50.0%
Ambiguous NI	40.0%	15.1%	4.5%	100.0%
Boss	53.3%	18.4%	3.5%	50.0%
Let child decide	10.0%	2.3%	1.4%	33.3%
Practical solutions	13.3%	3.1%	1.5%	33.3%

Note. NI = Non-interference. [a]Interparent frequency of strategy use = $(how\ many\ parents\ reported\ each\ category\ at\ least\ once)/30$. [b] Intraparent frequency of strategy use.

4.2 Testing the Hypotheses: Correlation Analysis

4.2.1 Demographic Variables and ToM

A two-tailed point-biserial correlation analysis was conducted to test the correlations between each demographic variable and each type of ToM score. As

demonstrated in Table 8, all the correlations were very weak (around .10) and not statistically significant. Hence, the non-significant p-values led the researcher to fail to reject the null hypothesis that there were no relationships between any demographic variables and ToM (H_0: $r = 0$). The results suggest that ToM scores did not differ based on the target child's age (11 or 12 years), gender, sibling dyad sex (same-sex or different-sex), family size (two-child or three-child), sibling age gap (zero, small, large), or birth order. This eliminates the need for a partial correlation analysis to control for these demographic variables.

Table 8

Correlations between ToM scores and demographic variables

ToM	Age	Sex	Sex-C	NoS	Second Born	Twin	Large gap	Zero gap
Cognitive	.09	.23	-.19	.18	-.17	-.14	-.07	-.14
Emotional[a]	.01	-.08	-.03	.13	.03	.07	-.07	.07
Total	.13	.11	-.18	.10	-.07	-.29	.03	-.29

Note. Sex-C = Gender composition. NoS = Number of siblings. [a]After transformation. $^{*}p < .05$.

4.2.2 SRQ-ToM Associations

A two-tailed Pearson's correlation analysis was performed to test the relationships between each SRQ and each ToM variable. Among all the correlations in the main analysis, only the association between the target children's self-rated sibling intimacy and their cognitive ToM performance proved statistically significant, $r(28) = -.42$, $p < .05$, as indicated in Table 9. This led the researcher to reject the null hypothesis that self-rated sibling intimacy was not related to cognitive ToM. As further illustrated in Figure 4, this was a moderate negative relationship, which implies that target children who rated their SRQ as warmer tended to score lower in cognitive ToM tasks.

Table 9

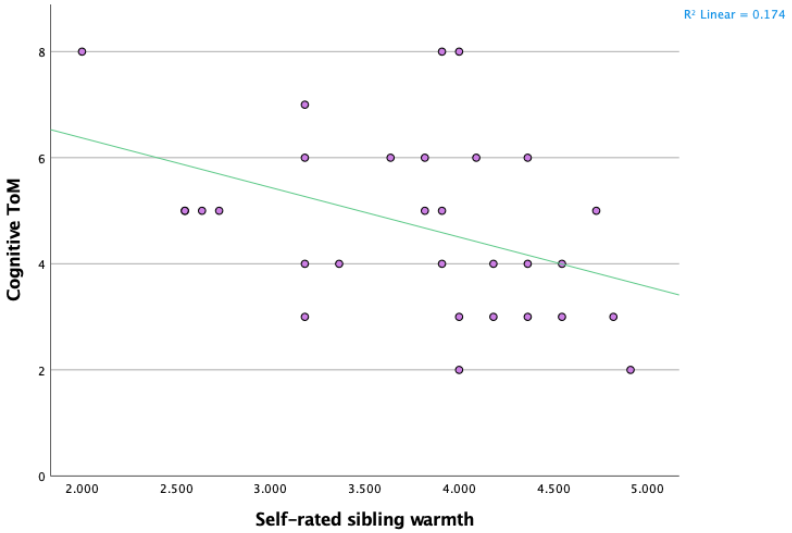
Correlations between ToM scores and SRQ variables

	Self-rated			Sibling-rated			Researcher-observed	
Main Analysis								
ToM	SI	SC	Total	SI	SC[a]	Total	Positive SRQ	Negative SRQ
Cognitive	-.42*	.00	-.28	.05	.08	-.03	.09	.15
Emotional[a]	-.36	.21	-.35	-.13	.12	-.17	-.18	-.20
Total	-.16	-.02	-.10	.03	.05	-.02	.15	.22
Sub-Analysis								
ToM[b]	SI	SC	Total	SI	SC	Total	Positive SRQ	Negative SRQ
Cognitive	-.51	.11	-.40	.19	-.24	.23	-.91	.23
Emotional	.03	.22	-.07	.15	.41	.01	-.60	.79
Total	-.13	.21	-.18	.18	.26	.08	-.82	.74
Sibling ToM[c]								
Cognitive	-.90*	.43	-.81	-.37	-.12	-.28	-.96*	.38
Emotional	.51	.05	.33	-.12	.57	-.26	.43	.32
Total	-.88	.52	-.83	-.48	.07	-.43	-.98*	.57

Note. SI = Sibling intimacy. SC = Sibling conflict. [a]After transformation. [b]Twin as target children. [c]Twin as siblings. *$p < .05$.

Figure 4

Scatterplot



4.2.3 Parenting-ToM Associations

A two-tailed Spearman's correlation analysis was conducted to examine the relationships between each parenting variable and each ToM variable. A statistically significant positive correlation was found between the proportion of *Parent Emotions* use and cognitive ToM (but not emotional and total ToM), where $r(28) = .44$, $p < .05$, as indicated in Table 10. This reveals that children whose parents used the *Parent Emotions* strategy more frequently tended to obtain higher cognitive ToM scores.

Table 10

Correlations of ToM, SRQ, and parental sibling status with parenting variables

		D	HF	PE	CD	SN	AD	PNI	AcNI	ANI	B	LCD	PS
Main Analysis													
1	Cognitive	.02	-.02	.44*	.08	-.07	-.04	-.07	-.13	-.10	.11	-.15	.02
	Emotional[a]	-.20	-.24	.07	-.08	-.11	.05	.30	-.40*	.28	.03	.13	.37*
	Total	.12	.14	.30	.09	-.08	-.02	-.17	.17	-.23	.03	-.11	-.21
2	SI	.35	.08	-.16	-.12	-.04	-.08	-.23	-.19	-.18	-.14	.04	-.15
	SC	-.09	.11	-.02	-.21	-.05	.23	.33	.32	.17	.12	.18	-.12
	Total	.27	-.02	-.09	.04	-.01	.06	-.33	-.03	-.22	-.16	-.05	-.03
3	SI	.16	-.02	.27	.08	-.02	.34	-.25	-.06	-.42*	-.19	-.27	-.13
	SC[a]	.13	-.14	-.16	-.26	-.05	-.06	.14	-.13	-.11	.21	.06	-.15
	Total	-.04	.05	.22	.25	.02	.30	-.22	.18	-.19	-.26	-.34	-.01
4	Positive	.28	.08	-.27	.20	.02	-.19	-.26	-.34	-.15	.08	-.12	.03
	Negative	.02	.09	.07	.02	-.19	-.30	.15	.49**	.10	-.06	.29	.22
5	Sibling status	.24	.02	.38*	-.09	-.06	-.29	-.13	.26	-.06	.09	-.02	.14
Sub-analysis													
6	Cognitive	-.73	.54	.54		-.60	.19		-.05	.86	-.92*	-.36	.54
	Emotional	.00	.00	.00		.36	-.39		.42	-.50	.50	.79	-.79
	Total	-.73	.73	.73		-.58	.09		.09	.92*	-.80	-.18	.36
7	Cognitive	-.61	.41	.41		-.76	.55		.33	.65	-.97**	-.61	.41
	Emotional	-.71	.35	.35		-.57	.33		.92*	.11	-.67	.00	-.35
	Total	-.71	.35	.35		-.70	.44		.86	.11	-.67	.00	-.35

Note. SI = Sibling intimacy. SC = Sibling conflict. D = Discuss. HF = How feel. PE = Parent emotions. CD = Cognitive MST. SN = Social norms. AD = Ambiguous discussion. PNI = Passive non-interference. AcNI = Active non-interference. ANI = Ambiguous non-interference. B = Boss. LCD = Let child decide. PS = Practical solutions. 1 = ToM scores. 2 = Self-rated. 3 = Sibling-rated. 4 = Observed SRQ. 5 = Parental. 6 = Sibling twins' ToM scores. 7 = Target-child twins' ToM. [a]After transformation. *$p < .05$. **$p < .01$.

Additionally, the proportion of *Active Non-interference* use was significantly negatively related to transformed emotional ToM, where $r(27) = -.40$, $p < .05$. Notably, due to the previously introduced reverse transformation of emotional ToM, this link suggests that children whose parents used *Active Non-interference* more frequently performed better in emotional ToM tasks. Moreover, the proportion of *Practical*

Solutions use was significantly positively correlated with transformed emotional ToM, where $r(27) = .37$, $p < .05$. This implies that children whose parents applied *Practical Solutions* more frequently scored lower in emotional ToM tasks. Nevertheless, all other relationships between parenting styles and ToM, as well as all the associations between different parenting style variables, were not statistically significant.

4.2.4 Parenting-SRQ Associations

A two-tailed Spearman's correlation analysis was conducted to verify the associations between each parenting variable and each SRQ variable. Only two statistically significant relationships were found, as demonstrated in Table 10. Firstly, sibling-rated sibling intimacy was significantly negatively related to the use of *Ambiguous Non-interference* by parents, $r(28) = -.42$, $p < .05$, suggesting that children whose parents used *Ambiguous Non-interference* more frequently tended to have less warm relationships with their siblings. Secondly, researcher-observed negative SRQ was significantly positively associated with the use of *Active Non-interference* by parents, $r(26) = .49$, $p < .05$. This indicates that children were inclined to have more negative SRQ when their parents employed *Active Non-interference* parenting more frequently.

4.3 Beyond Hypotheses

4.3.1 Cognitive ToM versus Emotional ToM

A Wilcoxon signed rank test showed that young adolescents' emotional ToM scores were not significantly different from their cognitive ToM scores, $z = 1.19$, $p = .23$.

4.3.2 Psychometric Properties of SRQ Measures

As shown in Table 11, the two-tailed Pearson's correlation tests revealed that, firstly, the target children's self-rated sibling conflict was significantly positively related to the sibling conflict scores reported by their siblings, $r(28) = .43$, $p < .05$. However, the self-rated sibling intimacy scores were not significantly correlated with sibling-rated sibling intimacy, $r(28) = .18$, $p = .35$. These results suggest that only the sibling conflict component in the current SRQ questionnaire demonstrated moderate

levels of inter-rater reliability, indicating that the two siblings' perceptions of their SRQ reached a moderate level of agreement.

Secondly, among the correlations between survey-rated SRQ and observational SRQ, only self-rated sibling conflict scores were significantly negatively related to observed positive SRQ scores, $r(26) = -.38$, $p < .05$. Thus, there is not enough evidence to confirm that either the current sibling conflict questionnaire or the Etch-a-Sketch observation demonstrates good concurrent validity.

Table 11

Correlations between SRQ variables

		1	2	3	4	5	6	7	8
Self-rated	1 SI		-.37*	.87**	.18	-.26	.29	.34	-.05
	2 SC			-.79**	.08	.43*	-.22	-.38*	.21
	3 Total				.08	-.40*	.31	.44*	-.15
Sibling-rated	4 SI					-.10	.76**	-.01	-.21
	5 SC[a]						-.72**	.04	-.18
	6 Total							-.02	-.04
Observed SRQ	7 Positive								-.20
	8 Negative								

Note. SI = Sibling intimacy. SC = Sibling conflict. [a]After transformation. *$p < .05$. **$p < .01$.

4.3.3 Does Self-rated Sibling Intimacy Predict Cognitive ToM?

Since the association between self-rated sibling intimacy and cognitive ToM was the only significant one among all the SRQ-ToM relationships, and *Parent Emotions* use was the only parenting variable related to cognitive ToM, a multiple regression was run to predict cognitive ToM from self-rated sibling intimacy and *Parent Emotions* use. The assumptions of independence of residuals, homoscedasticity, multicollinearity, normality of dependent variable, and linearity were met (see Appendix K).

Self-rated sibling intimacy and *Parent Emotions* use statistically significantly predicted cognitive ToM scores, $F(2, 27) = 5.74$, $p < .01$, adjusted $R^2 = .25$. Both predictors added statistically significantly to the prediction; that is, the slope coefficients of both independent variables were statistically significant ($p < .05$, as illustrated in Table 12). Cognitive ToM scores decreased by .87 units for each one-unit

increase in self-rated sibling intimacy. The results also indicate that 30% of the variability in cognitive ToM scores can be explained by the linear associations with self-rated sibling intimacy and *Parent Emotions* use.

Table 12

Multiple regression results

	B	SE B	β	t	Sig.	95% CI for B		R^2	Adj. R^2
						LB	UB		
Model								.30	.25
Constant	7.76	1.40		5.55	< .001	4.89	10.64		
Self-rated SI	-.87	.36	-.39	-.241	.02	-1.62	-.13		
Parent emotions	6.57	3.00	.35	2.19	.04	.41	12.73		

Note. SI = Sibling intimacy. B = Unstandardised regression coefficient. SE B = Standard error of the coefficient. β = Standardised coefficient. R^2 = Coefficient of determination. Adj. R^2 = Adjusted R^2. LB = Lower bound. UB = Upper bound.

4.3.4 Was Parental Sibling Status Related to Their Parenting Styles?

A two-tailed Spearman's correlation analysis was performed to test the associations between parental sibling status and each parenting variable, with lower coding for parents who had siblings and higher coding for parents who were only children. As shown in Table 10, among all the parenting variables, only *Parent Emotions* was significantly related to parental sibling status, $r(28) = .38$, $p < .05$. This positive relation suggests that parents who were only children tended to use the *Parent Emotions* strategy more frequently than parents who had siblings.

