Dialogic Reading with Mandarin-Speaking Children: Effects on Expressive Vocabulary and Narrative Competence

by

© Ling Li

A Dissertation submitted

to the School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Faculty of Education

Memorial University of Newfoundland

October 2020

St. John's Newfoundland and Labrador

Abstract

This research study examines the effects of a home-based dialogic reading (DR) intervention on the expressive vocabulary and narrative competence of Mandarin-speaking children. DR is a shared-reading approach that emphasizes verbal interactions between adults and young children. Expressive vocabulary is generally conceived as children's knowledge of word meanings in speaking or writing, and oral expressive vocabulary is focused on in the present study. Narrative competence refers to their ability to produce and comprehend stories. Three aspects of narrative production were measured—macrostructure, microstructure, and evaluation. Narrative comprehension was measured using comprehension questions that targeted story grammar and problem resolution. In the present study, 81 Mandarin-speaking kindergartners, aged four to five, were pretested on expressive vocabulary and narrative competence and assigned to one of two conditions—DR or customary reading. Parents of the children assigned to the DR group participated in three DR workshops and conducted at least 48 one-on-one shared-reading sessions at home with their child during the 12-week intervention period. Expressive vocabulary and narrative competence were retested immediately after the intervention and again four months later. In-home observations of shared-reading sessions recorded on video were used to track the changes in parents' reading behaviours, and multiple measures were employed to collect

program-evaluation data. The DR intervention yielded a greater effect on the children's expressive vocabulary at the delayed posttest stage. However, its effect on indicators of narrative competence was not significant except for one indicator of narrative microstructure—adverb density. Positive changes were found in the reading behaviours of parents in the DR group over the intervention period, and positive feedback was also obtained from the comprehensive evaluation of the DR program.

Acknowledgements

I wish to express my deepest gratitude to my supervisor, Dr. Antoinette (Toni) Doyle, for providing guidance and feedback at every stage of my Ph.D. journey, including the completion of this dissertation. She has inspired me to be professional and to do the right things as I develop as a scholar, and she has also provided valuable help and support during my time living in Newfoundland. Her professionalism, honesty, and generosity have greatly influenced me and will definitely continue to affect my life in the future.

I wish to convey my sincere appreciation to my research committee. I may be the most fortunate Ph.D. student in the world because I have had not only the best supervisor but also the best committee. Without their patience and support, I could not have overcome the numerous obstacles I faced as I conducted my research. Especially, I thank Dr. Jing Zhang for her help with expertise in literacy, methodology, and statistics. I am also grateful for the support of Dr. Jackie Hesson with regard to psychological knowledge, ethical considerations, and measurement decisions. Also, I thank Dr. Leslie Redmond for her expert input in statistics and linguistics.

I am indebted to Associate professor, Jianhui Zhang, the dean at the *Guizhou Professional Development Center for Early Childhood Educators,* and Xuefei Wang, an administrator at the centre, who helped me contact the participating kindergarten. I also thank all the staff of the participating kindergarten—especially, Yan Gong, the

headmistress, Bei Chen, the dean, teachers Xuemin Du and Tingting Zhang, and the other 28 teachers and childcare workers of the 10 participating classes—for their cooperation and hard work. Also, I wish to convey my deep gratitude to the 80 parents and 81 children who were involved in my dissertation research. Without their cooperation, the goals of this research project could not have been achieved.

I thank Xuefei Wang and her parents—Deshun Wang and Ping Zhang—Aiyu Liu and her parents—Tinghu Liu and Weihong Zhang—and Dr. Kun Yang for taking care of me in Guiyang and also for their friendship. I would also like to thank Nan Ma for her friendship and also for her support in Beijing. I also thank Yanhua Xu, a Ph.D. candidate at Capital Normal University in Beijing, for his support in providing methodological and statistical resources.

I wish to express special thanks to Tanyan Ye, for her company, support, comfort, and encouragement throughout these past years. I thank my parents—Lesheng Li and Junying Chen—and my younger brother—Zhuo Li—for their selfless and unconditional support for my research and my life. I also thank my good friends in Newfoundland—Aedon Young, Marissa Farahbod, and Dr. Eleanor (Ellie) King—for their support and friendship. All of these people have kept me going, and this dissertation project would not have been possible without their support.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

Shared book reading is a common parent-child activity in home settings. *Dialogic reading* (DR) is a shared picture book reading approach that emphasizes verbal interaction between adults and young children (Huebner, 2000; Mol, Bus, De Jong, & Smeets, 2008; Whitehurst et al., 1988). There are three features of this reading approach: (a) it is evocative—it incites children to talk more in the reading process; (b) it is responsive—it encourages parents or other adults to give more verbal feedback to their children; and (c) it is progressive—it emphasizes that children's verbal development should be guided with scaffolding methods or occur within their zone of proximal development (Whitehurst et al., 1988).

Early literacy refers to the precursor stage of conventional literacy skills. The National Early Literacy Panel identified 11 early-literacy skills as being moderately related to later literacy achievement even when accounting for children's socioeconomic status (SES) and IQ (Lonigan & Shanahan, 2008). One early-literacy skill that has been widely studied is *oral language*—the ability to understand and use spoken language to communicate ideas (Harris & Hodges, 1995; Lonigan & Shanahan, 2008; Soifer, 2011). One aspect of oral language is *vocabulary knowledge*, the ability to understand and use words for communication (Agdam & Sadeghi, 2014; Vinco, 2013).

Numerous studies have found that DR can enhance children's vocabulary, especially with regard to their expressive vocabulary development (Crain-Thoreson & Dale, 1999; Cronan, Cruz, Arriaga, & Sarkin, 1996; Fielding-Barnsley & Purdie, 2003; Huebner, 2000; Lonigan, Shanahan, & Cunningham, 2008; Whitehurst et al., 1988). Most studies, however, focus on young children in Western countries, especially English-speaking countries. A small number of studies have examined the effect of DR on Cantonese-speaking children's receptive vocabulary, character recognition, and phonological awareness in that language (Chow & McBride-Chang, 2003; Chow, McBride-Chang, & Cheung, 2010; Chow, McBride-Chang, Cheung, & Chow, 2008; Fung, Chow, & McBride-Chang, 2005). Mandarin speakers comprise a much greater proportion of the Chinese population; however, the expressive vocabulary of Mandarin-speaking children has received little attention from DR researchers. Only three studies (Lin & Peng, 2014; Lin, Wu, & Su, 2005; Shiu & Tsai, 2016) have investigated the effects of DR on the vocabulary of Mandarin-speaking children in Taiwan, and these studies found positive effects. However, due to the methodological limitations of these studies, (small sample size, mix of study settings, or the descriptive nature of the study), causal relationships between DR and Mandarin-speaking children's vocabulary require further investigation.

Another aspect of oral language is *narrative competence*, which refers to the ability to understand and produce stories (Hipfner-Boucher, 2011; Lever & Sénéchal,

2011; Reese, Leyva, Sparks, & Grolnik, 2010; Trabasso & Nickels, 1992). Studies have found that narrative competence varies with age (Hutson-Nechkash, 2001; Paul, 2007; Speery & Speery, 1996). From ages three to nine, children's narrative language develops from the expression of simple ideas, such as labeling objects or yes/no responses, to stories containing complete plots. Studies have also found that children from families with high socio-economic status (SES) have an advantage over their peers from low-SES families with regard to narrative competence (Dickinson & Tabors, 2001; Heath, 1983; Lai, 2013; Vernon-Feagans, Hammer, Miccio, & Manlove, 2001). Other studies revealed that parent-child interactions during reading time or in other context (e.g., talking about past events) impact the development of children's narrative competence (Griffin, Hemphill, Camp, & Wolf, 2004; Caspe & Melzi, 2008; Nelson & Van Meter, 2007; Peterson & McCabe, 1992, 1994, 2004; Peterson, Jesso, & McCabe, 1999; Snow, 1983; Uccelli, Hemphill, Pan, & Snow, 1999; Vygotsky, 1978). Only a few studies have investigated the relationships between DR and English-speaking children's narrative competence (Lever & Sénéchal, 2011; Reese et al., 2010; Zevenbergen, Whitehurst, & Zevenebergen, 2003). Zevenbergen et al. (2003) reported that DR can significantly improve one aspect of children's narrative competence—the use of evaluative devices. Reese et al. (2010) suggested that adding elaborative reminiscing—asking children to talk about their past experiences—into DR can make DR more powerful in facilitating

children's narrative development. Lever and Sénéchal (2011) found that children in the DR group produced and retold stories that had more structure than those of their peers in the control group. One case study in Taiwan (Shiu & Tsai, 2016) explored the relationships between DR and the narrative competence of Mandarin-speaking kindergarteners in a school setting. The small sample size (five children) means that the causal relationship between DR and narrative competence among Mandarin-speaking children remains in question. Therefore, much is still unknown about the relationships between DR and narrative competence, especially in Mandarin-speaking children.

To summarize, it is well established in the literature that DR can enhance vocabulary development in English-speaking children, especially the development of their expressive vocabulary. Only a few DR studies, however, focus on Mandarin-speaking children. In addition, little is known about the relationships between DR and narrative competence. To address the aforementioned gaps, the present study aims to examine the effects of DR on the expressive vocabulary and narrative competence of Mandarin-speaking children.

Chapter 2: Literature Review

This chapter begins with discussion of the concepts of DR, vocabulary, and narrative competence. Next, research on the relationships between DR and vocabulary will be reviewed. Then, the relationships between DR and narrative competence will be examined. Finally, prior DR research with Chinese children will be reviewed with a focus on their vocabulary development and narrative competence.

2.1 The DR Approach: Background and Development

Shared reading, or *joint reading*, refers to an early-childhood practice whereby an adult reads a book to a child or a group of children (Harris & Hodges, 1995). In home settings, the adult may be a parent or another caregiver such as an elder sibling or grandparent. Early descriptive studies showed that shared reading is related to young children's literacy skills before they learn to read formally, especially to their oral-language skills (McCormick & Mason, 1984; Wells, 1985a). Moreover, the quality of shared reading is believed to play a more important role in enhancing children's language development than its quantity (Bus, van Ijzendoorn, & Pellegrini, 1995; Scarborough & Dobrich, 1994). One important feature of a high-quality joint-reading session is its interactivity, which type of shared reading have been defined by researchers as *shared interactive reading* (SIR; Towson, Fettig, Fleury, & Abarca, 2017).