4.4 Sub-analysis for Twins

The sample size for the twin analysis was five. This is because, instead of including the other five twins as new target children in this sample, the researcher added the ToM scores of those twins (as siblings) as three new variables (i.e., sibling's cognitive ToM, sibling's emotional ToM, sibling's total ToM) for each existing twin as target child. In this new sample, the ToM variables, SRQ variables, *Social Norms, Ambiguous Discussion*, and *Active Non-interference* passed the normality checks. Hence, the researcher ran two-tailed Pearson's correlation tests for the correlations of

these SRQ and parenting variables with ToM scores. Notably, the parents from these twin families did not report any conduct in *Cognitive MST* and *Passive Non-interference* strategies. Moreover, the rest of the parenting variables violated the normality assumptions. Therefore, two-tailed Spearman's correlation tests were performed to explore the relationships between these parenting variables and the ToM scores.

The results indicate that all types of ToM scores for twins as target children were not significantly related to any types of ToM scores for twins as siblings. Regarding the SRQ-ToM links, as indicated in Table 9, the ToM scores for sibling twins were not related to any SRQ scores rated by themselves. However, the cognitive ToM of sibling twins was significantly related to sibling intimacy rated by their target-child twins, $r(3) = -.90$, $p < .05$. Moreover, the cognitive ToM and total ToM scores for sibling twins were significantly linked with observed positive SRQ, $r(2) = -.96$, $p < .05$, and $r(2) = -.98$, $p < .05$. Notably, these SRQ-ToM relations were negative, which is consistent with findings from the previous larger sample (see section 4.2.2). However, the ToM scores for twins as target children were not significantly related to any SRQ variables.

Concerning the parenting-ToM relations, as indicated in Table 10, the results reveal that the cognitive ToM scores for sibling twins were significantly correlated with *Boss* use by parents, $r(3) = -.92$, $p < .05$. Crucially, the cognitive ToM scores for target-child twins also showed a negative relationship with *Boss*, $r(3) = -.97$, $p < .01$. These strong relations suggest that the *Boss* strategy might be particularly detrimental to twins' ToM development. Additionally, the total ToM scores for sibling twins were related to *Ambiguous Non-interference*, $r(3) = .92$, $p < .05$. However, twins as target children did not show the same pattern for this relation, $r(3) = .11$, $p = .86$. They indicated the significant positive relation for Active Non-interference, $r(3) = .92$, $p < .05$.

Chapter 5: Discussion

5.1 Interpretations of the Main Findings

5.1.1 Negative Relations Between Positive SRQ and Cognitive ToM

5.1.1.1 Why Negative Relations?

The primary aim of this dissertation was to conduct a multimethod, multi-informant investigation into the relationship between Chinese young adolescents' SRQ and their ToM. The current study identified only significant negative relations between positive SRQ and ToM scores in both the main analysis and sub-analysis. This finding contrasts with the positive association identified in previous Western research and the sole Chinese study by Hou et al. (2022). One possible explanation for this discrepancy is that earlier studies primarily focused on preschoolers. For these younger children, better sibling relationships might serve as a secure attachment base, allowing them to explore social interactions and develop foundational social skills (Samuels, 1980). However, young adolescents require more sophisticated social skills, which are cultivated through more ToM-intensive social interactions (Foley & Hughes, 2021). Contrary to previous predictions based on Western studies, it is possible that in the Chinese context, more positive sibling relationships may not necessarily lead to constructive mental state-related problem solutions, such as explaining intentions. Instead, these positive relationships might promote general constructive solutions, like apologising and forgiving, thereby reducing interactions that necessitate the use of ToM, such as arguing.

This revised prediction seems plausible given that Doan and Wang (2010) discovered that Chinese parents, compared to their Western counterparts, tend to use fewer mental state terms in daily interactions with their children. Furthermore, the use of mental state language mediated the relationship between culture and children's EU. The current study also supports the idea that Chinese parents are more inclined to discuss situations with their children by referencing the consequences of behaviors (i.e., Discuss) and social norms (i.e., Social Norms) rather than others' mental states (e.g., How Feel). As a result of this limited exposure to mental state language, Chinese

children might reference mental states less frequently when addressing social situations than Western children do (Han et al., 1998; Wang, 2004).

5.1.1.2 Why Not Emotional ToM?

Furthermore, the newly uncovered negative correlations were between positive SRQ and cognitive ToM, not emotional ToM. Wang et al. (2022) also found non-significant associations between sibling effects (though not SRQ) and emotional ToM but did find significant relations with cognitive ToM in Chinese school-aged children. The non-significant correlations for emotional ToM in the current study cannot be attributed to the ceiling effect in emotional ToM scores, given that transformed scores were employed in the correlation analysis. However, the lower sensitivity of affective ToM tasks, which might capture minimal variations in individual differences, could still explain the non-significant correlations.

Importantly, this finding can be construed to mean that the daily interactions and negotiations of young adolescents with their siblings predominantly entail cognitive perspective-taking training, rather than emotional training. There's a theoretical foundation for this interpretation, as Shamay-Tsoory et al. (2010) posited that cognitive ToM is a prerequisite for the development of emotional ToM. The combination of both cognitive ToM and empathy facilitates the functioning of emotional ToM. Therefore, it is plausible that affective ToM is more challenging to cultivate and be influenced by sibling interactions.

5.1.1.3 Why Not Sibling-Rated SRQ?

In both the main and sub-analyses, significant associations with cognitive ToM were observed for both self-rated and researcher-observed positive SRQ, but not for sibling-rated positive SRQ. The results regarding correlations between each SRQ variable also indicate that perceptions of positive SRQ from these three groups varied significantly. The significant findings for observed SRQ could be due to observation bias on the part of the researcher, who was aware of the research hypotheses. However, the credibility of this interpretation is limited by the fact that significant findings for observed SRQ were only found in the sub-analysis and only for positive SRQ. An alternative explanation is that self-rated SRQ provides the most accurate reflection of

SRQ within each family. This is because the researcher only observed the sibling dyads for approximately eight minutes, which may not have been sufficient to capture the full scope of their relationships. This study also confirms that the observation tool did not have good concurrent validity. Furthermore, the siblings completed the SRQ questionnaire after the cooperative drawing game, and their emotions during the game may have influenced their SRQ ratings.

Although the target children were blind to the hypotheses, they might have guessed the researcher's expectations regarding the relationship between ToM and SRQ, and unconsciously rated SRQ according to certain patterns. However, the likelihood of this explanation is very low. Firstly, it would be challenging for participants to predict the negative SRQ-ToM associations that even the researcher did not anticipate. Secondly, if this explanation were true, it would be likely that this study would also find significant relations between sibling twins' ToM and the SRQ rated by themselves in the sub-analysis, which was not the case.

5.1.1.4 Why Not Negative SRQ?

In line with Hou et al. (2022), neither the main analysis nor the sub-analysis of twins in the current study found any associations between negative SRQ variables and children's ToM abilities. Hou et al. (2022) attributed this phenomenon to Chinese parenting styles, as Asian parents often employ authoritarian strategies that resolve sibling conflicts through disciplinary conduct (e.g., criticising children based on parents' values), thereby eliminating opportunities for sibling conflicts to serve as lessons in perspective-taking for children. Consequently, higher levels of sibling conflict do not lead to more training in understanding mental states, and vice versa. This explanation is supported by the current study, which found that the *Boss* category (i.e., controlling and punishing without any communication) was the most frequently used parenting strategy, and all mental-state-related categories were much less used by this sample of Chinese parents (see Table 7).

Importantly, the observed negative SRQ not only reflects the extent of sibling conflict but also other dimensions such as power imbalance (e.g., mild orders). However, daily power imbalance among siblings is less noticeable to parents and may be resolved or unresolved solely by the children themselves. Given the previous

explanation, this raises the question of why observed negative SRQ, which involves less parental intervention, also has no relation with ToM scores. Firstly, it might be because children tend to learn from their parents and use their parents' previously used methods to solve problems, thereby maintaining the impact of parenting strategies. Secondly, the researcher proposed another explanation for the overall non-significant relations between negative SRQ and ToM: the measures of negative SRQ variables may lack sufficient sensitivity to detect a significant relationship. In other words, the current measures, namely, the questionnaire for sibling conflict and the coding scheme of NAC indicators, only evaluate the frequency of negative SRQ events rather than the content of the events, and not all negative events are relevant for ToM development.

5.1.2 Other Findings

The second research objective aimed to delve into the associations between parenting and ToM. Parent Emotions was the sole parenting strategy related to mental states that showed a correlation with ToM. A direct comparison between the significant positive relationship found for Parent Emotions (pertaining to the parents' own feelings) and the non-significant correlations for How Feel (relating to the feelings of anyone involved in social incidents) suggests that parental emotions might serve as a stronger stimulus for children to contemplate mental states than emotions of others (e.g., peers). Moreover, the employment of Parent Emotions emerged as a significant positive predictor of ToM, reinforcing the idea that parenting practices related to mental states promote children's ToM development. This evidence challenges previous findings suggesting that Chinese children's proclivity to reference others' social roles and actions, rather than their mental states, significantly aided their ToM development (Lu et al., 2008).

Notably, Shahaeian et al. (2014) did not identify any significant correlations between Parent Emotions and ToM scores in their Iranian sample. First, the Parent Emotions in Shahaeian et al.'s study pertained to situations where parents discussed their negative feelings with their children. In contrast, this study also included instances where parents mentioned their positive feelings. This difference suggests that parents' positive feelings might stimulate children's reflections on others' mental states more than negative feelings do. Additionally, the difference in findings could stem from the

age of the participants: the Iranian sample comprised preschoolers who may not have developed the foundational social and cognitive skills to fully grasp their parents' emotions. As a result, these emotions might have been too complex to inspire reflections on mental states in younger children.

Moreover, this study identified a significant positive relationship between Active Non-interference and ToM. Additionally, there was a significant negative association between Practical Solution and ToM. Active Non-interference denotes the parenting strategy where parents overlook incidents, believing their involvement is not necessary for resolving the situation. Under these circumstances, children might interpret their parents' silence as positive signals (e.g., trust), encouraging them to handle the situation on their own. Without relying on their parents, children could develop ToM insights by navigating complex emotional scenarios. Conversely, when parents use non-disciplinary methods (Practical Solutions) to aid their children in conflict resolution (e.g., addressing perceived unfairness), the children's issues are immediately addressed, leaving no opportunity to develop mental state understanding. Notably, the positive relationship between Active Non-interference and ToM suggests that parents' decision to remain silent in certain situations does not always hinder children's ToM development, contrary to what was indicated in a previous study by Shahaeian et al. (2014).

However, it remains a mystery as to why the significant correlation of Parent Emotions is limited to cognitive ToM, while the pronounced correlations of Active Non-interference and Practical Solutions were confined to emotional ToM. Notably, the significant associations for Parent Emotions and Practical Solutions were discerned in the main analysis but were absent in the sub-analysis. In contrast, the sub-analysis highlighted significant relationships for both Boss and Ambiguous Non-interference. One plausible reason might be the varied sibling age gaps in the main analysis sample, while the sub-sample consisted solely of twins. As a result, sibling pairs of different age compositions might be influenced differently by the same parenting styles. Another consideration is the small sample sizes for both analyses, which could heighten the likelihood of incurring a Type I or a Type II error (Cohen et al., 2018). This may explain the restrictions to certain ToM type and significant finding for different parenting strategies in two analyses.

Despite the positive relationship of Active Non-interference with emotional ToM, it also exhibited a positive association with observed negative SRQ. This contrasts with a previous study which suggested that parenting styles granting more freedom correlated with more favourable sibling bonds (e.g., sharing, mutual respect; Krejčová et al., 2023). Additionally, this finding also contrasts with the previous conclusion in the literature review that positive parental non-involvement leads to better SRQ for adolescents. To sum up, the current findings suggest that employing the Active Non-interference strategy might enable children to develop insights into mental state understanding during independent interactions with one another. However, it might not necessarily facilitate their ability to address underlying tensions and rebuild close relationships.

5.2 Limitations and Implications

This study has several limitations in its research design. Firstly, the sample size of 30 is small for a quantitative study, which diminishes the statistical power and increases the likelihood of producing a false negative result (Cohen et al., 2018). Furthermore, a small sample size limits the ability to generalise results to a larger population. Secondly, for participant-rated SRQ by questionnaire, this study only investigated two aspects of SRQ that were primarily examined by previous studies: sibling intimacy and sibling conflict. However, Waddell et al. (2001) suggest that sibling relationships, as one type of social relationship, can be assessed using nine dimensions applicable to any social relations: supportive–non-supportive; cooperative–noncooperative; competitive–non-competitive; aggressive–nonaggressive; nurturing–nonnurturing; conflictual–nonconflictual; equal power base–power imbalance; effective social skills–ineffective social skills; and appropriate role relationship–inappropriate role relationship. Future studies should explore as many dimensions of SRQ as possible.

Moreover, this study did not control for several potential covariates, namely, target children's verbal abilities, EFs, and family socioeconomic status (SES). Numerous studies suggest that verbal abilities are strongly correlated with children's ToM performance, particularly for ToM tasks that rely heavily on children's verbal

responses, such as the Strange Stories task (Devine & Hughes, 2013; Milligan et al., 2007). EFs are another critical variable researchers have found to be closely associated with ToM abilities (Aboulafia-Brakha et al., 2011). For instance, Hughes and Ensor (2005) found that children's ToM and EFs were significantly correlated even after accounting for verbal abilities. Concerning family SES, many studies have not only highlighted the significant relationship between family SES and children's ToM but also the moderating role of family SES when exploring the influence of family factors on children's ToM. Ronald et al. (2006) discovered that family SES, as indicated by parental education levels and occupational status, could significantly explain the variance in 9-year-olds' ToM assessed by Strange Stories. A more recent study by Ebert et al. (2017) also demonstrated that the advantages of family elaborative MST on preschoolers' ToM were limited to high SES families. Future research should account for these potential covariates.