Early investigations of SIR and child-directed speech in the 1970s and 1980s revealed many characteristics of parent-child verbal interactions. First, it was observed that parent-child reading provides a rich context for children's language development (Moerk & Moerk, 1979; Ninio & Bruner, 1978; Snow & Goldfield, 1983; Wells & Gordon, 1985). For example, Ninio and Bruner (1978) demonstrated that 75.6% of a mother's labelling activity happened in joint-book-reading time. Second, individual differences were found in parent-child interaction. Ninio (1980) observed the joint-book-reading experiences of 40 dyads from either lower- or middle-SES families with children aged 17 to 22 months. She found that lower-SES parents talked less than higher-SES parents during reading time and that lower-SES children, correspondingly, had a smaller productive vocabulary than higher-SES children. Third, it was reported that the sophistication of parent-child verbal interactions increased with children's age and language competence (DeLoache & DeMendoza, 1987; Moerk, 1985; Pellegrini, Brody, & Sigel, 1985; Wheeler, 1983). For instance, Wheeler (1983) observed the child-directed speech used by the mothers of 10 children aged 17 months to five years old. She found that as their age increased, children's verbal expression in joint book reading with their mothers changed from simple labelling to meaningful two-to-three-word utterances, to complete sentences associated with things outside the picture. Accordingly, the sophistication of mothers' verbal attempts increased; the focus of "motherese" changed

from single elements to multiple elements to interpretive comments about the picture book as children aged. DeLoache and DeMendoza (1987) also found that the language used by mothers was more sophisticated and demanding in joint book reading with their elder children compared to that with their younger children. Further, Pellegrini et al. (1985) suggested that it is children's language competence—not their age—that results in changes in the language demanding level of child-directed speech; that is to say, age is just a proxy for children's language competence in affecting parents' language use when interacting with their children. However, most of these early studies on SIR and child-directed speech are descriptive or correlational in nature and cannot be used to prove the causal relationships between parent-child reading and children's language development.

A major advance in addressing this void in existing knowledge occurred when Whitehurst et al. (1988) developed an interactive parent-child reading approach—DR—to test experimentally whether maternal picture-book reading could have an immediate effect on young children's vocabulary development. In developing the DR approach, Whitehurst et al. (1988) devised the approach based on several principles that earlier researchers had found effective in enhancing children's language. The first principle is that the parent-child interaction should be *evocative*; that is, encouraging children to speak more by asking questions. As suggested in Wells (1985b), the number of questions

asked by parents played a major role in the effectiveness of shared reading. Also, Ninio and Bruner (1978) found that the most labelling activities by children occur at the time when they initiated the dialogue cycle or when the mother used "what" questions. The second principle—the *responsive* principle—is that parents should be responsive during joint book reading, meaning that they should provide immediate and accurate feedback to their children as much as possible. Scherer and Olswang (1984) investigated the conversation of four parent-child dyads and found that when a mother added expansions to her speech, this was systematically related to an increase in their child's spontaneous imitation of vocabulary, which brought about an increase in their productive vocabulary. The third principle identified by Whitehurst et al. (1988) is *progressive*; that is, language difficulty is adjusted so as to be appropriate to the age of the child. As stated by Vygotsky (1978) in his social development theory and his theory of the zone of proximal development, children learn from their interactions with the social environment, and parents' fine-tuning of the level of their own language difficulty and of cognitive strategies they use can promote children's internalization of these skills and strategies. Similarly, according to Bruner's scaffolding theory (Wood, Bruner, & Ross, 1976), parents' tutoring behaviours can help children focus on the task at hand and keep them motivated in the learning process. An empirical study by Wheeler (1983) also showed that mothers tended to adjust their language so that it was age-appropriate for their children.

In sum, the early developers of the DR approach identified and implemented three primary principles: evocative, responsive, and progressive.

In addition to these principles, Whitehurst et al. (1988) also proposed a set of DR techniques, which were used by many succeeding studies (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Harkins, Koch, & Michel, 1994; Valdez-Menchaca & Whitehurst, 1992; Whitehurst, Arnold et al., 1994; Whitehurst, Epstein et al, 1994). To make the DR techniques easier to learn, Zevenbergen and Whitehurst (2003) summarized and perfected the set of techniques used. They summarized DR techniques in two sets—one for children of two to three years of age and another for four- and five-year-olds. The set of techniques for two- to three-year-old children can be taught to parents or teachers in two sessions. The first session comprises seven key techniques: (a) asking "what" questions to let the child name the objects or find specific information in the book, (b) following up on the child's answer with further questions on the topic, (c) repeating what child has just said to show your recognition of their answer and reinforce their oral expression, (d) providing help to the child if necessary, (e) giving praise and encouragement for the child's efforts, (f) following the child's interest rather than the exact picture/text sequence in the book to evoke more responses, and (g) always keeping the child's reading enjoyment in mind. The second session comprises three key techniques: (a) asking open-ended questions to prompt the child to produce and express their ideas about the story, (b) extending the

child's answer and encouraging them to do their imitation, and (c) similarly to session one, enhancing the child's enjoyment of the reading process.

The DR techniques for four-to-five-year-old children can be summarized by two acronyms, "PEER" and "CROWD." The acronym *PEER* refers to the DR procedure of prompt, evaluate, expand, and repeat. *Prompt* means to encourage the child to name the objects in the book or to talk about the story. *Evaluate* refers to the adult's evaluation of the child's response. *Expand* means to enhance the child's verbalization by enriching the information, such as by expanding children's response "Cat" to "A red cat," or developing their oral response, such as by expanding their answer "It cat" to "It is a cat." Finally, parents can urge the child to repeat the expanded answer, which is the final technique, *repeat.* The acronym "CROWD" represents the five types of prompting questions that parents can use in shared-reading interactions with their children. The first type is the *completion question,* whereby parents pause before the end of a sentence and then ask the child to use a word or two to complete the sentence. The second type is the *recall question*, whereby parents ask the child, in the middle of the story or at the end, what has happened so far. The third type is the *open-ended question* that has no simple answer and requires the child to formulate an appropriate one in their own words. An example of this type of question is "What do you think of the sheep's behaviour?" The fourth type is the *WH- question*, referring to questions containing "where," "what," "when," "who," and

"why." Questions beginning with "how" also fit into this question type. A WH- question is different from an open-ended question: a WH- question has a very specific answer that can be located in the book, while the answer to an open-ended question is usually not in the book. The last type of question is the *distancing question*, which asks the child to link the story elements to their real life; for example, "Have you ever eaten this kind of food?"

When teaching the DR techniques to the intervenors, such as parents, educators, or other professionals, the trainer often uses role modelling or videos to demonstrate the skills (as in Arnold et al, 1994; Whitehurst et al., 1988). In role-modelling, the trainer usually played the role of a parent and the training assistant played that of a child. Using role-modelling as the demonstration method was costly and not suitable for programs involving large numbers of participants; therefore, researchers developed modelling videos for the DR programs which were much less expensive and more efficient for training large groups and disseminating DR techniques. Arnold et al. (1994) compared the effect of video training to that of role-modelling in a DR study. Interestingly, for the group where intervenors were trained using video, the language outcomes of the children were significantly higher than those of their peers whose intervenors were trained using role-modelling. One possible reason is that a real parent and a real child demonstrated the techniques in the video while the role modelling by the trainer and the assistant simulated actual interactions; thus, the modeling quality of role modeling was lower than videos

(Arnold et al., 1994; Zevenbergen & Whitehurst, 2003). The video-training method was widely adopted by many researchers and has generated positive outcomes among children (Blom-Hoffman, O'Neil-Pirozzi, Volpe, Cutting, & Bissinger, 2007; Fielding-Barnsley & Purdie, 2003). For example, Blom-Hoffman et al. (2007) conducted a DR intervention with 18 parents and their preschool children in a daycare setting using video training. They found that children in the DR group talked more about the story than their peer in the control group, and also than when their own baseline data had been collected.

Apart from the comparison of modelling methods, a recent study (Beschorner & Hutchison, 2016) compared the effects of delivering DR skills online and face-to-face. It turned out that parents in the online training group acquired dialogic skills equivalent to those in the face-to-face group, and that their use of dialogue in joint-book reading increased to the same degree after the intervention. However, in contrast to video training, the effects of which have been verified in many studies, online delivery of training requires more replication studies that test its effectiveness in a DR program. In addition, as suggested by Beschorner and Hutchison (2016), factors impacting online delivery should also be studied further.

To date, most DR research has focused on its effects on children's oral-language skills, which have been usually measured with vocabulary tests (Rahn, Coogle, & Storie, 2016; Towson & Gallagher, 2014; Valdez-Menchaca & Whitehurst, 1992; Whitehurst,

Arnold et al., 1994; Whitehurst et al., 1988). This might be because the DR approach is designed to promote young children's language development (Whitehurst et al., 1988). Other aspects of children's oral language, such as narrative skills or verbal response have also been considered in a few studies (Fleury & Schwartz, 2017; Huennekens & Xu, 2010; Lever & Sénéchal, 2011; Zevenbergen et al., 2003). A number of other studies have looked at a wide range of emergent literacy skills, such as concepts about print, phonological awareness, alphabetic knowledge, orthographic awareness, or emergent writing (Fielding-Barnsley & Purdie, 2003; Mol et al., 2009; Pillinger & Wood, 2014; Whitehurst et al., 1999; Whitehurst, Epstein et al., 1994; Zevenbergen et al., 1997), but there are only a few studies on each of these different topics. Towson and her colleagues (2017) reviewed experimental or quasi-experimental studies and studies with a single-case design which were published in English language on peer-reviewed journals, which were related to DR. They found that 27 of the 30 studies reviewed reported a significant increase in children's language and/or emergent literacy skills. It should be noted that, except for vocabulary, only a few of the studies focused on other language aspects and emergent literacy skills, which challenges the validity of Towson's review. and so further research is needed. As the number of DR studies increased, some researchers turned their attention to the effects of DR on parent-child reading behaviours, children at risk (language-delayed, deaf or hard-of-hearing, from low-income families, or

with autism), home literacy environment, cognitive abilities such as concept information, the role of non-researcher as trainers/interventionists, or children's reading interest (Crain-Thoreson & Dale, 1999; Cronan et al., 1996; Fleury & Schwartz, 2017; Fung et al., 2005; Huebner, 2000; Niklas, Cohrssen, & Tayler, 2016; Pillinger & Wood, 2014). In sum, over time, the research focus of DR shifted from its effects on language development, especially vocabulary development, to its effects on the DR reading process or on non-language factors.

To summarize, DR is a shared-reading approach that underlines the importance of interactivity and is based on evocative, responsive, and progressive principles. For children from four to five years old, the DR procedure can be summarized with the acronym PEER—prompt, evaluate, expand, and repeat. Types of DR prompt can be summarized with the acronym CROWD—completion questions, recall questions, open-ended questions, WH- questions, and distancing questions. For children three years old and under, parents' prompt questions should be easier than those for older children, and simple labelling questions can be used more frequently. In an intervention, the DR techniques can be demonstrated using role-modelling method or video-training materials. It seems that DR programs can also be delivered online; however, this delivery model is still in need of further exploration. Up to now, most DR studies have focused on the effects of this approach on children's oral-language or print-related emergent literacy

skills, and most these have reported significant results. Recently, the research focus has shifted to the effects of DR on the reading process and on non-language variables.

2.2 Vocabulary

This section focuses on young children's vocabulary knowledge. Starting with an introduction to the definition and developmental process of vocabulary, this section will then present a review of the relationships between children's vocabulary and reading development. Besides, effective instruction methods of vocabulary will also be reviewed and summarized in the following.

2.2.1 Definition and Development

Vocabulary is the knowledge of word meanings that are known or used by human beings, or the ability to retrieve word meaning by its oral pattern or written form (Biemiller & Boote, 2006; Harris & Hoges, 1995). In the literature, many terms, such as "word knowledge," "vocabulary knowledge," "vocabulary," "word learning," and "word meaning knowledge," are often used interchangeably to refer to vocabulary (August & Shanahan, 2006; Biemiller & Boote, 2006; Harris & Hoges, 1995; National Institute of Child Health and Human Development [NICHD], 2000). In the present study, the term *vocabulary* will be used. Words that can be understood by listening or reading are termed as *receptive vocabulary*, and those that can be produced by writing or speaking as *expressive vocabulary* (Burger & Chong, 2011). Vocabulary is one important component of language

skills, and vocabulary measurement is often used as a proxy for assessing overall language development (Lonigan & Shanahan, 2008; Rudd & Kelley, 2011; Soifer, 2011).

Vocabulary or word knowledge can be complex, and the complexity is characterized by the incremental, polysemous, multi-dimensional, interrelated, and heterogeneous features of vocabulary development (Nagy & Scott, 2000). *Incremental* vocabulary development refers to the notion that word knowledge is not a matter of all or nothing but a matter of degree; the degree here refers to not only the degree of the number of words but also that of the depth of word knowledge. *Polysemous* vocabulary development focuses on the fact that one word can have more than one meaning. *Multi-dimensional* vocabulary development refers to the concept that knowledge of a word is not limited to its definition but also includes knowledge of its oral/written form, frequency of its use, and grammatical requirements. *Interrelated* vocabulary development addresses the fact that word meanings are not independent but related to each other. For example, the term "Earth" is a sub-concept of "planet." *Heterogeneous* vocabulary development refers to the idea that different types of words make different demands on understanding. For instance, "and" as a function word is much easier to understand than "atom."

The complexity of vocabulary has two implications for researchers and practitioners: (a) explicit instruction is not sufficient for vocabulary learning, and (b)

learning words implicitly in complex contexts is necessary. In addition, word knowledge is not only a matter of "what" but also "how." Procedural knowledge—how to use words—is an essential component of vocabulary knowledge.

Young children's vocabulary develops with age and across several stages. In infancy (birth to 18 months), children pick up phonologically simple and context-based words, such as the phrase, "Say bye-bye," which is acquired as a whole speech stream and links to their daily environment. At this stage, the word use is often over-extended. In other words, the word is often overly generalized to refer to other meaning-related objects or phenomena. "Say bye-bye," for example, might be used to refer to the disappearance of objects. It is neither easy for infants to decode and extract the sounds of single words, nor is it possible for them to grasp the precise and well-refined meaning of vocabulary at this stage (Carlisle & Katz, 2005; Soifer, 2011).

For children between 18 and 36 months, vocabulary size expands, and their understanding of word meanings become more well-refined and specific after encountering the words many times in different contexts. Children become more sensitive to the internal structure of words. Invented words, such as "storyer" or "spensive," appear at this stage. At the same time, *extra-contextual* language use with their parents—language dialogue that is not connected with the immediate environment—increases their extra-contextual vocabulary—words that are not connected

with the objects or events in children's immediate environment. The characteristics of this stage are vocabulary growth in size, further precision of meaning, and sensitivity to the internal structure of words and to extra-contextual language (Carlisle & Katz, 2005; Snow, 1983; Soifer, 2011; Wasik & Bond, 2001).

As children progress beyond the age of three, the increase in the size of vocabulary, including extra-contextual vocabulary, and the refinement of word meanings continue and accelerate. Their invented vocabulary also evolves to conventional spelling as they grow older (Carlisle & Katz, 2005; Dikinson & Snow, 1987; Sénéchal, Ouellette, & Rodney, 2006; Soifer, 2011; Wasik & Bond, 2001).

Apart from age, three factors that impact vocabulary development—SES, social support, and gender—are also well documented in the literature. In the literature, SES is typically defined as family income, parental education, and/or occupation of the head of household, and the age of child participants (White, 1982). Children's vocabulary varies across socio-economic backgrounds (Baker, Simmons, & Kame'enui, 1998; Fenson et al., 1994; Graves, Brunetti, & Slater, 1982; Gravers & Slater, 1987; Hart & Risley, 1995; Hoff, 2003; Lai, 2013). Hart and Risley (1995), for example, investigated the language use of 42 Midwestern American families and found that there was a strong vocabulary gap between children from different SES backgrounds, and that the results favoured children from professional families rather than low-income families. Graves et al. (1982)

found the vocabulary size of fifth graders from higher-SES backgrounds to be twice as large as that of their peers from lower-SES backgrounds. Hoff (2003) tested the vocabulary growth rate of 33 children from high-SES families and 30 children from low-SES families across a 10-week period and found that high-SES children's vocabulary increased more rapidly than that of their peers from low-SES families.

A second factor—*social support*—refers to the environmental supporting factors for young children's early language, which is deemed as almost the most influential factor in the early development of vocabulary by many researchers. In a literature review, Hoff (2006) identified 18 social factors that can affect children's language development including vocabulary. These factors cover aspects of home, school, and community settings, such as birth order, cultural aspects, school, child-directed speech by parents, and joint attention. Notably, SES was categorized as one subset of home social factors in Hoff's study (2006). Differently from Hoff, the present study has defined SES as a factor independent from social-support factors, although the two could be closely connected. The studies of Whitehurst and Lonigan (1998) and Sénéchal and LeFevre (2002) showed that the home literacy environment contributes to young children's vocabulary development. Huttenlocher (1991) found that the variations in early vocabulary growth are largely related to the amount of parent speech instead of to children's ability. Pan, Rowe, Singer, and Snow (2005) found that differences in parental speech and level of

language and literacy skills are related to growth differences in their children's expressive vocabulary. Weizman and Snow (2001) analyzed 263 sessions of parent-child conversations from 53 low-income families and found that the density of sophisticated words in parental speech had more influence than the total amount of speech on children's vocabulary performance from kindergarten to Grade two.

Gender also plays a role in the development of vocabulary (Bauer, Goldfield, & Reznick, 2002; Fenson et al., 1994; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Reznick & Goldfield, 1998). Bauer et al. (2002) assessed the vocabulary development of 26 children at two-month intervals from the age of eight, nine, or ten months old through to the age of 14 months and found that the vocabulary growth rate of girls was faster than that of boys. Fenson et al. (1994) reported that girls performed significantly better than boys on receptive and expressive vocabulary and also achieved higher composite scores on two types of vocabulary. Huttenlocher et al. (1991) reviewed the literature and found that gender differences were constantly found among children under two years old. These authors also found that there was a significant interactive effect between children's age and gender on vocabulary development, especially before the age of 20 months, and that girls' vocabulary increased earlier than that of boys. Thus, it is important to consider SES, social factors, and gender when investigating the vocabulary development of young children.

2.2.2 Vocabulary and Reading / Language Development

Vocabulary is central to language comprehension and plays an important role in reading development (Cunningham & Stanovich, 1997; Perfetti & Adolf, 2012; Storch & Whitehurst, 2002). Perfetti and Adolf (2012) posited that the quality of one's vocabulary knowledge—whether the knowledge is accessible, precise, and flexible—determines the success of language comprehension. Reading entails both *code-related skills* to decode and encode words and *meaning-oriented skills*—language comprehension—to understand text meaning (August & Shanahan, 2006; Gough, 1996; Gough & Tunmer, 1986; Hoover & Gough, 1990). Therefore, vocabulary plays an important role in the reading process.

Many correlational studies have shown that individuals' vocabulary is related to their reading ability (Cunningham & Stanovich, 1997; Davis, 1944, 1968; Muter, Hulme, Snowling, & Stevenson, 2004; Nation & Snowling, 2004; Perfetti, 2007; Sénéchal et al., 2006; Singer, 1965; Storch & Whitehurst, 2002). For example, Singer (1965) tested the relationship between vocabulary and reading comprehension of over 250 third- to sixth-grade pupils; the results showed that vocabulary accounted for significant variance in reading achievement. Davis (1968) explored the component skills of reading comprehension and found that vocabulary made the largest contribution to variance in reading comprehension. Nation and Snowling (2004) found that the expressive vocabulary of 72 children at the age of 8.5 years can predict their concurrent

reading-comprehension and word reading skills and also those skills 4.5 years later. Sénéchal et al. (2006) showed that the depth of children's vocabulary knowledge in Grade four predicts their concurrent reading comprehension. To find a causal link between vocabulary and reading development, some researchers tested the effects of vocabulary intervention on children's reading comprehension and observed positive effects (Beck, Perfetti, & Mckeown, 1982; Draper & Moeller, 1971; Kame'enui, Carnine, & Freschi, 1982; Margosein, Pascarella, & Pflaum, 1982). It should be noted, however, that some studies found that vocabulary intervention did not improve children's reading comprehension (Jackson & Dizney, 1963; Jenkins, et al., 1978; Tuinman & Brady, 1974). This might be because reading comprehension is a comprehensive and complex cognitive activity and requires many skills other than vocabulary. Or, to affect achievement in reading, vocabulary interventions need to be comprehensive and long-term (Nagy, 2005). Generally speaking, up to the present, the empirical evidence for causality is relatively weak compared with that for correlational relationships (Baumann, 2009; Baumann, Kame'enui, & Ash, 2003).