Another limitation of this study is that only one parent (either the mother or the father) was invited to participate, thereby overlooking the potential differences in parenting styles between mothers and fathers. Some studies have indicated that mothers and fathers often possess distinct parenting beliefs, exhibit different parenting behaviors, and can even influence their children's ToM development differently across cultures (Chuang & Su, 2009; LaBounty et al., 2008; Yaffe, 2023). Moreover, concerning the sub-analysis of twins, this study assumed that parents employed a consistent parenting style for both twins and did not evaluate the parenting approach for each child individually within the sibling dyads. However, it is plausible that parents adapt their parenting strategies based on each child's unique characteristics. For instance, Newton et al. (2014) identified a bidirectional relationship between young adolescents' prosocial behaviors and maternal sensitivity. Therefore, future research should be attuned to potential gender differences in parenting when conducting related studies.

A limitation of this study was that, aside from coding for ToM scores, the research assistant's coding for observational SRQ and parenting was incomplete. Although the inter-rater reliability of these measures was satisfactory, the coded scores might still be imprecise. Additionally, the inter-rater reliability of measures in this study was not as robust as the data reported by other studies. This discrepancy could be

attributed to the fact that both the researcher and the assistant are master's students, possibly lacking extensive experience in coding. Alternatively, there may be shortcomings in the current coding system and procedure. For instance, it might have been more effective for the researcher to use transcripts of audio recordings from parenting interviews for coding, rather than coding directly from the recordings.

A flaw in the current coding scheme for parenting styles is the categorisation of Social Norms. Upon reviewing the coding during the final stage, the researcher identified two distinct situations within this category. First, when parents used social norms to explain incidents or consequences of certain behaviors, such as saying, "people won't like you if you show off too much," the Social Norms category closely aligned with the Discuss category. In both Discuss and Social Norms, parents explained situations to children. The distinguishing feature of Social Norms, however, was the explicit reference to social norms or public opinions during the discussion. Second, when parents employed social norms as guidelines for children's behaviour, stating for instance, "you should not be impolite to your grandparents," the Social Norms category paralleled the Boss category. Here, parents directly criticised children using social norms without explanations of behaviours. Since the Discuss and Boss approaches are incompatible, it would have been advantageous for the researcher to differentiate the two situations that involve the use of Social Norms into separate categories. Doing so would have facilitated a more accurate and nuanced classification of parenting styles within the coding scheme.

5.3 Contributions and Implications

The first section of this chapter examines the core findings related to the relationships among the three main variables: ToM, SRQ, and parenting styles. These findings close a significant research gap and offer unique insights in comparison to previous studies on the subject. Potential explanations for the novel findings and discrepancies when compared to earlier studies suggest avenues for new research and highlight emerging gaps (detailed further in section 5.5). Beyond the central results, this study also delves into the role of demographic factors as secondary variables. Specifically, this research provides further evidence, similar to the findings of Hou et

al. (2022), suggesting that the number of siblings, birth order, gender composition, and sibling age gap do not significantly influence Chinese children's ToM development. Consequently, subsequent studies might not need to account for these sibling-related factors when exploring similar research themes. Collectively, the evidence garnered from this study aligns more closely with the Reciprocal and Dynamic Development Model than with the Apprenticeship Model and Age Threshold Model.

The current study also revealed that parental sibling status significantly influenced their parenting styles. Specifically, parents without siblings tended to employ the Parent Emotions strategy more frequently than those with siblings. This discovery represents a crucial insight into the relationship between Chinese parental sibling status and specific parenting practices. Earlier research only identified these relationships based on classic classifications of parenting styles, noting that Chinese mothers who were only children reported more authoritative and less authoritarian styles than mothers with siblings (Fan & Chen, 2019). While the reasons why single-child parents in China use mental-state-related strategies, especially Parent Emotions, more frequently than parents with siblings remain unclear, this finding paves the way for future studies to consider the role of parental sibling status (e.g., as a moderator) when examining the impact of Chinese parenting styles on child development.

In addition to research outcomes, the current study offers valuable insights for methodologies in the field. Firstly, this study was the first attempt at applying Fletcher et al.'s (1995) selected subset of Strange Stories to Chinese children aged 11 to 12. The observed ceiling effect for the emotional ToM tasks suggests that future studies should adjust the difficulty level of emotional ToM for this age group. Moreover, the divergent findings between children's cognitive and emotional ToM scores highlight the importance of assessing them separately, rather than combining them into a total score. To the best of the researcher's knowledge, this study is also the first to incorporate white questions preceding mental-state questions in Strange Stories tasks. Based on observations, these white questions effectively alleviate children's stress. However, in the current approach, even if children answered the white questions incorrectly, their ToM responses were directly solicited and included in the data analysis. In future studies, should children answer white questions incorrectly on the first attempt, the

recording could be replayed. If the child errs again, researchers could then provide the correct answer for white questions before soliciting and recording their ToM responses.

Moreover, during the Strange Stories tasks, the researcher observed that most Chinese children tended to provide very concise responses, occasionally resorting to just a phrase or idiom to answer the ToM questions. Such a phenomenon has not been documented in Western samples. From observations made by the researcher, this manner of responding was not due to shyness or a lack of expressive language skills. The researcher posits that this might be because Chinese children have limited experience answering open-ended questions. Some studies have reported that classrooms in China have a higher student-to-teacher ratio compared to Western countries, which results in fewer opportunities for each student to respond to teachers' questions in school (Wanless et al., 2011). Another possibility is that in the Chinese language, the use of concise expressions, such as idioms, is viewed as a sophisticated form of communication, and is thus encouraged among primary school students. This observed phenomenon underscores the significance of making cultural adaptations when implementing measures in a new country. For instance, in this study, a new coding requirement was introduced for the idiom *Zuo Zei Xin Xu*.

Another contribution of this study is the adaptations made to parenting strategies, with the inclusion of six new categories such as Active Non-interference and Practical Solutions, both of which were found to be significantly related to ToM. However, future studies should further refine the interview questions and coding scheme for these parenting categories. For example, for the three non-interference categories, an added question during the interview could determine whether Ambiguous Non-interference was active or passive. Moreover, the current study did not verify whether children recognised their parents' reasons for using passive or active non-interference. This recognition may be essential since the manifestations of the three non-interference strategies in parental behaviors were all silent. The distinct effects on children's ToM development could hinge on whether children recognise their parents' underlying beliefs about parenting.

Moreover, during the interviews, the researcher observed that Chinese parents frequently recounted incidents in which their children were depicted as victims. For instance, in the first scenario, many parents narrated stories where their children were

insulted by peers. Additionally, a significant number of Chinese parents expressed difficulty in recalling specific events, instead offering general descriptions of typical situations and their usual responses. This trend might diminish the credibility of the parents' responses. Their answers could reflect their expectations about their own reactions, rather than their genuine actions in real-life situations. It is also noteworthy that prior research using this methodology often collected parental feedback through questionnaires rather than interviews (Lewis et al., 2006; Shahaeian et al., 2014). Both verbal and written methods have their strengths and drawbacks. A benefit of verbal interviews over written surveys is that the interviewer can immediately seek clarification on answers, and participants can request clarifications on questions, thereby avoiding potential misunderstandings (Cohen et al., 2018). Conversely, a strength of written surveys is that they maintain standardisation, as they do not permit follow-up inquiries. Future researchers should carefully weigh the importance of these factors and determine potential trade-offs when selecting data collection methods.

Lastly, concerning the sources of literature, this study incorporated several studies in Chinese that were published in China and have not been translated into English. The researcher surmised that the significant research gap concerning siblings in Chinese studies might be attributed to language barriers and access limitations in China. Specifically, some Chinese research on this topic is exclusively published in Chinese on Zhi Wang (CNKI). Due to challenges in accessing the website and reading the papers by international researchers, these studies have been largely overlooked.

5.4 Future Directions

Most importantly, future research should replicate this study to determine whether consistent results can be achieved with three main improvements: (1) a larger sample size, at least 100 as recommended by Foley and Hughes (2021); (2) the use of more sensitive, valid, and reliable culturally adapted measures, such as those that can assess the content of sibling conflict in addition to its frequency; and (3) controlling for key extraneous variables, such as family SES. Only after these steps can research findings offer practical suggestions for family and school education. Building on these

replications, future research could also explore the effects of the three variables in a longitudinal design.

The second research direction suggests that future studies should validate the credibility of the current interpretations of research outcomes and delve deeper into the potential reasons behind unexplained findings from this study. First, future research could examine the specific conflict-resolution strategies employed by Chinese children with varying levels of SRQ and assess how these strategies relate to their training in understanding mental states. Second, researchers could investigate the distinct proportions of training for cognitive and emotional ToM during the daily interactions of Chinese siblings. Third, it would be valuable to investigate if siblings adopt their parents' values, beliefs, and conflict-resolution strategies when addressing their own disagreements or conflicts. Fourth, it would be crucial to determine whether children's ToM development is more significantly influenced when they recognise their parents' mental states compared to others' mental states. Moreover, to figure out the reason behind the link between Chinese parents' sibling status and the strategy of Parent Emotions, future studies could start from investigating the different parenting style patterns Chinese parents adopt toward a single child compared to siblings, and how these styles vary depending on the age and gender composition of sibling groups. There may be a unique intergenerational transmission of parenting styles in China.

Additionally, the current study focused solely on one parental factor: parenting styles. Future research might consider including other parental factors that have been demonstrated to impact both ToM and SRQ, such as parent-child relationship quality. Furthermore, this study recommends that future research should place a greater emphasis on studies involving twins. Beyond offering insights into genetic and environmental influences on child development, research on twins can shed unique light on sibling relationships. Given their similarities (e.g., physical appearance, age), twins may be more sensitive to competition and parental intervention in their interactions compared to regular sibling dyads.

5.5 Conclusion

This study, utilising a cross-sectional multimethod multi-informant correlational design, was the first to explore the relationships between young adolescents' ToM development, their SRQ, and parenting styles in China. As anticipated, significant moderate relations emerged among these three variables. Young adolescents' self-rated sibling intimacy negatively predicted their cognitive ToM scores, while the Parent Emotions strategy employed by their parents positively predicted cognitive ToM. Additionally, young adolescents whose parents frequently used the Active Non-interference and Practical Solutions strategies scored higher and lower on emotional ToM tasks, respectively. The study also indicated that a greater use of Active Non-interference by parents was linked to more negative researcher-observed SRQ. Regarding demographic information, young adolescents' age, gender, sibship size, birth order, sibling age gap, and sibling gender composition were found to be unrelated to their ToM abilities. Nonetheless, parents' sibling status was found to be related to parenting styles. Considering the small sample size and methodological limitations of the current study, future research should aim to replicate these findings with enhanced research design and measures. Simultaneously, it is vital to verify the credibility of interpretations derived from the current results.

References

Aboulafia-Brakha, T., Christe, B., Martory, M. D., & Annoni, J. M. (2011). Theory of mind tasks and executive functions: A systematic review of group studies in neurology. *Journal of neuropsychology, 5*(1), 39-55. https://doi.org/10.1348/174866410X533660

Bandura, A. (1977). *Social learning theory.* Prentice Hall.

Baumrind, D. (1978). Parental disciplinary patterns and social competence in children. *Youth & Society, 9*(3), 239-267. https://doi.org/10.1177/0044118X7800900302

Baumrind, D. (1991). Parenting styles and adolescent development. In J. Brooks-Gunn, R. Lerner, & A. C. Peterson (Eds.), *The Encyclopedia of Adolescence* (pp. 746-758). Garland.

Blake, J. (1981). Family size and the quality of children. *Demography, 18*, 421-442. https://doi.org/10.2307/2060941

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

Blyth, D. A., & Foster-Clark, F. S. (1987). Gender differences in perceived intimacy with different members of adolescents' social networks. *Sex Roles, 17*(11-12), 689-718. https://doi.org/10.1007/BF00287683

Blyth, D., Hill, J., & Thiel, K. (1982). Early adolescents' significant others: Grade and gender differences in perceived relationships with familial and nonfamilial

adults and young people. *Journal of Youth and Adolescence, 11*, 425-450. https://doi.org/10.1007/BF01538805

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185-216. https://doi.org/10.1177/135910457000100301

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Harvard University Press.

Buhrmester, D., & Furman, W. (1990). Perceptions of sibling relationships during middle childhood and adolescence. *Child Development, 61*, 1387-1398. https://doi.org/10.1111/j.1467-8624.1990.tb02869

Buist, K. L., Deković, M., & Prinzie, P. (2013). Sibling relationship quality and psychopathology of children and adolescents: A meta-analysis. *Clinical Psychology Review, 33*(1), 97-106. https://doi.org/10.1016/j.cpr.2012.10.007

Calero, C. I., Salles, A., Semelman, M., & Sigman, M. (2013). Age and gender dependent development of theory of mind in 6-to 8-years old children. *Frontiers in Human Neuroscience, 7*, 1-7. https://doi.org/10.3389/fnhum.2013.00281

Carr, A., Slade, L., Yuill, N., Sullivan, S., & Ruffman, T. (2018). Minding the children: A longitudinal study of mental state talk, theory of mind, and behavioural adjustment from the age of 3 to 10. *Social Development, 27*, 826–840. https://doi.org/10.1111/sode.12315

Cassidy, K. W., Fineberg, D. S., Brown, K., & Perkins, A. (2005). Theory of mind may be contagious, but you don't catch it from your twin. *Child development, 76*(1), 97-106. https://doi.org/10.1111/j.1467-8624.2005.00832.x

Chandler, M. J., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role-taking competence. *International Journal of Behavioral Development, 7*(2), 145-156. https://doi.org/10.1177/016502548400700203

Chao, R., & Tseng, V. (2002). Parenting of Asians. In M. Bornstein (Eds.), *Handbook of parenting (Second Edition). Volume 4: Social conditions and applied parenting* (pp. 59–93). Lawrence.