In addition to its centrality in language comprehension and reading comprehension, vocabulary also makes it possible for one to produce language (Nation, 2001; Schmitt, 2000; Widdowson, 1989). Widdowson (1989), in his review, pointed out that the ability to use language in oral or written form is integrated firmly with language

knowledge, which includes vocabulary knowledge. Nation (2001) asserted that vocabulary makes language use possible and that language use, in turn, can increase vocabulary. Schmitt (2000) argued that vocabulary is central to the acquisition of spoken language and the learning of a second language. Thus, vocabulary determines how one can communicate effectively to comprehend others and to express oneself with oral language.

2.2.3 Vocabulary Support and Instruction

Given the importance of vocabulary, it is important to offer effective vocabulary instructions for children. The central question here is "What is good vocabulary instruction?" The principles of effective vocabulary instruction are well documented in the literature (Beck & McKeown, 2007; Biemiller & Boote, 2006; Blachowicz & Fisher, 2000; Carlisle & Katz, 2005; Nagy, 2005; NICHD, 2000; Penno, Wilkinson, & Moore, 2002). First, effective vocabulary instruction should allow children to engage actively: interactive instructional techniques such as immediate feedback or vocabulary discussion can be used to achieve this goal. Second, vocabulary instruction should be personalized so that each child makes connections with the vocabulary being learned. Obviously, this can promote the internalization process of word knowledge. The third principle is learning in context. The complexity of word knowledge means that vocabulary learning can only be effective when a word's meaning is clearly demonstrated in the application

scenario or context where it is used. Thus, it is necessary to immerse children in a rich vocabulary learning context. Fourth, full acquisition of a word requires repetition, which means encountering the word repeatedly in different contexts. Thus, good vocabulary instruction should provide varied contexts for children to learn the target word. Fifth, vocabulary instruction should be long-term. This is because word knowledge is multi-dimensional and polysemous, and internalization requires repeated encounters with a word in different contexts, which takes time. Also, only large enough vocabulary growth can make obvious and significant difference in children's reading. Sixth, effective vocabulary instructions can be direct (e.g., giving explanation of vocabulary definitions) or indirect (e.g., reading stories), and using multiple instructional methods is better than using a single method. Finally, several other effective components can be included in a vocabulary instruction, such as computer-assisted or multi-media learning materials, effective vocabulary assessment, and clarification of learning tasks.

These principles have led to many vocabulary instruction methods being devised and used in research and practice. Few of them address vocabulary only; most are text- or reading-based (Biemiller & Boote, 2006; Carlisle & Katz, 2005; Nagy, 2005; NICHD, 2000). *Vocabulary-only* instructional methods refer to instruction focusing on individual words or word groups without the aid of reading context. For example, the *keyword method*, in which the sound of a word is converted into a concept and then connected to

the real meaning of the word with visual aids, focuses on a single word at a time. Another example, *semantic mapping*, in which the targeted word and its relations with other words are presented visually in a graphic, focuses on interrelated word groups. *Text-based* methods usually use stories or expository articles as the context for vocabulary learning. One example is the text talk method, in which teachers select several vocabulary words that students might be not familiar with in advance and then give explanations and examples to students during the reading process. Along with these two types of instructions, generative vocabulary knowledge or word meaning derivation strategies such as morphological or structural analysis, contextual analysis, or word-feature analysis, are usually taught during the instructional process. For instance, when teaching context analysis in text-based instruction, teachers instruct students to locate contextual clues that can help them derive word meanings from the text.

In summary, vocabulary refers to knowledge of word meanings that can be retrieved in the process of communication. The development of young children's vocabulary extends from contextual to extra-contextual, from over-extended to well-refined, from invented to conventional, and from small in size to large and rapidly increasing. Moreover, the development of vocabulary varies with SES backgrounds, gender, and level of social support. In addition, vocabulary is central to language comprehension and production, which ultimately impacts reading achievement and

oral-language development. Finally, many principles of effective vocabulary instruction have been identified in the literature, such as active engagement, repetition, personalized learning, and contextualization. Based on these principles, various vocabulary instructional methods have been constructed that either focus on vocabulary only or embed vocabulary learning in a text. Strategies for deriving word meaning are often taught along with these two types of instruction.

2.3 Narrative Competence

This section begins with an introduction to the definition and the classification system of narrative as a noun and the definition of narrative competence. Next, it will move on to the topic of the relationships between children's narrative competence and later reading development. The third part will turn to review the indicators and measurement of narrative competence. Towards the end, the impacting factors and the intervention methods of children's narrative competence will be reviewed and summarized.

2.3.1 Narrative and Narrative Competence: Definitions and Taxonomies

To understand narrative competence, the meaning of the noun "narrative" should first be explored. According to Bruner (1986), *narrative* is a way of knowing or a mode of thought, connected to human intentionality and goal-directed actions. More simply, narrative is a kind of cognitive functioning about what a human wants and how they achieve what they want. Narrative is also regarded as a way of discourse, a verbal

description of past events in temporal sequence (Hoff, 2013; McCabe & Peterson, 1991; Stein, 1997). In addition, Richardson's (1990) definition combined these two aspects, conceiving narrative as a way of reasoning and of representation that human beings use to understand and express the world. Similarly, another definition characterized narratives as accounts of past events stored in memory, constructed by cognition, intentionally selected by the teller or the writer before transmitting to the audiences or readers, and organized in a knowledge structure that can be anticipated by these audiences or readers (Graesser, Golding, & Long, 1996; Heath & Branscombe, 1986). From these definitions, it can be concluded that narrative possesses both cognitive and verbal qualities. To be clearer, narratives are expressions of past events in temporal sequence, the construction and comprehension of which are deeply rooted in human cognitive processes.

Narrative, as expressions of past events, can take different forms. Heath and Branscombe (1986) divided narratives into four types: recounts, accounts, eventcasts, and stories. *Recounts* refer to a verbal description of a series of events with a scaffolding prompt provided by the teller or a third party, such as a child's retelling of a well-known story or an adult's retelling of a folktale he heard in his childhood. *Accounts* also are verbal description of a range of events but without scaffolding prompts, such a child's description of what they did in the afternoon. *Eventcast* refers to the verbal description of the ongoing activities of an event, such as a child's explanation of his cookie-baking

process. *Stories*, or fictional stories, usually involve a main character who goes through a series of events and takes many actions to achieve some goal. In later research, Hedberg and Westby (1993) added another type of narrative to the classification system—s*cripts*, meaning events that regularly happen or reoccur in daily life, such as getting dressed in the morning or a bedtime-story routine before sleep.

Another widely accepted taxonomy of narratives was devised by Hudson and Shapiro (1991). They categorized narrative as scripts, personal narratives, and stories. Similar to the definition of Hedberg and Westby (1993), *scripts* refer to what usually happens as a daily routine. *Personal narratives* are a child's descriptions of the events they personally experienced in the past. *Stories* refer to imaginative or fictional tales that usually contain a plot and essential story elements such as a formal beginning, an orientation (the introduction of the characters and the setting), initiating event (the event that sets off the rest of the events), a problem or a goal, resolution of the problem, and a formal ending.

Some researchers employ a simple dichotomy of narratives: i.e., narratives can be *autobiographical*—recounts of personal experiences—or *fictional*—narratives of story events in a book (Chang & McCabe, 2013; Hipfner-Boucher, 2011; Lever & Sénéchal,2011).

Some researchers in recent years simply and conceptually defined narrative as stories, and narrative competence as the ability to either comprehend or produce stories (Hipfner-Boucher, 2011; Lever & Sénéchal, 2011). These two definitions are used in the present study. Moreover, children's fictional narratives, or stories, are the focus of the present study because autobiographical narrative is easily influenced by cultural bias, which might be a confounding variable in the measurement of children's real narrative ability (Heath, 1983; Hipfner-Boucher, 2011; Pellegrini & Galda, 1993).

2.3.2 Developmental Changes

With regard to developmental changes in children's narrative skills, Hudson and Shapiro (1991) suggested that different narratives—scripts, personal narratives, and stories—have different developmental trajectories. Because of the focus on fictional narrative of the present study, the developmental changes in children's fictional narratives will be emphasized in the following section.

One well-established developmental framework, which identified six developmental levels for children's story narratives, was constructed by Applebee (1978) and further extended and interpreted by many succeeding researchers (Hedberg & Stoel-Gammon, 1986; Hutson-Nechkash, 2001; Klecan-Aker & Kelty, 1990; Klecan-Aker & Swank, 1987; Westby, 1982, 1984). The development of story narratives involves two essential processes: *centring*—forming a central idea that runs through the whole

story—and *chaining*—forming a logical or temporal order that links events together (Applebee, 1978). There are two ways or principles for linking events / objects in a story to form a central idea—similarity or complementarity (Applebee, 1978; Hedberg & Stoel-Gammon, 1986). Centring by *similarity* means the events / objects in the story are linked by their common attributes. For example, the story "Giraffe is running. Turtle is running. Rabbit is running. That is all." forms its central idea from the common attribute "is running." Centring by *complementarity* means that the events / objects in the story are combined on the basis of abstract underlying complementary relations. In the story "The elephant yelled at me. I cried. My mom ran out and asked me, 'What happened?' My brother saw this. He got angry with the elephant." the central idea is "elephant yelling at me", and all the events are derived from this centre and complement each other. Centring appears earlier than chaining, and similarity earlier than complementarity in the development of children's story narratives.

Apart from the two essential processes and principles, children's fictional narrative develops across six stages (Applebee, 1978). From birth to two years old, children typically only give yes/no responses, label objects, or express simple ideas with short, plain sentences or sentence fragments (Hedberg & Stoel-Gammon, 1986; Hutson-Nechkash, 2001; Klecan-Aker & Kelty, 1990; Klecan-Aker & Swank, 1987; Westby, 1982, 1984). This stage is called the *heaps* story stage because children's

expressions are idea heaps at this stage, only a “pile” of ideas arranged in an arbitrary way.

Between the age two and three, children’s narrative competence shifts from heaps to a *sequence* stage. At this stage, children can describe the events around a theme and a character, but the sequence of events is not always correct. The similarity principle is employed at this stage to link the statements.

At around three to four years old, children can tell stories including a character, a setting, a topic, or a central theme, and cause-effect logic emerges in the story. Children begin to use the complementarity principle to link events and objects. This stage is called *primitive narrative* because all essential elements are not included in the story.

Between the ages of four and five years, there is a short period during which chaining appears in children’s stories. Centring diminishes or even disappears temporarily because children spend more cognitive energy on forming the chain while continuously shifting from one event to another. Children usually go through this stage, called *unfocused chain* because of the absence of centring, quite quickly.

At around five years, children find a balance between chaining and centring, and the two appear together in their story narratives. In this *focused chain* stage, children’s stories become more logical, and more story attributes, such as character motivation and abrupt endings, are added. An *abrupt ending* means that an unnatural ending statement

appears at the end of the story, i.e., it is still difficult for the listener to interpret the relationship between this type of ending and the beginning of the story.

Between five and seven years old, children' stories usually contain a complete plot—a problem identified at the beginning and solved at the end—in a logical event sequence. Additionally, the goals and motivation of the character are connected to the plot, and the relationship between the beginning and the ending is established and can be interpreted by the listeners. This stage is called *true narrative*, which means that the narrative is mature and seems quite like adult narratives, but this does not mean that the development of narrative skills is complete. Children's narrative skills will continue to develop into their adolescent or even adult years (Nippold, 2016).

2.3.3 Relationship to Later Reading Development

Numerous studies have found that good readers perform better in story production and story comprehension than poor readers or reading-disabled readers (Cameron, Hunt, & Linton, 1988; Feagans & Short, 1984; Geva & Olson, 1983; Klop, 2011; Oakhill, 1984; Roth & Spekman, 1986; Roth, Spekman, & Fye, 1995). For example, Klop (2011) found that children from groups with normal reading comprehension significantly outperformed their peers from groups that exhibited low reading comprehension or that had reading comprehension deficits with regard to narrative skills. Feagans and Short (1984) showed that the content and complexity of the story production of children with language

impairments were lower than that of children whose reading development was progressing normally. Geva and Olson (1983) found that good readers, at age six, tended to use more developed strategies to retell the story, that more causal links appeared in their retelling, and that they were able to adjust their retelling according to the listeners' knowledge of the story. Cameron et al. (1988) replicated their study with 32 seven-year-old children and observed similar results.

It also has been claimed that young children's narrative competence is a good predictor of their later reading development and academic success (Feagans & Appelbaum, 1986; Griffin et al., 2004; Paul & Smith, 1993; Roth, Speece, & Cooper, 2002; Snow, 1991; Snow, Tabors, Nicholson, & Kurland, 1995; Snyder & Downey, 1991; Stephens, 1988). For example, Feagans and Appelbaum (1986) demonstrated that children with good narrative skills had better scores in reading-recognition and reading-comprehension tests and that this advantage lasted for the whole research span (three years). Griffin et al. (2004) found that the narrative and expository discourse of five-year-old children can predict their reading-comprehension performance at eight years old. Paul and Smith (1993) found that a delay in children's narrative skills at age two persisted until the late preschool period, putting them at risk of future academic difficulties such as learning disabilities. Roth et al. (2002), however, found that children's narrative skills in kindergarten contributed to the variance in their reading comprehension

in Grade 1, but not in kindergarten and Grade 2. In fact, the bivariate correlation *r* between kindergarten narrative skills and Grade 1 reading comprehension was negative. These researchers explained that the types of elicitation tasks and coding procedures used might have affected the results. They also posited a second possible reason—that narrative skills might be more related to the reading comprehension of children in middle or high grades than in Grades one and two. This, they argue, is because children in in lower grades are still at the stage of learning to read, when decoding skills are being consolidated. In other words, they have not reached the stage of reading to learn, when reading comprehension plays a critical role. Snyder and Downey (1991) to some extent confirmed this second explanation by showing that young children's narrative competence significantly explained the variance in the reading comprehension of eight- to eleven-year-old children's reading comprehension and that the explanatory power was even larger for children of 11 to 14 years of age. In sum, the narrative skills young children make an ever-increasing contribution to their later reading comprehension.

The theoretical basis for the relationship between narrative and reading lies in the literate language nature of narrative skills. The development of language is believed to be a continuum, one end of which is oral language and the other end literate language, so that one's language develops from oral to literate (Westby, 1985; Snow, 1983). *Oral language* is defined as conversational language, emphasizing the process of learning to

speak with phonological, morphological, syntactic, or pragmatic knowledge (Westby, 1985). *Literate language* is conceived as book-like language, underlining the process of speaking to learn and the process of monitoring and reflecting experiences using language (Westby, 1985). One prominent property of literate language is its *decontextualized* nature (Greenhalgh & Strong, 2001; Paul, 2001; Snow 1991, 1983; Snow & Dickson 1991). Being decontextualized means that this type of oral discourse is content-extended, syntactically complex, and removed from the immediate environment (Paul, 2001; Snow, 1983). *Content-extended* means that the decontextualized language contains all the information that is necessary for comprehension, in contrast to the abbreviatory feature of conversational oral language. Decontextualized language is ubiquitous in teachers' instructional language, and therefore children who are skillful in understanding and using this type of language can better match their teacher's expectations and gain more from classroom learning (Wallach & Butler 1994; Snow 1983). In addition, the properties described in the definition of decontextualized language are also those of the language used in the written text, and children having good decontextualized language skills can understand more in their reading and writing process (Snow 1983). From this perspective, literate or decontextualized oral discourse serves as a bridge between oral language and formal reading and writing and benefits children's oral communication in an academic environment (Greenhalgh & Strong, 2001; Snow 1983). Narrative, as a noun here, is one

important type of literate language and is predominately decontextualized (Greenhalgh & Strong, 2001). Therefore, narrative competence can function as decontextualized language and contribute to children's reading achievement and academic success.

2.3.4 Indicators and Measurement

Narrative competence is sophisticated and comprehensive; therefore, its measurement comprises multiple capacity indicators. In a review of research studies on narrative—their focus and the measures used—Justice et al. (2006) found that narrative skills are usually evaluated from two aspects: macrostructure and microstructure. *Narrative macrostructure* refers to the overall story structure—the temporal-causal sequence in which events are bound into a story. *Narrative microstructure* refers to the quality of language used in the narrative, including linguistic features like story length, cohesion, and syntactic organization. Many researchers believe that narrative evaluation is also an important aspect of children's narrative competence (Bamberg & Damrad-Frye, 1991; Labov, Cohen, Robins & Lewis, 1968; Peterson & McCabe, 1983). *Narrative evaluation* refers to the linguistic devices embedded in the narrative that explains why some events happen and what the point of the narrative is (Labov, 1972). A detailed review of these three aspects will be presented in the following sections. It should be noted that the three aspects are not always in a complementary relation to each other. Researchers have found that the capacity to produce and comprehend narrative is finite, i.e., spending too many

cognitive resources on one aspect might lower the narrator's performance on other aspects (Curenton & Justice 2004; Justice et al., 2006; Wang, 2018). This indicates that measuring one aspect only cannot reveal the panorama of narrative competence and that a comprehensive evaluation of all three aspects is necessary.

2.3.4.1 Macrostructure

The first aspect of narrative competence that is often evaluated in the literature is macrostructure, which is often measured from three perspectives: story grammar, episodic complexity, and coherence. The first facet, *story grammar*, refers to the rewrite rules of story information and the elements of the episodic structure, such as setting, characters, and the outcome (Peterson & McCabe, 1983; Stein, 1978; Stein & Glenn, 1975). Research on story grammar can be traced back to Bartlett's (1932) study, which showed that the study participants employed an inner psychological schema to reconstruct the story information in their retelling. Later, Propp (1958) identified the *morphology*, or the relative fixed patterns and elements, of folktales. Based on these studies, Rumelhart (1975) analyzed the information contained in folktales, fables, and myths and proposed a story-grammar system to describe the underlying representation structure in such stories. Due to the complexity and the difficulty of use of Rumelhart's (1975) system, certain investigators proposed modifications to the story-grammar system (Mandler & Johnson, 1977; Stein, 1978; Stein & Glenn, 1979; Westby, Van Dongen, & Maggart, 1989).

This revised system, which includes the information categories of setting, initiating event, internal response, initial plan, attempt, consequence, and reaction was widely used in subsequent studies (Hayward, Schneider, & Gillam, 2009; Lever & Sénéchal, 2011; Schneider, Dubé, & Hayward, 2005). Also, story grammar has often been measured by tallying the information categories included in the story (Peterson & McCabe, 1983; Stein & Glenn, 1979). The following example illustrates how the information categories can be incorporated into a story:

> The dog and the cat were good friends. (*Setting*)
> The dog was playing a ball. (*Initiating event*)
> The cat liked the ball. (*Internal response*)
> The cat wanted to play with the ball together with the dog. (*Internal plan*)
> The cat asked the dog whether he could play. (*Attempt*)
> The dog said no. (*Consequence*)
> The cat was very sad. (*Reaction*)

Episodic complexity, or episodic level, is the second facet of children's narrative macrostructure. Certain researchers have divided information constituents of story grammar into two categories: (a) setting and (b) episode (all elements of story grammar except setting) (Davies, Shanks, & Davies, 2004; Merritt & Liles, 1989; Peterson & McCabe, 1983; Rumelhart, 1977). An episode in a story can be *complete* (including all essential story grammar constituents) or *incomplete* (including only some of them). A story can comprise more than one episode, and some researchers have used the complexity and the number of episodes as indicators of the level of story macrostructure,

or episodic complexity (Davies et al., 2004; Hayward & Schneider, 2000; Hughes, 1997; Merritt & Liles, 1989).