Chen, B. B. (2019). Chinese adolescents' sibling conflicts: Links with maternal involvement in sibling relationships and coparenting. *Journal of Research on Adolescence, 29*(3), 752-762. https://doi.org/10.1111/jora.12413

Chen, B. B., Chen, X., & Wang, X. (2021). Siblings versus parents: Warm relationships and shyness among Chinese adolescents. *Social Development, 30*(3), 883-896. https://doi.org/10.1111/sode.12509

Chen, X., Dong, Q., & Zhou, H. (1997). Authoritative and authoritarian parenting practices and social and school adjustment. *International Journal of Behavioural Development, 21*, 855–873. https://doi.org/10.1080/016502597384703

Chuang, S. S., & Su, Y. (2009). Do we see eye to eye? Chinese mothers' and fathers' parenting beliefs and values for toddlers in Canada and China. *Journal of Family Psychology, 23*(3), 331–341. https://doi.org/10.1037/a0016015

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education (8th edition)*. Routledge.

Cole, K., & Mitchell, P. (2000). Siblings in the development of executive control and a theory of mind. *British Journal of Developmental Psychology, 18*(2), 279-295. https://doi.org/10.1348/026151000165698

Cox, M. J., & Paley, B. (2003). Understanding families as systems. *Current Directions in Psychological Science, 12*, 193–196. https://doi.org/10.1111/1467-8721.01259

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70*(4), 853–865. https://doi.org/10.1111/1467-8624.00061

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child development, 84*(3), 989-1003. https://doi.org/10.1111/cdev.12017

Devine, R. T., & Hughes, C. (2018). Family correlates of false belief understanding in early childhood: A meta-analysis. *Child Development, 89*(5), 971–987. https://doi.org/10.1111/cdev.12682

Dirks, M. A., Persram, R., Recchia, H. E., & Howe, N. (2015). Sibling relationships as sources of risk and resilience in the development and maintenance of internalizing and externalizing problems during childhood and adolescence. *Clinical psychology review, 42*, 145-155. https://doi.org/10.1016/j.cpr.2015.07.003

Doan, S. N., & Wang, Q. (2010). Maternal discussions of mental states and behaviors: Relations to emotion situation knowledge in European American and immigrant Chinese children. *Child development, 81*(5), 1490-1503. https://doi.org/10.1111/j.1467-8624.2010.01487.x

Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review, 60*(5), 746-761. https://doi.org/10.2307/2096320

Downey, D. B. (2001). Number of siblings and intellectual development: The resource dilution explanation. *American Psychologist, 56*(6-7), 497-504. https://doi.org/10.1037/0003-066X.56.6-7.497

Downey, D. B., & Condron, D. J. (2004). Playing well with others in kindergarten: The benefit of siblings at home. *Journal of Marriage and Family, 66*(2), 333-350. https://doi.org/10.1111/j.1741-3737.2004.00024.x

Dunn, J. (1983). Sibling relationships in early childhood. *Child Development, 54*, 787 – 811. https://doi.org/10.2307/1129886

Ebert, S., Peterson, C., Slaughter, V., & Weinert, S. (2017). Links among parents' mental state language, family socioeconomic status, and preschoolers' theory of mind development. *Cognitive Development, 44*, 32–48. https://doi.org/10.1016/j.cogdev.2017.08.005

Eccles, J., Templeton, J., Barber, B., & Sotone, M. (2003). Adolescence and emerging adulthood: The critical passage ways to adulthood. In M. H. Bornstein, L. Davidson, C. L. M. Keyes, & K. A. Moore (Eds.), *Well-being: Positive development across the life course* (pp. 383-406). Lawrence Erlbaum Associates.

Fan, J., & Chen, B. B. (2020). Parenting styles and coparenting in China: The role of parents and children's sibling status. *Current Psychology, 39*(5), 1505-1512. https://doi.org/10.1007/s12144-019-00379-7

Fang, F. X., Henry, M. W., Liu, Y. J., Liu, G. X., & Kang, R. (2009). Longitudinal perspectives: The sequences of theory-of-mind development in Chinese preschoolers. *Acta Psychologica Sinica, 41*(8), 706-714. https://journal.psych.ac.cn/acps/EN/Y2009/V41/I08/706

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M. A., Kumar, V., & Gazanizad, N. (2010). Theory of mind, birth order, and siblings among preschool children. *American Journal of Scientific Research, 7*(3), 25-35.

Feng, Y., Whiteman, S. D., Xu, S., Li, L., Jin, S., & French, D. C. (2019). Chinese adolescents' relationships with mothers, fathers, and siblings: Associations with youth's internalising and externalising problems. *Journal of Relationships Research, 10*, 1-11. https://doi.org/10.1017/jrr.2019.11

Field, A. P. (2018). *Discovering statistics using IBM SPSS Statistics (5th edition)*. SAGE Publications.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57*(2), 109-128. https://doi.org/10.1016/0010-0277(95)00692-R

Foley, S., & Hughes, C. (2021). Family influences on theory of mind in middle childhood. In R. T., Devine, & S. Lecce (Eds), *Theory of Mind in Middle Childhood and Adolescence* (pp. 101-121). Routledge.

Fox, J. (2016). *Applied regression analysis and generalized linear models (3rd edition)*. Sage Publications.

Furman, W., & Buhrmester, D. (2009). Methods and measures: The Network of Relationships Inventory: Behavioural Systems Version. *International Journal*

of Behavioural Development, 33(5), 470-478. https://doi.org/10.1177/0165025409342634

Gass, K., Jenkins, J., & Dunn, J. (2007). Are sibling relationships protective? A longitudinal study. *Journal of child psychology and psychiatry, 48*(2), 167-175. https://doi.org/10.1111/j.1469-7610.2006.01699.x

Han, J. J., Leichtman, M. D., & Wang, Q. (1998). Autobiographical memory in Korean, Chinese, and American children. *Developmental Psychology, 34*(4), 701–713. https://doi.org/10.1037/0012-1649.34.4.701

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children, and adults. *Journal of autism and Developmental disorders, 24*(2), 129-154. https://doi.org/10.1007/BF02172093

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. *Behavioural and Brain Sciences, 33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Hou, X. H., Gong, Z. Q., Wang, L. J., Zhou, Y., & Su, Y. (2020). A reciprocal and dynamic development model for the effects of siblings on children's theory of mind. *Frontiers in Psychology, 11*, 1-12. https://doi.org/10.3389/fpsyg.2020.554023

Hou, X. H., Wang, L. J., Li, M., Qin, Q. Z., Li, Y., & Chen, B. B. (2022). The roles of sibling status and sibling relationship quality on theory of mind among Chinese preschool children. *Personality and Individual Differences, 185*, 1-8. https://doi.org/10.1016/j.paid.2021.111273

Howe, N., Rinaldi, C. M., Jennings, M., & Petrakos, H. (2002). "No! The lambs can stay out because they got cozies": Constructive and destructive sibling conflict, pretend play, and social understanding. *Child Development, 73*(5), 1460–1473. https://doi.org/10.1111/1467-8624.00483

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives, 9*(3), 149-153. https://doi.org/10.1111/cdep.12124

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology, 34*(5), 1026–1037. https://doi.org/10.1037/0012-1649.34.5.1026

Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2-year-olds: A family affair? *Developmental Neuropsychology, 28*(2), 645-668. https://doi.org/10.1207/s15326942dn2802_5

Hughes, C., & Ensor, R. (2006). Behavioural problems in 2-year-olds: Links with individual differences in theory of mind, executive function, and harsh parenting. *The Journal of Child Psychology and Psychiatry, 47*(5), 488–497. https://doi.org/10.1111/j.1469-7610.2005.01519.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections, and new directions for studies of typical and atypical development. *Social Development, 13*, 590-619. https://doi.org/10.1111/j.1467-9507.2004.00285.x

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind.

Personality and individual differences, 54(5), 616-621. https://doi.org/10.1016/j.paid.2012.11.022

Jambon, M., Madigan, S., Plamondon, A., Daniel, E., & Jenkins, J. M. (2019). The development of empathic concern in siblings: A reciprocal influence model. *Child Development, 90*(5), 1598–1613. https://doi.org/10.1111/cdev.13015

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology, 32*(1), 70–78. https://doi.org/10.1037/0012-1649.32.1.70

Jensen, A. C., Whiteman, S. D., Loeser, M. K., & Bernard, J. M. B. (2018). Sibling Influences. In R. J. R. Levesque (Ed.), *Encyclopedia of adolescence* (pp. 1-7). Springer. https://doi.org/10.1007/978-3-319-33228-4_37

Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. (2023). Is ChatGPT a good translator? A preliminary study. https://doi.org/10.48550/arXiv.2301.08745

Kennedy, K., Lagattuta, K. H., & Sayfan, L. (2015). Sibling composition, executive function, and children's thinking about mental diversity. *Journal of Experimental Child Psychology, 132*, 121-139. https://doi.org/10.1016/j.jecp.2014.11.007

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25-41. https://doi.org/10.1016/S0010-0277(03)00064-7

Kim, B. S., & Hong, S. (2004). A psychometric revision of the Asian Values Scale using the Rasch model. *Measurement and Evaluation in Counselling and Development, 37*(1), 15-27. https://doi.org/10.1080/07481756.2004.11909747

Kim, S. Y., Wang, Y., Orozco-Lapray, D., Shen, Y., & Murtuza, M. (2013). Does "tiger parenting" exist? Parenting profiles of Chinese Americans and adolescent developmental outcomes. *Asian American Journal of Psychology, 4*(1), 7–18. https://doi.org/10.1037/a0030612

Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, *6*(5), 26-41. https://eric.ed.gov/?id=EJ1154775

Kramer, L. (2014). Learning emotional understanding and emotion regulation through sibling interaction. *Early Education and Development, 25*(2), 160-184. https://doi.org/10.1080/10409289.2014.838824

Kramer, L., Perozynski, L. A., & Chung, T. Y. (1999). Parental responses to sibling conflict: The effects of development and parent gender. *Child Development, 70*(6), 1401-1414. https://doi.org/10.1111/1467-8624.00102

Krejčová, K., Chýlová, H., & Rymešová, P. (2023). Analysis of siblings' relationship and parenting style using structure modelling approach. *Plos one, 18*(2), e0281266. https://doi.org/10.1371/journal.pone.0281266

Kuntoro, I. A., Peterson, C. C., & Slaughter, V. (2017). Culture, parenting, and children's theory of mind development in Indonesia. *Journal of Cross-Cultural Psychology, 48*(9), 1389-1409. https://doi.org/10.1177/0022022117725404

LaBounty, J., Wellman, H. M., Olson, S., Lagattuta, K., & Liu, D. (2008). Mothers' and fathers' use of internal state talk with their young children. *Social Development, 17*(4), 757-775. https://doi.org/10.1111/j.1467-9507.2007.00450.x

Lagattuta, K. H., & Wellman, H. M. (2002). Differences in early parent-child conversations about negative versus positive emotions: Implications for the development of psychological understanding. *Developmental Psychology, 38*(4), 564–580. https://doi.org/10.1037/0012-1649.38.4.564

Lam, C. B., McHale, S. M., Lam, C. S., Chung, K. K. H., & Cheung, R. Y. M. (2021). Sibling relationship qualities and peer and academic adjustment: A multi-informant longitudinal study of Chinese families. *Journal of Family Psychology, 35*(5), 584–594. https://doi.org/10.1037/fam0000744

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. https://doi.org/10.2307/2529310

Lecce, S., Bianco, F., Devine, R. T., & Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: A short-term longitudinal study. *Journal of experimental child psychology, 163*, 69-86. https://doi.org/10.1016/j.jecp.2017.06.011

Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. M. (1996). Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child development, 67*(6), 2930-2947. https://doi.org/10.1111/j.1467-8624.1996.tb01896.x

Lewis, C., Huang, Z., & Rooksby, M. (2006). Chinese preschoolers' false belief understanding: Is social knowledge underpinned by parental styles, social interactions, or executive functions? *Psychologia, 49*(4), 252-266. https://doi.org/10.2117/psysoc.2006.252

Lu, H., Su, Y., & Wang, Q. (2008). Talking about others facilitates theory of mind in Chinese preschoolers. *Developmental Psychology, 44*(6), 1726–1736. https://doi.org/10.1037/a0013074

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interaction. In P. H. Mussen, & E. M. Heatherington (Eds), *Handbook of child psychology: Socialization, personality, and social development* (pp. 1-101). Wiley.