As suggested by its definition and in contrast to story grammar, which focuses on the inclusion of structural information units, episodic complexity places more emphasis on how well these structural elements are organized. To measure "how well," Hughes (1997) proposed a framework for rating episodic complexity. This framework divided episodic complexity into six levels, from simplest to most complex: (a) descriptive sequence, (b) action sequence, (c) reactive sequence, (d) abbreviated episode, (e) incomplete episode, and (f) complete episode. At the *descriptive sequence* level, the narrative describes the temporally related sequence of events and describes characters, settings, and habitual actions only, without any causal relations. At the *action sequence* level, the action events of a character are usually ordered temporally but causal connections between events still do not appear. At the *reactive sequence* level, the causal connections between characters' actions appear, and one action will lead to another action; however, there is no plan for the series of actions, and these actions are not goal-directed. At the *abbreviated episode* level, the goal of the character is well expressed, but the planning process to achieve the goal is still not obvious and must be inferred. At the *incomplete episode* level, all essential story information units are not included, i.e., certain of the following elements—initiating event, internal response, initial plan,

response, attempt, consequence, and reaction—are missing. At the final level, *complete episode*, the narrative includes both a temporally and causally related sequence of events, all the essential story-grammar constituents, and a clear goal of the character and planning.

The third facet of macrostructure that has been widely examined in the literature is *coherence*, which refers to the meaningful connectiveness of story events constructed on the basis of proper temporal and causal relations (Cain, 2003; Christiansen, 1995; Glosser & Deser, 1991; Hoff, 2013; Hudson & Shapiro, 1991; Shapiro & Hudson, 1991). Unlike story grammar and episodic complexity, coherence focuses more on the semantic appropriateness or thematic unity of the overall structure (Cain, 2003; Christiansen, 1995; Glosser & Deser, 1991). Glosser and Deser (1991) further divided coherence into two types: global and local. *Global coherence* is defined as the degree to which the narrative conforms with the universal coherent story grammar, while *local coherence* is the degree to which the narrative conforms with the specific theme of the story.

Coherence can be measured by the episodic complexity of the story (Cain, 2003), the degree of inclusion of story-grammar constituents (Davies et al., 2004; Diehl, Bennetto, & Young, 2006; Shapiro & Hudson, 1991), evaluation of its local and global coherence (Glosser & Deser, 1991), or analysis of the causal connections in a story and its causal chain (Diehl et al., 2006). *Causal connections* refer to the causal relationships

between story events, and a *causal chain* refers to a series of events that are causally connected in the story.

Since coherence can be measured by the level of episodic complexity and the inclusion of story grammar constituents, the implication is that the three facets of macrostructure overlap and are closely connected. Cain's study (2003) confirmed this by showing that children with poor performance in story grammar differed in the types of connectives they used. It should be noted that certain researchers have uncovered a factor that may confound measurement results (Cain, 2003; Cain & Oakhill, 1996; Shapiro & Hudson, 1991): children sometimes cannot produce a coherent story, not because of the poor story conception in their mind but due to the high cognitive demands of other aspects of telling a story, e.g., memory. These researchers suggested that the use of pictorial aids can relieve such cognitive demands and help children to present a coherent story.

2.3.4.2 Microstructure

The second aspect of narrative competence is microstructure, which is also evaluated from three perspectives: productivity, syntactic complexity, and cohesion (Hipfner-Boucher, 2011; Wang, 2018). With regard to microstructure, *productivity* refers to the length and diversity of story language and is usually measured by the total number of words (TNW), the total number of different words (NDW), the total number of T-units

(TNT), and lexical density (LD) (Justice et al., 2006; Wang, 2018). As suggested by its name, TNW is calculated by counting the number of words in a narrative sample, and NDW by counting the number of the novel words in the sample. LD is usually measured by the ratio of content words to TNW, or, the ratio of some types of content words, such as verb, adverb, or adjective, to TNW. To analyze a narrative, researchers usually first parse the narrative sample into smaller pieces or T-units, also called "individual propositions," "individual statements," or "C-units" in the literature.

A *T-unit* is comprised of an independent clause and other dependent clauses or phrases, and it usually has only one main subject and one main verb (Hunt, 1965; Justice, Bowles, Pence, & Gosse, 2010; Justice et al., 2006; Wang, 2018). For instance, the sentence, "Giraffe cried when he saw his plane in the water" is one T-unit; "giraffe cried" is the independent clause, and "when he saw his plane in the water" is the dependent clause. Another sentence, "Giraffe saw his plane in the water, and giraffe cried" comprises two T-units because both are independent clauses, each containing a main subject and a main verb. A T-unit is usually a complete sentence, and thus TNT is usually the total number of individual sentences in a narrative sample.

It should be noted that different indicators might weigh differently in measuring productivity. Justice et al. (2006) conducted a factor analysis and then devised a formula for calculating a score for narrative productivity. The formula reflected their finding that

the TNW, NDW, and TNT contributed differently to narrative productivity. Uccelli & Páez (2007) also found that NDW was a more sensitive measure than TNT when measuring English narrative productivity.

With regard to developmental changes, certain researchers found that children's narrative productivity increases with their age and language competence (Curenton & Justice, 2004; Justice et al., 2006). Curenton and Justice (2004) found that the number of T-units in a group of three-year-old children was significantly lower than that in a group of five-year-olds. The mean number of T-units in the four-year-old group was higher than that in three-year-old group but lower than that in five-year-old group; however, the number of T-units in the four-year-old group did not differ significantly from the other two groups. It could be that what occurs at age four is a transitional phase between what happens at ages three and five. Justice et al. (2006) collected narrative samples from 250 5-12-year-old children and found that their fictional narrative productivity increased with age and peaked at the age of 10. Nonetheless, some studies reported that narrative productivity does not increase with age (Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003; Shapiro & Hudson, 1991). Justice et al. (2006) found that children's narrative productivity decreased slightly from age 11 to age12 after the peak at age 10. The authors ascribed this phenomenon to motivational or testing reasons rather than to a real decrease in productivity related to narrative competence: older children might have been

uninterested in telling elaborate fictional stories, or the assessment materials might not have captured the highest levels of their narrative productivity. Shapiro and Hudson (1991) also found that the number of T-units in a group of four-year-olds did not differ significantly from that in a group of six-year-olds. Muñoz et al. (2003) compared the TNT, NDW, and TNT of children below five years with those of children above five years. They found that the three measures did not differ significantly between the two groups, although children in the over-five-year-old group did have a higher TNT. One possible reason for the non-significant results is that the age difference between the two groups was not large enough to see a significant difference in narrative productivity. The other possible reason is that the development of other aspects of narrative competence compromises children's performance in narrative productivity, given that the cognitive capacity to complete linguistic tasks is finite (Justice et al., 2006). In summary, narrative productivity generally increases with children's age, but the increase is not steady and might decrease with age, possibly due to factors such as decreased motivation, testing issues, or the development of other aspects of narrative competence.

The second facet of narrative microstructure is *syntactic complexity*, meaning the complexity of the sentence structure in the narrative (Justice et al., 2006; Wang, 2018). It is usually evaluated by the mean length of utterance in words (MLU-w), the mean length of utterance in morphemes (MLU-m), the mean length of the five longest utterances in

words (MLU5), the number of complex T-units (COMT), or the proportion of complex T-units (PROCOMT). An *MLU-w* score is obtained by calculating the ratio of words to utterances, while *MLU-m* is the ratio of morphemes to utterances (Parker & Brorson, 2005). A *morpheme* is the smallest semantic unit of language that can stand alone to convey meaning (Farrell & White, 2018). One morpheme can be a whole word, such as "cat," and it also can be part of a word, e.g., "dogs" has two morphemes, "dog" and "-s"; the former referring to the animal and the latter to the plural status of the dog. Justice et al. (2006) found that MLU-w and MLU-m are highly correlated. Due to differences between the two languages, morphemes in an utterance are easily segmented in English but with difficulty in Chinese; therefore, researchers often use the *mean length of utterance in characters* (MLU-c) or MLU-w to analyze Chinese narrative samples instead of MLU-m (Wang, 2018). However, one researcher (Cheung, 1998) have posited that MLU-w is a better index than MLU-c in measuring Chinese narrative syntactic complexity because the basic developmental units of Chinese language in early childhood are words rather than characters. A *MLU5* score is obtained by dividing the total number of words /morphemes in the five longest utterances by five (Zhou & Zhang, 2009). The MLU5 score is usually higher than but also positively correlated with that for MLU (MacWhinney, 2000; Pan, 1994). Some studies have shown that MLU5 is more sensitive in capturing the development of syntactic complexity than MLU. This is especially the

case when children are above four years in age (Zhou & Zhang, 2009) because individual variation in MLU increased greatly after they reached the age of four (Eisenberg, Fersko, & Lundgren, 2001), and also because increases in MLU scores after the age of four turned out to be too gradual to indicate children's linguistic changes (Cheung, 1998). A *COMT* score is obtained by calculating the total number of complex T-units which comprises of one independent clause and at least one dependent clause (Hipfner-Boucher, 2011; Justice et al., 2006). A *PROCOMT* score is the ratio of COMT to the total number of T-units (Hipfner-Boucher, 2011; Justice et al., 2006). Similar to narrative productivity, children's syntactic complexity also increases with age (Curenton & Justice, 2004; Justice et al., 2006; Muñoz et al., 2006).

The third facet of microstructure is *cohesion*, which refers to the connectiveness among individual sentences that is often realized by employing cohesive devices (Glosser & Deser, 1991; Hudson & Shapiro, 1991; Shapiro & Hudson, 1991). *Cohesive devices* are usually the conjunctions in a narrative sample—the connectives between clauses—and character referentials—pronouns, articles, and nominals. Conjunctions can be further divided as *coordinating conjunctions*—connective ties used to coordinate two clauses, such as "for," "and," "nor," "but," or "yet"—and *subordinating conjunctions*—connective ties used to subordinate clauses, such as "since," "though," "unless," or "if." It should be noted that cohesion relates to but is different from narrative coherence. Coherence is more

about the connectiveness among story grammar constituents while cohesion refers to the connectiveness between syntactic elements (Diehl et al., 2006; Shapiro & Hudson, 1991). Coherence contributes mainly to the flow of meaning in the whole story while cohesion contributes mainly to sentence flow of the story language. Further, Karmiloff-Smith (1985) posited that it is possible for a narrative to be coherent but without cohesive devices, and also is possible that a narrative be cohesive but without coherent structure. Although the existence of the two can be independent of each other, researchers have found that in children coherence and cohesion are closely connected (Cain, 2003; Peterson & McCabe, 1991; Shapiro & Hudson, 1991). One study revealed that a more coherent story contains greater proportion of dependent connectives, such as “but ” or “because,” a type of sophisticated cohesive connectives that indicates a dependent relation between clauses (Cain, 2003). The first explanation comes from Shapiro and Hudson (1991) who posited that good coherence allows children to focus more on cohesion in their storytelling. That is, a good conception of the macrostructure of their story demands less cognitive resources from children, and so they have more cognitive power to process the cohesive units in their story telling process. The second explanation is that cohesive devices such as inter-sentence connectives are the building material for not only cohesion but also coherence (Peterson & McCabe, 1991). In other words, coherence and cohesion share the same linguistic basis to some extent.

2.3.4.3 Evaluation

The third aspect of narrative competence is evaluation. Labov (1972) was among the first who tried to identify and define this concept, positing that fictional narratives are comprised of two essential components: basic story events and evaluation. The *basic story events* are the chain of events that are linked by temporal or causal connectives and that describe what happens in the story. *Evaluation* refers to the contextualizing clauses that identify the setting and characters of the story and comment on the story events, emphasizing why these events happen and what the point of the narrative is. Labov (1972) also pointed out that evaluative devices can exist in any place in a story but usually cluster around the climax or the high point.

Based on the frameworks of Labov et al. (1968) and Peterson and McCabe (1983), Bamberg and Damrad-Frye (1991) identified five types of evaluative devices commonly used in a story: (a) references to internal states, (b) character speech, (c) distancing devices, (d) negative qualifiers, and (f) causal connectors. *References to internal states* refer to comments on internal emotions or feelings, such as "happy," "sad," "interested," or "want." *Character speech* refers to the dialogue used by the characters. It could be direct, such as "The dog said, 'I love it!'," or indirect, "The dog told the rabbit that he loved it." *Distancing devices* refer to the linguistic devices, mainly adverbs, that express uncertainty, such as "probably," "mostly," and "kind of." *Negative qualifiers* refer to

linguistic devices that indicate negative actions or states, such as “no,” “not,” or “-ab” in words like “abnormal.” *Causal connectors* refer to connectives that indicate causal links between story events, such as “because,” “therefore,” “so,” or “thus.” Many studies showed that the number of children’s evaluative devices increases with age (Bamberg, 1987; Bamberg, 1997; Bamberg & Damrad-Frye, 1991). For example, Bamberg and Damrad-Frye (1991) reported that the number of evaluative devices used by adults is three times that used by five-year-old children and two times that by nine-year olds. Although the number of evaluative devices used generally increases with age, the developmental process for each device varied (Bamberg & Damrad-Frye, 1991; Zhu & Li, 2015). Bamberg and Damrad-Frye (1991) reported that references to internal states did not become the main evaluative device used by children until age nine. In a summary of the literature, Zhu and Li (2015) found that references to internal states increase with age, character speech decreases with age, and the use of negative qualifiers does not change much overtime.

It has also been reported that children’s use of evaluative devices is impacted by linguistic / cognitive / cultural factors and by the level of parents’ oral support (Zhu & Li, 2015). With regard to parents support, children will use evaluative devices more often and more easily if parents ask more questions or provide more feedback in a communicative

context such as storybook reading (Lever & Sénéchal, 2011; Peterson et al., 1999; Zevenbergen et al., 2003).

2.3.4.4 Measurement and Analysis

Narrative studies usually employ language sample analysis (LSA) procedures to collect and analyze narrative samples (Cain, 2003; Lever & Sénéchal, 2011; Curenton & Justice, 2004; Justice et al., 2006; Zevenbergen et al., 2003). LSA consists of four basic steps: (1) eliciting and collecting narrative samples, (2) transcribing oral narratives into text, (3) coding and analyzing the narrative text, and (4) interpreting the analysis (Leadholm & Miller, 1994). Miller (1991) posited that LSA is the only valid measure of children's expressive language. Based on their review, Justice et al. (2006) pointed out that narrative assessment, which is usually achieved by LSA procedures, is more sensitive, more valid, and less biased compared with many norm-referenced language assessments.

In step one of LSA, narrative samples collected in a study may be fictional—stories comprising of imaginary events—or autobiographical—accounts of personal experiences. Researchers have found that autobiographical narrative is easily influenced by cultural bias (Heath, 1983; Hipfner-Boucher, 2011; Pellegrini and Galda, 1993); i.e., children can show large variability in their autobiographical narratives because of differences in their familial backgrounds rather than their narrative skills. Thus,

the present study focused on children's fictional narratives instead of autobiographical ones to eliminate the confounding effects of familial and cultural factors.

Two types of tasks—story retelling and story generation—are often used to elicit narrative samples from children (Lever & Sénéchal, 2011). In *story retelling*, a child is told a story and then asked to retell it. A shortcoming with the retelling task is that narrative ability is often confounded with working memory (Feagans & Short, 1984; Hipfner-Boucher, 2011; O'Neill et al., 2004), i.e., children who have better working memory can better retell a story than their peers whose narrative competence is comparable but whose working memory is not as good. When children are required to retell without the aid of visual aids such as story pictures, the confounding influence of working memory is stronger (Cain, 2003; Cain & Oakhill, 1996; Hudson & Shapiro, 1991). In the story generation, children tell a personal story or create a story based on wordless pictures. In contrast to story retelling, narrative competence is not confounded with working memory. Therefore, in the present study, a story-generation task was used to elicit children's narrative samples.

2.3.5 Impacting Factors and Intervention Methods

In addition to the fact that children's narrative competence generally develops with age, two other factors that influence children's narrative competence—SES and gender—have also been widely researched. As with vocabulary development, SES plays an important

role in the development of narrative competence (Dickinson & Tabors, 2001; Heath, 1983; Lai, 2013; Vernon-Feagans et al., 2001). For example, Lai (2013) found that Taiwanese children from middle-SES backgrounds outperformed their peers from low-SES backgrounds on the use of temporal conjunctions, such as "然后" ("then" in Mandarin), and also on the narrative macrostructure of their Mandarin narratives. Other researchers have also found that the narrative skills of children from lower-SES backgrounds are less sophisticated than those of their peers from middle-SES backgrounds (Heath, 1983; Vernon-Feagans et al., 2001).

Another factor that can impact children's narrative development is gender. Many studies have shown that girls do better in the development of narrative competence (Buckner & Fivush, 1998; Haden, Haine, & Fivush, 1997; Peterson & Biggs, 2001; Wang, 2018). For example, Peterson and Biggs (2001) investigated the narrative samples of three-, five- and eight-year-old children and found that three-year-old girls use more evaluative devices than boys of the same age. They also found that boys tended to use more emotion labels, i.e., direct terms expressing emotions, such as "happy," "sad," or "angry." Yet, girls preferred to use more references to emotional states, that is, terms or sentences directly implying the emotional responses of the character, such as "thank goodness" or "she is screaming." Haden et al. (1997) showed that the narratives of 40-month-old girls were lengthier and more evaluative than those of boys of the same age,

and girls' narrative competences also developed faster. Buckner and Fivush (1998) analyzed the personal narrative samples of 22 seven-and-one-half-year-old children and found that girls' personal narratives were longer, more coherent, and more detailed than boys. In addition, compared with the boys of the same age, girls tended to place their narrative in a social context, referencing more people and emotions with stronger rationality. From another perspective, Frank, Baron-Cohen, and Ganzel's (2015) study suggested that the gender differences in children's narrative competence might have a neural basis. Due to these potential impacts, SES status and gender should be considered when conducting intervention research that targets narrative competence.

With regard to intervention methods, three methods have been commonly reported in the literature: storybook/video-book reading, prompted recall of past events, and parent-child pretense play (Zeng & Li, 2006). The most commonly used approach is adult-child storybook/video-book reading (Beck & Clarke-Stewart, 1998; Clarke-Stewart & Beck, 1999; Dale, Crain-Thoreson, Notari-Syverson, & Cole, 1996; Harkins et al., 1994; Huennekens & Xu, 2010; Wallach & Butler, 1994). For example, Wallach and Butler (1994) pointed out that exposure to storybooks is an important way for children to learn decontextualized language. Beck and Clarke-Stewart (1998) and Clarke-Stewart and Beck (1999) both found that video-book reading plus pedagogical mother-child discussions, using techniques such as questioning about main events, plots, causal links,

actions, internal states, moral issues, or order of events can help children better retell stories.

The second common approach is encouraging children to retell their personal stories. Peterson et al. (1999) carried out an intervention study in which parents were encouraged to prompt their children to talk more about their personal experiences by using techniques like asking complicated "WH-" questions, carefully listening, and repeating.

The third intervention approach (Baumer, Ferholt, and Lecusay, 2005) consists of three parts: story enactment, general discussion about the story, and retelling the story by drawing, printing, or role playing. The researchers found that this type of adult-child pretend play can improve the length and the coherence of children's narratives and their ability to comprehend the story. Further examination of the three intervention approaches reveals that discussion techniques such as questioning, topic extension, corrective feedback, and joint attention are usually used in intervention implementation.

To summarize, narrative competence can be simply conceived as the ability to produce and comprehend a story. Operationally, it consists of three dimensions: macrostructure, microstructure, and evaluation. Children's narrative competence generally increases with their age and across several developmental stages, and the development is affected by their SES status and gender. Three intervention

approaches—joint-book reading/video-book watching, prompted recall of the past events, and parent-child pretense play—are found to be useful in promoting narrative skills, according to the literature.

More research on this topic is required in the future. First, measures of narrative competence should be more comprehensive—taking indicators of different aspects into consideration. In addition, children's SES status and gender should be controlled in any intervention study, given that children from different SES backgrounds and of different genders benefit differently from the program on narrative competence. Finally, discussion techniques that can enhance narrative competence should be emphasized and incorporated into future intervention programs to make them more powerful in improving children's narrative language.