McAlister, A., & Peterson, C. C. (2006). Mental playmates: Siblings, executive functioning, and theory of mind. *British Journal of Developmental Psychology, 24*(4), 733-751. https://doi.org/10.1348/026151005X70094

McAlister, A., & Peterson, C. (2007). A longitudinal study of child siblings and theory of mind development. *Cognitive Development, 22*(2), 258-270. https://doi.org/10.1016/j.cogdev.2006.10.009

McAlister, A., & Peterson, C. (2013). Siblings, theory of mind, and executive functioning in children aged 3–6 years: New longitudinal evidence. *Child development, 84*(4), 1442-1458. https://doi.org/10.1111/cdev.12043

McHale, S. M., Updegraff, K., A., & Whiteman, S. D. (2012). Sibling relationships and influences in childhood and adolescence. *Journal of Marriage and Family, 74*(5), 913-930. https://doi.org/10.1111/j.1741-3737.2012.01011.x

Newton, E. K., Laible, D., Carlo, G., Steele, J. S., & McGinley, M. (2014). Do sensitive parents foster kind children, or vice versa? Bidirectional influences between children's prosocial behavior and parental sensitivity. *Developmental Psychology, 50*(6), 1808–1816. https://doi.org/10.1037/a0036495

Milevsky, A., Schlechter, M. J., & Machlev, M. (2011). Effects of parenting style and involvement in sibling conflict on adolescent sibling relationships. *Journal of Social and Personal Relationships, 28*(8), 1130-1148. https://doi.org/10.1177/0265407511406894

Miller, S. A. (2013). Children's understanding of second-order false belief: Comparisons of content and method of assessment. *Infant and Child Development, 22*(6), 649-658. https://doi.org/10.1002/icd.1810

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), 622-646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Noller, P. (2005). Sibling relationships in adolescence: Learning and growing together. *Personal relationships, 12*(1), 1-22. https://doi.org/10.1111/j.1350-4126.2005.00099.x

Oliver, B. R., & Pike, A. (2021). Introducing a novel online observation of parenting behaviour: Reliability and validation. *Parenting, 21*(2), 168-183. https://doi.org/10.1080/15295192.2019.1694838

Olson, S. L., Lopez-Duran, N., Lunkenheimer, E. S., Chang, H., & Sameroff, A. J. (2011). Individual differences in the development of early peer aggression: Integrating contributions of self-regulation, theory of mind, and parenting. *Development and psychopathology, 23*(1), 253-266. https://doi.org/10.1017/S0954579410000775

O'Reilly, J., & Peterson, C. C. (2014). Theory of mind at home: Linking authoritative and authoritarian parenting styles to children's social understanding. *Early*

child development and care, 184(12), 1934-1947. https://doi.org/10.1080/15295192.2019.1694838

Osborne, J. W., & Overbay, A. (2008). Best practices in data cleaning: How outliers and "fringeliers" can increase error rates and decrease the quality and precision of your results. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 205-213). Sage Publications.

Paine, A. L., Pearce, H., van Goozen, S. H., de Sonneville, L. M., & Hay, D. F. (2018). Late, but not early, arriving younger siblings foster firstborns' understanding of second-order false belief. *Journal of Experimental Child Psychology, 166*, 251-265. https://doi.org/10.1016/j.jecp.2017.08.007

Pavarini, G., de Hollanda Souza, D., & Hawk, C. K. (2013). Parental practices and theory of mind development. *Journal of Child and Family Studies, 22*, 844-853. https://doi.org/10.1007/s10826-012-9643-8

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child development, 65*(4), 1228-1238. https://doi.org/10.1111/j.1467-8624.1994.tb00814.x

Perner, J., & Wimmer, H. (1985). "John Thinks That Mary Thinks That…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*, 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Peterson, C. C. (2000). Kindred spirits: Influences of siblings' perspectives on theory of mind. *Cognitive Development, 15*(4), 435–455. https://doi.org/10.1016/S0885-2014(01)00040-5

Peterson, C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*, 502-517. https://doi.org/10.1111/j.1467-8624.2005.00859.x

Piaget, J. (1957). *Construction of reality in the child*. Routledge & Kegan Paul.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind. *Behavioural and Brain Sciences, 1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Prime, H., Plamondon, A., Pauker, S., Perlman, M., & Jenkins, J. M. (2016). Sibling cognitive sensitivity as a moderator of the relationship between sibship size and children's theory of mind: A longitudinal analysis. *Cognitive Development, 39*, 93–102. https://doi.org/10.1016/j.cogdev.2016.03.005

Recchia, H. E., & Howe, N. (2009). Associations between social understanding, sibling relationship quality, and siblings' conflict strategies and outcomes. *Child Development, 80*(5), 1564-1578. https://doi.org/10.1111/j.1467-8624.2009.01351.x

Ronald, A., Viding, E., Happé, F., & Plomin, R. (2006). Individual differences in theory of mind ability in middle childhood and links with verbal ability and autistic traits: A twin study. *Social Neuroscience, 1*, 412–425. https://doi.org/10.1080/17470910601068088

Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. A. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*(1), 161–174. https://doi.org/10.1037/0012-1649.34.1.161

Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social development, 8*(3), 395-411. https://doi.org/10.1111/1467-9507.00103

Samuels, H. R. (1980). The effect of an older sibling on infant locomotor exploration of a new environment. *Child Development, 51*(2), 607-609. https://doi.org/10.2307/1129305

Shahaeian, A. (2015). Sibling, family, and social influences on children's theory of mind understanding: New evidence from diverse intracultural samples. *Journal of Cross-Cultural Psychology, 46*(6), 805-820. https://doi.org/10.1177/0022022115583897

Shahaeian, A., Nielsen, M., Peterson, C. C., & Slaughter, V. (2014). Iranian mothers' disciplinary strategies and theory of mind in children: A focus on belief understanding. *Journal of Cross-Cultural Psychology, 45*(7), 1110-1123. https://doi.org/10.1177/0022022114534772

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology, 47*(5), 1239–1247. https://doi.org/10.1037/a0023899

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex, 46*(5), 668–677. https://doi.org/10.1016/j.cortex.2009.04.008

Song, J. H., & Volling, B. L. (2018). Theory-of-Mind development and early sibling relationships after the birth of a sibling: Parental discipline matters. *Infant and child development, 27*(1), 1-17. https://doi.org/10.1002/icd.2053

Sulloway, F. J. (1996). *Born to rebel: Birth order, family dynamics, and creative lives*. Pantheon Books.

Sun, L. H., & Zhang, A. R. (2018). Investigation on the sibling relationship of two children in different parenting styles (in Chinese). *Journal Shanghai Education Research, 8*, 59–63. https://doi.org/10.16194/j.cnki.31-1059/g4.2018.08.014

Tafreshi, D., & Racine, T. (2016). Children's interpretive theory of mind: The role of mothers' personal epistemologies and mother-child talk about interpretation. *Cognitive Development, 39*, 57–70. https://doi.org/10.1016/j.cogdev.2016.04.003

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child development, 77*(2), 465-481. https://doi.org/10.1111/j.1467-8624.2006.00882.x

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: The growing complexity of recursive thinking ability. *Europe's Journal of Psychology, 11*(1), 112-124. https://doi.org/10.5964/ejop.v11i1.829

Vinden, P. G. (2001). Parenting attitudes and children's understanding of mind: A comparison of Korean American and Anglo-American families. *Cognitive development, 16*(3), 793-809. https://doi.org/10.1016/S0885-2014(01)00059-4

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Waddell, J., Pepler, D., & Moore, T. (2001). Observations of sibling interactions in violent families. *Journal of Community Psychology, 29*(3), 241-258. https://doi.org/10.1002/jcop.1016

Waite, E. B., Shanahan, L., Calkins, S. D., Keane, S. P., & O'Brien, M. (2011). Life events, sibling warmth, and youths' adjustment. *Journal of Marriage and Family, 73*, 902–912. https://doi.org/10.1111/j.1741-3737.2011.00857.x

Wang, Q. (2004). The Emergence of Cultural Self-Constructs: Autobiographical Memory and Self-Description in European American and Chinese Children. *Developmental Psychology, 40*(1), 3–15. https://doi.org/10.1037/0012-1649.40.1.3

Wang, S., Andrews, G., Pendergast, D., Neumann, D., Chen, Y., & Shum, D. H. (2022). A cross-cultural study of theory of mind using strange stories in school-aged children from Australia and mainland China. *Journal of Cognition and Development, 23*(1), 40-63. https://doi.org/10.1080/15248372.2021.1974445

Wanless, S. B., McClelland, M. M., Lan, X., Son, S. H., Cameron, C. E., Morrison, F. J., Chen, F., Chen, J., Li, S., Lee, K., & Sung, M. (2013). Gender differences in behavioural regulation in four societies: The United States, Taiwan, South Korea, and China. *Early Childhood Research Quarterly, 28*(3), 621-633. https://doi.org/10.1016/j.ecresq.2013.04.002

Wellman, H. M., Fang, F. X., Liu, D., Zhu, L. Q., & Liu, G. X. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychology Science, 17*(12), 1075–1081. https://doi.org/10.1111/J.1467-9280.2006.01830.X

Wellman, H. M., & Liu, D. (2004). Scaling theory-of-mind tasks. *Child Development, 75*, 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*, 245-275. https://doi.org/10.10160010-0277/(90)90024-E

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development, 80*(4), 1097-1117. https://doi.org/10.1111/j.1467-8624.2009.01319.x

Whiteman, S. D., McHale, S. M., & Soli, A. (2011). Theoretical perspectives on sibling relationships. *Journal of family theory & review, 3*(2), 124-139. https://doi.org/10.1111/j.1756-2589.2011.00087.x

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. https://doi.org/10.1016/0010-0277(83)90004-5

Wright, B. C., & Mahfoud, J. (2012). A child-centred exploration of the relevance of family and friends to theory of mind development. *Scandinavian Journal of Psychology, 53*(1), 32-40. https://doi.org/10.1111/j.1467-9450.2011.00920.x

Xie, S., & Li, H. (2019). 'Tiger mom, panda dad': A study of contemporary Chinese parenting profiles. *Early Child Development and Care, 189*(2), 284-300. https://doi.org/10.1080/03004430.2017.1318870

Yaffe, Y. (2023). Systematic review of the differences between mothers and fathers in parenting styles and practices. *Current Psychology, 42*, 1-14. https://doi.org/10.1007/s12144-020-01014-6

Yeates, K. O., Bigler, E. D., Dennis, M., Gerhardt, C. A., Rubin, K. H., Stancin, T., Taylor, H. G., & Vannatta, K. (2007). Social outcomes in childhood brain disorder: A heuristic integration of social neuroscience and developmental

psychology. *Psychological Bulletin, 133*(3), 535–556. https://doi.org/10.1037/0033-2909.133.3.535

Yu, J. J., & Gamble, W. C. (2008). Pathways of influence: Marital relationships and their association with parenting styles and sibling relationship quality. *Journal of Child and Family Studies, 17*, 757-778. https://doi.org/10.1007/s10826-008-9188-z

Yucel, D., & Yuan, A. V. (2015). Do siblings matter? The effect of siblings on socio-emotional development and educational aspirations among early adolescents. *Child Indicators Research, 8*, 671-697. https://doi.org/10.1007/s12187-014-9268-0

Appendix A: Power Analysis

A power analysis was conducted for two-tailed Pearson Product Moment correlation analysis with power of 0.8 and Pearson correlation parameter of 0.5 through SPSS. The estimated sample size was 29 as the following figure showed.

Power Analysis Table

	N	Actual Power[b]	Test Assumptions			
			Power	Null	Alternative	Sig.
Pearson Correlation[a]	29	.814	.8	0	.5	.05

a. Two-sided test.

b. Based on Fisher's z-transformation and normal approximation with bias adjustment.

Appendix B: An Invitation Letter to the Headmasters

Dear Sir/Madam,

My name is [redacted]. I am a Chinese student pursuing a Master of Science in Education (Child Development and Education) at the University of Oxford. In my current programme, I am supervised by [redacted], [redacted]. I am writing to request your assistance with my dissertation research project. My research explores the associations between sibling relationship quality, children's theory-of-mind development (i.e., abilities to understand other's mental states such as intentions and beliefs), and parenting styles. I would like to ask you to be our facilitator/gatekeeper to access students and parents of your school.

If your school decides to participate, you will be asked to distribute a video to the headteachers of classes in grade 5 and 6 for them to briefly introduce this research to students and parents. I am looking for around 30 students aged 11-12 years who have at least one sibling (8 years and over). These students, their one sibling, and their one parent will be invited to our online research. Each family unit will participate in one online research session for around 1 hour.

I would be extremely grateful if your school could assist us with this project. Given that four-decade-long One Child policy in our country, your school will contribute a lot to this understudied area of sibling relationships in China. I have attached the informational sheet detailing the research project in greater detail. Whether or not you feel it would be appropriate for your school to participate, I would be grateful if you would complete the pro-forma below and return it to me via [redacted]. If you have any further questions about this project, please do not hesitate to contact me. I am looking forward to hearing from you. Have a nice day!

Best regards,
[redacted]

[School name]

[School address]

[Headmaster's name]

☐ We do not wish to participate in this project.

☐ We would like to find out more about this project.

☐ We would like to take part in this project.

If you would like further information, or are interested in taking part, please give the name of a contact person for your school, and details of the best way to contact him or her.

Contact name: _____

Contact email: _____

Contact telephone number: _____

Thank you for your help.

Appendix C: Participant Information Sheet (For Children)

Central University Research Ethics Committee Approval Reference:
EDUC_C1A_23_173

Information sheet for Child Participants

We are inviting you to join in a research study. My name is [redacted], and I work at the University of Oxford in the Department of Education.

Before you decide if you would like to join in, it's important to understand what the research is about, why we're doing it and what it would involve for you. Please read and think about this leaflet carefully. Please feel free to talk to your family, friends, or the researcher about it if you want. If anything isn't clear or you have more questions you can ask your parent to give us a call and we can discuss it with you and your parent.

Why are we doing this research?

My research explores the associations between sibling relationship quality, young adolescents' abilities to understand other's mental states such as intentions and beliefs, and your parents' parenting styles.

Why have I been invited to take part?

If you are from [redacted] primary school, you are included in this research because you are 11-12 years old and have a sibling who is 8 years old and above. If you are the sibling, you are included in this research because you are 8 years old and above. In total, we recruit 30 sibling pairs like you guys are.

Do I have to take part?

No - It is up to you to decide if you want to take part in this research. You are free to stop taking part at any time during the research without giving a reason by telling your teacher, the researcher, or your parent/guardian. You do not have to say why, and this will not affect your education. If you decide to stop, no one will be upset with you.

What will happen if I take part in the research?

You and your sibling and one of your parents will be invited to our online research session (around 1 hour in total). You and your sibling will be asked to complete a short questionnaire (around 5 minutes) separately. Then, you two will be invited to an interesting online drawing game (around 10 minutes), and this will require you guys to work together. And then, you will be asked to listen to a few stories and answer some questions regarding the beliefs of the characters in the story (around 10 minutes). Your parents won't be around you if you don't like. You will only be recorded for the drawing activity.

What are the possible disadvantages and risks in taking part?

There are no known high risks to participate in this study. Nevertheless, you may quarrel with your sibling in the drawing game.

Are there any benefits in taking part?

Your family will gain a participation certificate with the signature of [redacted] and [redacted].

What information will be collected and what happens to the results of the research?

Results are kept strictly confidential, and only the people doing the research, or helping with the research, can look at the data. Only a number will be used to identify you, and all information and results are kept in in a password-protected electronic file on University of Oxford secure servers. I will change the names of your school, teacher, and all the students when I write about my research. No one will know that you have taken part unless you tell them yourself. The findings from the research will/may be written up in my master's thesis. If I want to use the information for anything else, I will ask your permission. At the end of my research, I will write to your school about what I found out in my research. You are welcome to read this if you are interested. All research data and records will be stored for 3 years after publication or public release of the work of the research.

Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research.

The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest.

Further information about your rights with respect to your personal data is available from https://compliance.web.ox.ac.uk/individual-rights.

Will anyone else know I'm doing this?

We will keep your information private. This means we will only tell those who have a need or right to know, such as the research team and your parent/guardian. We will only share information that has your name and address removed.

What if I don't want to take part in the research anymore?

Just tell your parent/guardian and the people carrying out the research that you don't want to take part. You don't have to give a reason, and no one will be annoyed with you. It is YOUR choice.

Who has reviewed the research?

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee.

What do I do now?

Please tell your parents, guardians and/or teacher whether you are happy to take part.

What if there is a problem or something goes wrong?

Please tell us if you are worried about any part of this research, by contacting the researcher [redacted] You may also talk to your teacher/parent/guardian who will let the researcher know. If you are still unhappy or wish to make a complaint, either you or your teacher/parent/guardian can contact the chair of the Research Ethics Committee at the University of Oxford:

Chair, **Social Sciences & Humanities Interdivisional Research Ethics Committee**; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB

Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

[redacted]

Thank you for reading – please ask me any questions.

Appendix D: Strange Stories

This study adapted practice story's coding scheme from Wang et al. (2022), and eight stories' coding scheme from White et al. (2009).

Practice story:

Sarah and Tom are going on a picnic. It is Tom's idea. He says it will be a lovely sunny day. But when they get the food out, it rains, and the food gets all wet. Sarah says: "Oh yes, a lovely day for a picnic alright!"

Chinese translation: 丽丽和汤姆要去野餐。今天去野餐是汤姆的主意。他说这将是一个美好的晴天。但是当他们把食物拿出来的时候，天下雨了，食物都湿了。丽丽说:"哦，是的，真是野餐的好天气!"

Back translation: Lili and Tom are going for a picnic. Today, going for a picnic was Tom's idea. He said it would be a beautiful sunny day. But when they took out the food, it started to rain, and the food got wet. Lili said, "Oh, yes, what a great picnic weather!"

ToM Q: Why did Sarah (Lili) say that?

A: To understand sarcasm, children first need to infer that Sarah knows it is not a lovely day and then to infer that she is unhappy now.

Story 1: Double Bluff (cognitive)

Simon is a big liar. Simon's brother Jim knows this, he knows that Simon never tells the truth! Now yesterday Simon stole Jim's ping-pong paddle, and Jim knows Simon has hidden it somewhere, though he can't find it. He's very cross. So, he finds Simon and he says, "Where is my ping-pong paddle? You must have hidden it either in the cupboard or under your bed because I've looked everywhere else. Where is it, in the cupboard or under your bed"? Simon tells him the paddle is under his bed.

Chinese translation: 小雨是个大骗子，大雷知道这一点。昨天，小雨偷了大雷的乒乓球拍。大雷找不到这个球拍，他知道小雨肯定把它藏起来了。大雷很生气。于是，他找到了小雨，问道："我的乒乓球拍在哪里？你一定是把它藏在橱柜里或者床底下了，因为其他地方我都找过了。它在哪里？在橱柜里还是在你的床底下？"小雨告诉大雷，球拍在床底下。

Back translation: Xiaoyu is a big liar, and Dalei knows this. Yesterday, Xiaoyu stole Dalei's ping pong paddle. Dalei couldn't find the paddle and he knew that Xiaoyu had definitely hidden it. Dalei was very angry. So, he found Xiaoyu and asked, "Where is my ping pong paddle? You must have hidden it in the cupboard or under the bed

because I've searched everywhere else. Where is it? Is it in the cupboard or under your bed?" Xiaoyu told Dalei that the paddle was under the bed.

White Q: What did Jim (Dalei) lose?

Answer: ping-pong paddle.

ToM Q: Why will Jim (Dalei) look in the cupboard for the ping-pong paddle?

2 points—reference to Jim knowing Simon lies.

1 point—reference to facts (that's where it really is, Simon's a big liar) or Simon hiding it without reference to implications of lying.

0 points—reference to general nonspecific information (because he looked everywhere else)

Story 2: Double Bluff (cognitive)

During the war, the Red army captures a member of the Blue army. They want him to tell them where his army's tanks are; they know they are either by the sea or in the mountains. They know that the prisoner will not want to tell them, he will want to save his army, and so he will certainly lie to them. The prisoner is very brave and very clever, he will not let them find his tanks. The tanks are really in the mountains. Now when the other side asks him where his tanks are, he says, "They are in the mountains."

Chinese translation: 战争期间，红军俘虏了一名蓝军士兵。他们想让他说出蓝军的坦克在哪里，他们知道那些坦克不是在海边就是在山上。红军知道这名囚犯想拯救他的蓝军，所以他不会告诉他们真相。这个囚犯既勇敢又聪明，他不会让红军找到在山里的坦克。当红军问囚犯坦克在哪里时，囚犯说:"在山里。"

Back translation: During the war, the Red Army captured a soldier from the Blue Army. They wanted him to reveal where the Blue Army's tanks were located, knowing that the tanks were either near the seaside or in the mountains. The Red Army knew that the prisoner wanted to protect his own army, so he wouldn't tell them the truth. The prisoner was both brave and clever, and he wouldn't let the Red Army find the tanks in the mountains. When the Red Army asked the prisoner where the tanks were, he said, "In the mountains".

White Q: How many armies are there?

Answer: two (red and blue).

ToM Q: Why did the prisoner say that?

2 points—reference to fact that other army will not believe and hence look in other place, reference to prisoner's realization that that's what they'll do, or reference to double bluff.

1 point—reference to outcome (to save his army's tanks) or to mislead them.

0 points—reference to motivation that misses the point of double bluff (he was scared)

Story 3: Persuasion (emotional)

Brian is always hungry. Today at school it is his favourite meal—sausages and beans. He is a very greedy boy, and he would like to have more sausages than anybody else, even though his mother will have made him a lovely meal when he gets home! But everyone is allowed two sausages and no more. When it is Brian's turn to be served, he says, "Oh, please can I have four sausages, because I won't be having any dinner when I get home!"

Chinese translation: 小白总是很饿。今天在学校里有他最喜欢吃的香肠。他是一个非常贪婪的男孩，他比任何人都想要更多的香肠，即使他妈妈会在他回家时为他做一顿美味的晚餐。可惜按规定每个人只能拿两根香肠，不能多。当轮到小白盛饭时，他说："哦，请给我四根香肠，因为我回家后没有晚饭吃。"

Back translation: Xiaobai is always very hungry. Today at school, they had his favourite sausage. He's a very greedy boy and wants more sausage than anyone else, even though his mother will make him a delicious dinner when he gets home. Unfortunately, according to the rules, each person can only take two sausages, no more. When it was Xiaobai's turn to get food, he said, "Oh, please give me four sausages because I won't have any dinner to eat when I get home."

White Q: How many sausages can each student get?

Answer: two

ToM Q: Why does Brian (Xiaobai) say this?

2 points—reference to fact that he's trying to elicit sympathy, being deceptive.

1 point—reference to his state (greedy), outcome (to get more sausages) or factual.

0 points—reference to a motivation that misses the point of sympathy elicitation/deception, or factually incorrect.

Story 4: Persuasion (emotional)

Jill wanted to buy a kitten, so she went to see Mrs. Smith, who had lots of kittens she didn't want. Now Mrs. Smith loved the kittens, and she wouldn't do anything to harm them, though she couldn't keep them all herself. When Jill visited, she wasn't sure she wanted one of Mrs. Smith's kittens, since they were all males and she had wanted a female. But Mrs. Smith said, "If no one buys the kittens I'll just have to drown them!"

Chinese translation: 丽丽想去找王奶奶买小猫，因为王奶奶有很多不想要的小猫。王奶奶是个爱猫的人，她不会做任何伤害小猫的事，尽管她不能把它们都养在身边。当丽丽来拜访时，她开始犹豫要不要买，因为王奶奶的猫都是公的，而丽丽想要一只母的。但是王奶奶说："如果没有人买小猫，我就只能把它们淹死了。"

Back translation: Lili wants to go to Grandma Wang's house to buy a kitten because Grandma Wang has many unwanted kittens. Grandma Wang loves cats, and she won't do anything to harm them, even though she can't keep them all with her. When Lili came to visit, she hesitated whether to buy one or not because all of Grandma Wang's cats were male, and Lili wanted a female one. But Grandma Wang said, "If no one buys the kittens, I will have to drown them".

White Q: Does Lili want to buy a male cate or a female cat?

Answer: female

ToM Q: Why did Mrs. Smith (Grandma Wang) say that?

2 points—reference to persuasion, manipulating feelings, trying to induce guilt/pity

1 point—reference to outcome (to sell them or get rid of them in a way which implies not drowning) or simple motivation (to make Jill sad)

0 points—reference to general knowledge or dilemma without realization that the statement was not true (she's a horrible woman)

Story 5: White lies (emotional)

One day Aunt Jane came to visit Peter. Now Peter loves his aunt very much, but today she is wearing a new hat; a new hat which Peter thinks is very ugly indeed. Peter thinks his aunt looks silly in it, and much nicer in her old hat. But when Aunt Jane asks Peter, "How do you like my new hat?" Peter says, "Oh, it's very nice."

Chinese translation: 一天，姨妈来看望小明。小明非常爱他的姨妈，但是今天她戴了一顶小明觉得很丑的新帽子。小明认为姨妈戴着新帽子看起来很傻，而戴她的旧帽子会漂亮得多。但是当姨妈问小明:"你觉得我的新帽子怎么样"，小明说:"嗯，非常好。"

Back translation: One day, Auntie came to visit Xiao Ming. Xiao Ming loves his aunt very much, but today she was wearing a new hat that Xiao Ming thought was very ugly. Xiao Ming thought that his aunt looked silly wearing the new hat, and that she would look much prettier wearing her old one. But when Auntie asked Xiao Ming, 'What do you think of my new hat?', Xiao Ming said, “Hmm, it looks very nice”.

White Q: What does Xiaoming think is ugly?

Answer: the new hat

ToM Q: Why does he say that?

2 points—reference to white lie or wanting to spare her feelings; some implication that this is for aunt’s benefit rather than just for his, desire to avoid rudeness or insult.

1 point—reference to trait (he’s a nice boy) or relationship (he likes his aunt); purely motivational (so she won’t shout at him) with no reference to aunt’s thoughts or feelings; incomplete explanation (he’s lying, he’s pretending).

0 points—reference to irrelevant or incorrect facts/feelings (he likes the hat; he wants to trick her).

Story 6: White lies (emotional)

Helen waited all year for Christmas, because she knew at Christmas, she could ask her parents for a rabbit. Helen wanted a rabbit more than anything in the world. At last Christmas Day arrived, and Helen ran to unwrap the big box her parents had given her. She felt sure it would contain a little rabbit in a cage. But when she opened it, with all the family standing round, she found her present was just a boring old set of encyclopedias, which Helen did not want at all! Still, when Helen’s parents asked her how she liked her Christmas present, she said, “It’s lovely, thank you. It’s just what I wanted.”

Chinese translation: 悦悦一整年都在等待生日，因为她知道在生日时，她可以向父母要一只兔子。悦悦最想要的是一只兔子。生日到了，悦悦跑去打开父母送给她的大盒子，她发现盒子里只是一套她根本不想要的百科全书。尽管如此，当父母问她是否喜欢这份生日礼物时，她说:“这很好，谢谢你们。这正是我想要的。”

Back translation: Yueyue had been waiting for her birthday all year because she knew that on her birthday, she could ask her parents for a rabbit. A rabbit is what Yueyue wanted most. When her birthday arrived, Yueyue ran to open the big box her parents had given her, but she found a set of encyclopaedias that she didn't want at all. Even so, when her parents asked her if she liked the birthday gift, she said, “It's very good, thank you. This is exactly what I wanted.”

White Q: What gift did Yueyue receive?

Answer: encyclopaedias

ToM Q: Why did she say this?

2 points—reference to white lie or wanting to spare their feelings; some implication that this is for parent's benefit rather than just for her, desire to avoid rudeness or insult.

1 point—reference to trait (she's a nice girl) or relationship (she likes her parents); purely motivational (so they won't shout at her) with no reference to parent's thoughts or feelings; incomplete explanation (she's lying, she's pretending).

0 points—reference to irrelevant or incorrect facts/feelings (she likes the present, she wants to trick them).

Story 7: Misunderstanding (cognitive)

One late night, old Mrs. Peabody is walking home. She doesn't like walking home alone in the dark because she is always afraid that someone will attack her and rob her. She really is a very nervous person! Suddenly, out of the shadows comes a man. He wants to ask Mrs. Peabody what time it is, so he walks toward her. When Mrs. Peabody sees the man coming toward her, she starts to tremble and says, "Take my purse, just don't hurt me please!"

Chinese translation: 一天深夜，张奶奶走在回家的路上。她不喜欢在黑暗中独自走回家，因为她总是害怕有人袭击抢劫她。她高度紧张。突然，一个人从阴影中走了出来。他走向她是因为想问张奶奶现在几点了。张奶奶看到那个男人朝她走来，浑身发抖，说:"拿着我的钱包，请不要伤害我。"

Back translation: One late night, Grandma Zhang was walking back home. She didn't like walking alone in the dark because she was always afraid that someone would attack or rob her. She was highly nervous. Suddenly, a person came out of the shadows. He approached her and asked Grandma Zhang what time it was. Grandma Zhang saw the man walking towards her and she trembled, saying, "Take my wallet, please don't hurt me."

White Q: Where was Grandma Zhang going?

Answer: home.

ToM Q: Why did she say that?

2 points—reference to her belief that he was going to mug her or her ignorance of his real intention.

1 point—reference to her trait (she's nervous) or state (she's scared) or intention (so he wouldn't hurt her) without suggestion that fear was unnecessary.

0 points—factually incorrect/irrelevant answers; reference to the man actually intending to attack her.

Story 8: Misunderstanding (cognitive)

A burglar who has just robbed a shop is making his getaway. As he is running home, a policeman on his beat sees him drop his glove. He doesn't know the man is a burglar, he just wants to tell him he dropped his glove. But when the policeman shouts out to the burglar, "Hey, you! Stop!" the burglar turns round, sees the policeman and gives himself up. He puts his hands up and admits that he did the break-in at the local shop.

Chinese translation: 一个刚刚抢劫了一家商店的窃贼正在逃跑。当他跑回家的时候，一个巡逻的警察看到他掉了手套。警察不知道这个人是窃贼，他只是想告诉他他的手套掉了。但是当警察对窃贼喊道:"嘿，你，停下。"窃贼转过身来，对警察举手承认了是他闯入了当地的商店。

Back translation: A thief who had just robbed a store was running away. When he ran back home, a patrolling police officer saw him drop his glove. The police officer didn't know that the man was a thief, he just wanted to tell him that he had dropped his glove. But when the police officer yelled to the thief, "Hey, you, stop!", the thief turned around, raised his hands, and admitted to the police officer that he had broken into the local store.

White Q: What did the burglar drop?

Answer: glove

ToM Q: Why did the burglar do that?

2 points—reference to belief that policeman knew that he'd burgled the shop.

1 point—reference to something factually correct in story.

0 points—factually incorrect/irrelevant answers.

Appendix E: Sibling Relationship Questionnaire

欢迎来到我们的研究！请调整到一个舒适的姿势来回答接下来的问题。完成这个问卷大概需要五分钟。
Welcome to our research! Please prepare yourself a comfortable position for the following questions. It takes about five minutes to complete the questionnaire.

A. 被试编号 Participant ID: _____.

B. 你今年多大啦？Who old are you?

请以今天与你一同参加本次研究的这位兄弟姐妹为依据，来回答下面的问题。
Please answer the following questions based on the sibling who is taking part in the study with you today.

下面的表述是否符合你和你的兄弟姐妹呢？请凭直觉作答。
Please rate the following statements about you and your sibling based on your intuition.

1 = 完全不符合 strongly disagree
2 = 不太符合 somewhat disagree
3 = 部分符合，部分不符合 neither agree or disagree
4 = 比较符合 somewhat agree
5 = 完全符合 strongly agree

1. 我经常向他/她寻求建议和帮助。 I often ask him/her for advice and help.	1	2	3	4	5
2. 我希望自己像他/她一样。 I wish to be like him/her.	1	2	3	4	5
3. 我经常和他/她意见不一致。 I often disagree with him/her.	1	2	3	4	5
4. 不管我做什么，他/她都会接受我。 He/she will accept me no matter what I do.	1	2	3	4	5
5．我经常生他/她的气。 I often get angry with him/her.	1	2	3	4	5
6．我们经常吵架。 We often quarrel.	1	2	3	4	5
7. 我经常向他/她分享我的心情或者秘密。 I often share my feelings or secrets with him/her.	1	2	3	4	5
8. 他/她了解我的兴趣爱好。 He/she knows my interests and hobbies.	1	2	3	4	5

9. 他/她经常向我寻求意见和帮助。 He/she often asks me for advice and help.	1	2	3	4	5
10. 我经常对他/她不耐烦。 I often get impatient with him/her.	1	2	3	4	5
11. 他/她对我而言是重要的。 He/she is important to me.	1	2	3	4	5
12. 我们经常惹恼对方。 We often annoy each other.	1	2	3	4	5
13. 我对我们之间的关系是满意的。 I am satisfied with our relationship.	1	2	3	4	5
14. 他/她经常向我分享他/她的心情或者秘密。 He/she often shares his/her feelings or secrets with me.	1	2	3	4	5
15. 他/她经常对我不耐烦。 He/she often gets impatient with me.	1	2	3	4	5
16. 他/她经常生我的气。 He/she often gets angry with me.	1	2	3	4	5
17. 我了解他/她的兴趣爱好。 I know his/her interests and hobbies.	1	2	3	4	5
18. 我不管他/她做什么，我都会接受他/她。 I will accept him/her no matter what he/she does.	1	2	3	4	5

Appendix F: Observation Indicators

The following coding scheme was created by Waddell et al. (2001).

148

AFFECTIVE CLIMATE

Positive Affective Climate: The emotional climate of the interaction is positive and both siblings seem to feel "good" about the interaction. There is notable interest and pleasure. Behaviour is expressed with pleasurable facial expressions, affection, and /or enthusiastic interest.

5. Approval, compliments, praise: verbal/physical expression that praises, reinforces, shows approval/admiration of sibling. Nodding in approval, smiling, laughing. Positive statements made by sibling about the other. Examples: "Good for you", "You're good at this game", "That's it", "Way to go!"
Seeks attention from sibling for purpose of approval, reinforcement, reassurance, related to game or performance on task.

6. Offers or Displays Affection: Behaviours or conversations that reflect affectionate behaviour. Any positive physical contact such as touching affectionately, kissing, hugging, holding hands, patting, gentle playful touch. Whispering, telling secrets, affectionate teasing, giggling, moving closer together, verbal expression of caring/affection, esteem, humour mimicking. Caretaking behaviours.

7. Smiles: Smiles are exchanged between siblings. Eye contact should be observed during smiles.

8. Laughter: High intensity giggling. Any laughter that is in fun, not sarcastic or mocking. It does not have to be shared laughter, but it usually is and is enjoyed by both siblings.

9. Excitement: Boisterous fun. Showing strong positive feelings. Siblings conversations can get louder and sometimes high pitched. A high level of energy is evident. Can be silliness.

10. Cooperate, help, share, comply: Invites sibling or involves sibling in decision making/problem solving. Offers or gives an object spontaneously or on request. In collaboration with sibling refers to instructions, seeks assistance from adults for purpose of clarification or explanation of game\activity. Works collaboratively and cooperatively on game - participation in game/activity is a mutual process. Supports and acknowledges the efforts of sibling to help/assist

11. Sympathy, comfort, support: Statements meant to console, ease distress, or provide emotional support to one another. Verbal or physical consolation when sibling is in some way distressed, acceptance of comfort, affection and reassurance.

12. Teach: Offers helpful advice, guidance, suggestion.

Clarifies questions, instructions. Teaches sibling rules of game if sibling expresses that he/she is not familiar with rules. Responds to sibling request for information, help and guidance. 149

13. Conflict Management/Problem solving: Attempts to either keep a conflict from escalating, or to reduce a conflict that has already escalated between the siblings. May involve trying to resolve the argument, compromising, or employing behaviours that reduce the level of negative affect. Examples: referring to a rule of the game/task, referring to family rules re behaviour, compromising or making an offer of some kind. "Okay, we'll finish this game and then play with the Rubik's cube", "Okay, I'll do my name on the Etchosketch, and you do your name". There must be some kind of conflict (at least a brief one) to code conflict management/problem solving as occurring. The siblings must do something to manage the conflict/issue. Both siblings must agree to accept the solution/compromise.

14. Rough and Tumble Play: Physical contact that is neutral or nonaversive - clowning around, playfights with no domination or malice intent.

15. Subdued, Peaceful: The emotional climate is positive, yet subdued and quiet. There is quiet conversation or activity between siblings.

16. Other positive: Use this category to denote any other positive activity or affect occurring in the sibling interaction that is not listed above. List these events or affects on the coding sheet.

Negative Affective Climate: The emotional climate of the interaction is negative. Either or both voice tone and facial expressions indicate siblings displeasure, irritation, hostility.

21. Mild limits: Mild limits are expressed by one sibling to control the other sibling's behaviour. Mild limits involve a simple, direct, low-keyed instruction or prohibition for behaviour. Affect surrounding a mild limit is usually pretty mild and does not involve anger. Usually signalled with the words "Don't" or "Stop", "wait", "go" with a strong tone of voice, bossy but without apparent anger.

22. Major limits: Major limits involve stronger measures to control the other sibling's behaviour, and can be accompanied by more intense affects (anger, frustration, yelling). Can occur when milder limits are not effective, or the sibling may go straight to major limits. Examples: Grabbing game or other object from sibling without comment/request/permission, threatening, swearing.

23. Negative verbal/physical behaviour: -any negative expression, name calling, jeering, taunting with sarcastic, disdainful tone, threatening, quarrelling, demeaning remarks, insults, swearing/defensiveness, whining, statements which derogatory, pestering, temper tantrums. Complaints re sibling, game, general events, making fun of sibling, "egging on", nudging, tapping, grabbing objects such as the game, gentle physical reprimands, an uncooperative response to an initiating behaviour ie refusing to help sibling in any way, refusing to accept direction, suggestions, advice or help from sibling.

24. Physical Aggression: Aversive physical contact, spitting, pushing, hitting, pulling, shoving, kicking, biting, punching, pulling hair, hitting with an object. Usually associated with anger and conflict.

25. Ignore, no response, rejecting: Passive resistance to a request, demand by sibling. Includes a lack of response to comments made by sibling which are game related or general conversation. Verbal or behavioural overture made by sibling and is rejected or ignored by the other. Both positive and negative feeling states will be included here. These include likes, dislikes, wants, desires, distastes, opinions etc. An overture will be considered to have been ignored when no direct response is made by the second party. If no acknowledgement is directed to the overture, it was ignored. An overture will be considered to have been rejected when the response made by the second party challenges the validity of the verbal and physical overture. It gives the message that overture is wrong, bad, or inappropriate.

26. Negative help seeking: Tattling on sibling, demanding clarification of rules without discussion or negotiation with sibling.

27. Sadness, unhappiness: From tone of voice, or facial expression, siblings appear to be sad and unhappy.

28. Subdued, boredom: The affective climate is rather flat, subdued, dull or quiet and the siblings appear bored. Siblings seem not to know what to do, and sit unoccupied with a facial expression that seems blank, uninterested, and not particularly content.

29. Other negative: Use this category to denote any other negative events or affects occurring during the sibling interaction that are not listed above.

Appendix G: Parenting Scenarios

Revised by Shahaeian et al. (2014), in the Appendix.

Scenarios Presented to Parents

1	Can you remember a time recently that your child teased or hit another child?
2	Can you remember a time recently when your child shouted at you or your husband, made fun of either of you, or referred to you or your husband in some unflattering way?
3	Can you remember a time recently when your child damaged something that didn't belong to him?
4	Can you remember a time when you were going to a party and your child didn't want to get dressed as you want him to do?
5	Can you think of the last time your child disagreed with you?
6	Can you remember the last time your child acted impolitely?

Adaptations to the current study: (No adaptations to scenario 3 and 5)

1. Scenario 1: Since Chinese parents often portrayed their children as victims, the first scenario has been expanded to include mutual conflicts (such as quarrelling and fighting) instead of solely directional negative behaviors (teasing and hitting).
2. Scenario 2: Given that many Chinese families include grandparents in addition to parents and children, this study expanded the scope of shouting incidents in the second scenario to include any elder family members, not just parents.
3. Scenario 2: Removed the phrase "refer to you or your husband in an unflattering way" due to its ambiguous definition. Added "talk back to" to this scenario, as it is a behaviour commonly reported by Chinese parents.
4. Scenario 4: Broadened the scope of the scenario by replacing "get dressed" with "behave".

Appendix H: Parenting Styles Coding Scheme

The following coding scheme was created by Shahaeian et al. (2014), pp. 1115-1116.

Discuss, describe consequences (dubbed "Discuss"). This is when mothers tried to discuss the situation with their child, explaining why a behavior is not appropriate or telling them about the consequences of certain behavior. Any responses in which the mother treats the child as a mindful agent who can think and make decisions are included in this category. This category also includes situations where the mother gives the child opportunity to make up his or her mind or to explain to the mother what she or he has done. For example, "I explained to him that he shouldn't touch other people's belonging without them knowing." This example is included as the mother is explaining to the child that touching another's belonging without permission is not appropriate. If this example was something like "I told him he shouldn't touch other people's belongings," it would be coded into the *Boss* category. Conversely, a statement like "I told him people will think you are a bad boy if you touch their belongings" would be coded under *Social Norms*. Another example is, "I tried to talk to them and see why they have started fighting," in which the mother is giving her children a chance to explain to her the reason of their behavior. And, "I told her she can wear this shirt but then she might catch a cold" which is an example of consequence.

Let Child Decide. This code was used when parents said that in response to the situation in the question they let the child decide or left it to the child to choose. This category is used if the parent added no further explanation to make the response appropriate to appear as a *Discuss* answer, for example: "I try to let him choose what he wants."

How Feel. This category includes responses that refer to the feelings of someone involved in the scenario. This is when a mother is referring to the feelings of the victim; for instance, "what would you feel if it were you instead of her/him."

Avoidance, silence, or passivity (dubbed "Silence"). This category is for responses that show a passive approach where the parent is not directly involved or does not challenge the child's behavior. This is when mothers said that they do not talk to the child, they ignored the child's misbehavior, or they asked someone else to talk to the child. Sometimes they responded, "I didn't tell him anything," "I try to control my anger and don't mention it," or "I told his older sister to talk to him about it." Any response that does not actively involve the parent in the situation is included in this category. It should be noted that this category is different from neglecting the child when the child needs the parent (such as silent treatment). In these responses, the mother is avoiding telling the child he has done something wrong. If the mother uses silence as a punishment (e.g., one mother said, "I get angry at them and then ignore them when they come and ask for some- thing, this way they know they should behave themselves") this would be coded as a *Boss* strategy.

Deciding for the child, punishing or controlling (Boss). This category of responses refers to situations when parents do not discuss things with the child. This situation is

a “parent to child” decision- making situation, when parents try to control the child’s behavior and do not give the child a chance to respond, defend, or make decisions, without any further explanation or discussion. Examples are as follows: “I told him what to do,” “She should do what I ask her to,” “I told him he can’t watch TV tonight,” “I told her she is a bad girl.”

Social Norms. This is when parents refer to the importance of other people’s judgments of behavior, such as appropriate behavior in public or in front of others. Also included in this category are sole responses relating something socially accepted without further explanation, by referring to respect for adults and politeness. Examples are as follows: “I told him people will laugh at him if he wears that,” “So what would others think if you talk to your dad like this?” or “You should respect your older sister.”

Parent Emotions. This category includes responses referring to any sort of negative feelings that mothers expressed to the child, for instance, “I would be very unhappy and told her I am very upset with what you did.”

Appendix I: Pilot Study

Table 1
Demographic information of three pilot families

No.	Target Child		Sibling		Parent
	Age	Gender	Age	Gender	
1	13	Female	25	Male	Mother
2	8	Female	14	Male	Mother
3	13	Male	7	Male	Mother

Note. The pilot families were recruited from the researcher's friends and relatives.

Appendix J: Normality Checks for Main Analysis

Table 1

Tests of Normality for ToM and SRQ variables

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ToM_C	.137	30	.158	.937	30	.076
ToM_E	.233	30	<.001	.894	30	.006
ToM	.131	30	.198	.957	30	.253
Self_W	.167	30	.032	.949	30	.157
Self_C	.108	30	.200*	.945	30	.127
Self_Total	.104	30	.200*	.973	30	.616
Sib_SW	.136	30	.162	.942	30	.100
Sib_SC	.220	30	<.001	.904	30	.011
Sib_ST	.109	30	.200*	.941	30	.100

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

As shown in Table J1, all the ToM and participant-rated SRQ variables in the main analysis except emotional ToM and sibling-rated sibling conflict have passed the normality check as assessed by Shapiro-Wilk's test ($p > .05$). After these transformations, both variables passed the normality checks, as shown in Table J2. All the parenting variables violated the assumption of normality as assessed by Shapiro-Wilk's test ($p < .05$), as indicated in Table J3.

Table 2

Tests of normality for transformed variables

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ToM_E_sqrt	.194	30	.005	.938	30	.081
Sib_SC_sqrt	.190	30	.007	.941	30	.099

a. Lilliefors Significance Correction

Table 3

Tests of normality for parenting variables

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
General D	.317	30	<.001	.784	30	<.001
How Feel	.483	30	<.001	.452	30	<.001
Parent Emotions	.497	30	<.001	.484	30	<.001
Cogntive D	.510	30	<.001	.424	30	<.001
Social Norms	.477	30	<.001	.527	30	<.001
AD	.376	30	<.001	.674	30	<.001
Passive NI	.492	30	<.001	.476	30	<.001
Active NI	.420	30	<.001	.646	30	<.001
ANI	.329	30	<.001	.674	30	<.001
Boss	.299	30	<.001	.791	30	<.001
Let Child D	.522	30	<.001	.357	30	<.001
Practical Solutions	.510	30	<.001	.424	30	<.001

a. Lilliefors Significance Correction

Table 4

Descriptive statistics for observed SRQ

Statistics

		Obser_Pos	Obser_Neg
N	Valid	28	28
	Missing	2	2
Mean		2.71	2.82
Std. Deviation		1.243	1.416
Skewness		.092	.339
Std. Error of Skewness		.441	.441
Kurtosis		-.953	-1.269
Std. Error of Kurtosis		.858	.858

As shown in Table J4, both the skewness statistics of two observed variables are between -2 and +2, suggesting that the skewness of data distribution is acceptable (George & Mallery, 2010[3]); the kurtosis of two variables is also acceptable because all the kurtosis statistics are within the range from -7 to +7 (Hair et al., 2010[4]). Together these results with Figures 4 and 5 suggest that we can treat two variables as normally distributed.

[3] George, D., & Mallery, P. (2010). *SPSS for Windows step by step: A simple study guide and reference.* Pearson Education.

[4] Hair, J. F. J., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis*. Prentice-Hall.

Figure 1

Data distribution of emotional ToM scores

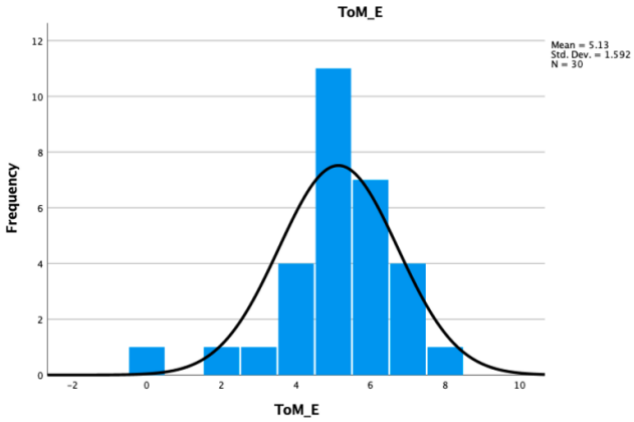


Figure 2

Data distribution of sibling-rated sibling conflict

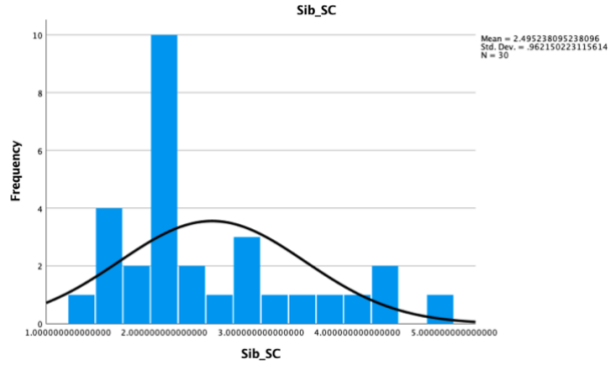


Figure 3

Data distribution of Discuss parenting strategy

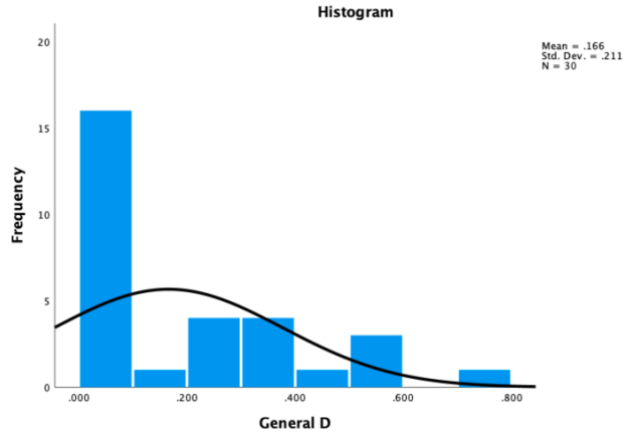


Figure 4

Data distribution of observed positive SRQ

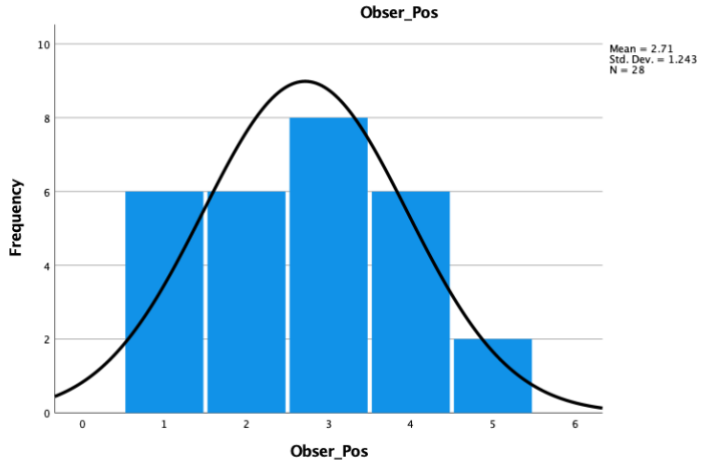
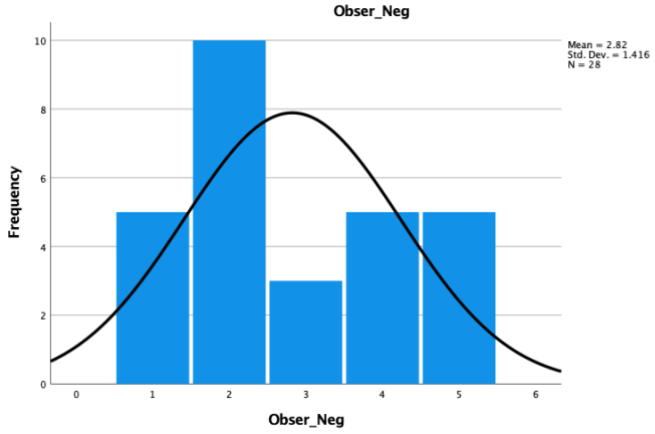


Figure 5

Data distribution of observed negative SRQ



Appendix K: Assumption checks for Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin–Watson
1	.546[a]	.298	.246	1.459	1.645

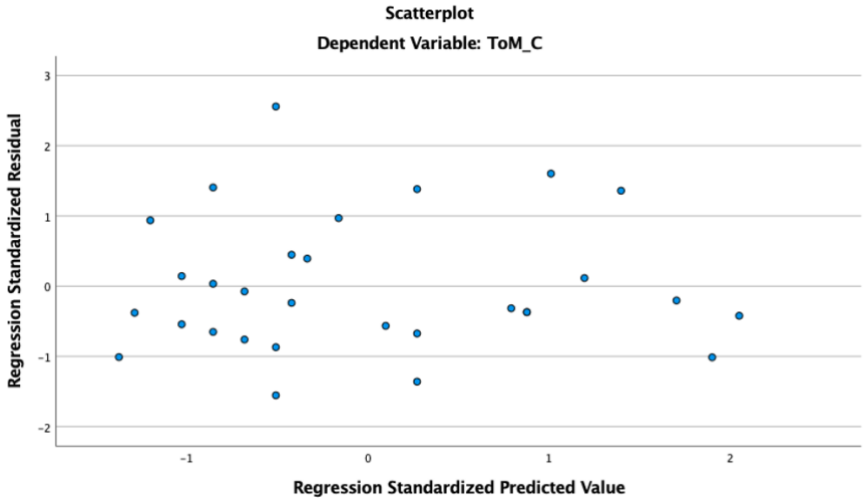
a. Predictors: (Constant), Parent Emotions, Self_W

b. Dependent Variable: ToM_C

There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.65 (between 0 to 4). As indicated in Figure K1, the residuals formed a horizontal band, suggesting that the relation between independent and dependent variables is likely to be linear. There was homoscedasticity as the spread of the residuals did not increase or decrease across the predicted values (see Figure K1). The check for multicollinearity was passed because all the VIF values of variables were around 1 (< 10).

Figure 1

Scatterplot



Appendix L: Ethics Approval

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
DEPARTMENTAL RESEARCH ETHICS COMMITTEE

Department of Education
15 Norham Gardens, Oxford OX2 6PY
student.curec@education.ox.ac.uk; staff.curec@education.ox.ac.uk

[redacted]
Department of Education, Social Sciences Division
University of Oxford

26 April 2023

Dear [redacted]

Research ethics approval

Research title: Exploring the Associations between Sibling Relationship Quality, Parenting Styles, and Theory-of-Mind Development in Chinese Adolescents: A Preliminary Analysis.

Research ethics reference: EDUC_C1A_23_173

The above application has been considered on behalf of the Education Departmental Research Ethics Committee (DREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the DREC, ethics approval has now been granted for this study.

Please note the following:

Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's website.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to staff.curec@education.ox.ac.uk / student.curec@education.ox.ac.uk or ethics@socsci.ox.ac.uk.

Yours sincerely

Gosia Marschall

DREC member

cc: [redacted]@education.ox.ac.uk