2.4 DR and Vocabulary

Numerous studies have shown that DR can promote the growth of young children's vocabulary (Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Lonigan, Purpura, Wilson, Walker, & Clancy-Menchetti, 2013; Strouse, O'Doherty, & Troseth, 2013; Whitehurst et al., 1988). For example, Whitehurst et al. (1988) compared the effects of parent-child DR reading with that of parent-child customary reading (CR) on the expressive and receptive vocabulary of two-year-old children from middle-SES families. The post-test results showed that DR significantly enhanced children's expressive

vocabulary. Although the differences in receptive vocabulary were not significant, the scores showed a trend favoring the DR group. In a follow-up test nine months later, the differences in receptive vocabulary between the two groups diminished, but that in expressive vocabulary remained. In another example from Lonigan and his colleagues (2013), a study of 324 preschool children found that DR can advance children's expressive vocabulary, the use of modifiers and attributes, and the composite score of these variables. Further, Strouse et al. (2013) suggested that just adding some DR elements to joint-book reading can significantly affect children's vocabulary development. Their study showed that either embedding dialogic questions into joint-book reading or making comments during the reading process can significantly improve 3-year-old children's expressive vocabulary compared to the children in a no-treatment group.

Researchers have uncovered an added value of DR in facilitating children's expressive vocabulary: the effect of DR on children's expressive vocabulary is more significant than that on their receptive vocabulary (Hargrave & Sénéchal, 2000; Lonigan & Whitehurst, 1998; Mol et al., 2008, 2009; Sim, Berthelsen, Walker, Nicholson, & Fielding-Barnsley, 2014; Whitehurst, Arnold, et al., 1994; Whitehurst, Epstein, et al., 1994). For instance, Sim and her colleagues (2014) reported that DR or DR plus print-referencing behaviours had a significant effect on children's expressive vocabulary in comparison with a no-treatment group but not on their receptive vocabulary. Also, a

meta-analysis on educator-implemented DR programs found that DR can explain 6% of the variance in children's general vocabulary, and the effect size on expressive vocabulary was much larger than that on receptive vocabulary (Mol et al., 2009). Similar results were found in another meta-analysis of home-based DR studies, in which researchers reported that the Cohen's *d* for expressive vocabulary was .59 but a much smaller number for receptive vocabulary, only .22 (Mol et al., 2008). It is possible that the interactive features of DR make it more powerful in facilitating expressive vocabulary.

The effects of DR on children's vocabulary development remain consistent across different SES backgrounds, interventionist types, instructional methods (role-modelling or video training), the number of children being read with, different intervention settings (home or school), and among children of different language levels (Arnold et al., 1994; Arnold & Whitehurst, 1994; Hargrave & Sénéchal, 2000; Lonigan & Whitehurst, 1998; Valdez-Menchaca & Whitehurst, 1992). For instance, one study (Valdez-Menchaca & Whitehurst, 1992) tested the effects of DR on the language development of 20 two-year-old children from low-income Mexican families in a daycare setting with professionals as interventionists. Their results showed that DR implemented by professionals enhanced children's receptive and expressive vocabulary and that the difference in expressive vocabulary remained in the follow-up test. Later, Arnold et al. (1994) replicated the study with children from middle- or upper-SES backgrounds, adding

another DR group for which videos were used to train the intervenors. They found that the children from whose intervenors were trained through either role-modelling or video outperformed their peers from the CR group on expressive and receptive vocabulary. Further, two studies showed that the efficacy of DR remained the same when children participated in this type of reading in a group of no more than five (Arnold et al., 1994; Lonigan & Whitehurst, 1998). Hargrave and Sénéchal (2000) extended this number to eight in their study, and they also found that DR can improve the expressive vocabulary of language-delayed children.

Although DR remains effective across different conditions, the effect size of DR was found to differ with risk status (e.g., receiving government support, low-income, or mother with a low education level), age, and intervention intensity (Mol et al., 2008; Whitehurst, Arnold et al., 1994). With regard to intervention intensity, it is claimed that conducting DR in both the home and the school setting is more effective than at school only (Whitehurst, Arnold et al., 1994). This study does not reveal whether the home setting is better than the school setting, but the root reason for the difference might be that reading in both home and school added to the intervention intensity, i.e., children in the home-plus-school setting were read to more frequently than children in the school-only group, which led to the larger increase in their performance.

The important role of intervention intensity was confirmed in another study by Cronan and his colleagues (1996), who examined the effects of a DR-plus conceptual-instruction program on Head Start children from low-income families. Participants from nine families were assigned to one of three conditions: (a) a group that received 18 interventional home visits, (b) one that received three home visits group, and (c) one that received no home visits. The results suggested that the intervention enabled greater gains on measures of parental reading behaviour, language comprehension, and conceptual development of children in the 18-home-visits group than for their peers in the other two groups. The difference in performance between the group that received three home visits and the one that received no home visits group was not significant, except for frequency of library-card use.

With regard to risk status, a meta-analysis of 16 DR studies found that the risk status of language and literacy impairments can affect intervention results significantly—children at risk benefited less from the DR program (Mol et al., 2008). Also, although there is variance within groups, overall lower-SES families tended to be less interactive during storybook-reading time (Bus et al., 1995; Heath, 1982; Ninio, 1980). With regard to children's age, Mol et al. (2008) found that younger children (2–3 years old) benefited more from the DR approach than older children (4–5 years old). They explained that this

is because the older children do not need as much parent instruction to comprehend the story and want to listen more and interact less, while young children want the opposite.

It should also be noted that not all study results support the positive effects of DR on children's vocabulary (Crain-Thoreson & Dale, 1999; Huebner, 2000; Rahn et al., 2016; Towson & Gallagher, 2014). Huebner (2000) taught community librarians to instruct parents to read with their toddlers in a dialogic way; however, no significant differences were found with regard to expressive and receptive vocabulary between the DR group and the CR group. There are two possible reasons for the lack of significant results. First, children in the study already had a very high level of language ability so that the effects of DR turned out not to be obvious. Second, educating librarians instead of reading researchers or graduate students to deliver the program might have diminished the effectiveness of the intervention. This is because researchers and graduate students usually have comprehensive professional knowledge in reading and so deal with the challenges of parent training more effectively than the librarians who only received several instructional sessions. Another study by Crain-Thoreson and Dale (1999) tested the effects of DR on 32 language-delayed children and using parents and special-education staff as interventionists. The authors found no significant effects on the expressive and receptive vocabulary of the children. They suggested that the small sample size made it hard to generate significant statistical results, and also that the overly

high-level instructional techniques used by special-education staff made it difficult for children with language problems to benefit from the program. Further, Towson and her colleagues (2014) compared the receptive and expressive vocabulary of children in a DR and a CR group but found no significant differences. The possible explanations for this result are the short duration of the intervention and the small sample size. Although some studies have shown no significant difference, overall the effects of DR on receptive and expressive vocabulary are quite well-established, given that many studies have confirmed its causal effects on the two types of vocabulary (Blom-Hoffman et al., 2007; Cronan et al., 1996; Fielding-Barnsley & Purdie, 2003; Hargrave & Sénéchal, 2000; Towson et al., 2017).

Why does DR facilitate children's vocabulary development? Several researchers (Leung, 1992; Sénéchal, 1997; Sénéchal, 2011; Sénéchal et al., 1996) have proposed four possible reasons. First, the language in storybooks is more complex than that used in daily life; therefore, children are exposed to a number of rare words in shared reading with their parents. Second, during storybook reading, parents pay close attention to their child and provide more informative and instructional interaction than in other daily activities, which is valuable to their child's language and vocabulary development. Third, storybook reading provides repeated exposure to new vocabulary, which has been found to be a very effective method of vocabulary development. Fourth, the shared-reading process provides

flexible representations of vocabulary for children; for example, children can infer word meanings from phonological (parent speech) or pictorial (illustrations in the book) presentation.

Other features of storybook reading might be beneficial to vocabulary learning. Parents can provide explicit teaching of novel vocabulary words that are less likely to be found in daily conversation (Hayes & Ahrens, 1988). Children can also acquire vocabulary implicitly from the stories they hear since the words may be embedded in supportive contexts that imply or define their meaning (NICHD, 2000; Schwanenflugel, Stahl, & McFalls, 1997). In addition, stories provide rich contextual information for children to understand vocabulary in a profound way. The plot of a story, for example, can contribute to the understanding of vocabulary items. Researchers also have shown that redundant information in the storybook is very important for vocabulary learning (Kame'enui et al., 1982).

In summary, theory and empirical research studies constantly suggest that DR provides a rich context and opportunity for children to develop their vocabulary. Many studies have demonstrated the causal relationship between DR and children's vocabulary development, especially their expressive vocabulary, across SES groups, treatment settings, interventionist types, instructional methods, the number of children being read with, and among normal developing and language-delayed children. Second, the efficacy

of a DR program can be affected by risk status, age, and intervention intensity. Finally, the effectiveness of DR lies in the rich words contained in storybooks, in interactions with parents, repeated exposure to vocabulary, various contexts of vocabulary, and rich contextual information that facilitates understanding.

2.5 DR and Narrative Competence

To a large extent, narrative language is decontextualized (Greenhalgh & Strong, 2001). That is, a narrative describes one's past experience or a story from a book that is remotely located from children's immediate environment. Many researchers believe that this type of decontextualized language can be acquired through interaction between children and competent speakers (Caspe & Melzi, 2008; Griffin et al., 2004; Nelson & Van Meter, 2007; Peterson & McCabe, 1992, 1994, 2004; Peterson et al., 1999; Snow, 1983; Uccelli et al., 1999; Vygotsky, 1978); i.e., the conversations between competent speakers and a child can help with the acquisition of the story structure and other linguistic knowledge of narratives. For young children, these speakers are usually their parents or other primary caregivers at home. Snow (1983) gave three possible reasons for this phenomenon. First, parent-child conversations involve semantic contingency—the topic can go back and forth between them continuously. Second, the conversation can act as scaffolding—fine-tuned to children's level with further guidance to move forward. Third,

the conversational process involves accountability procedures—the adult speaker showing children the best they can do in narrative language.

These three reasons, to some extent, align with the three principles of DR. Evocative questions and feedback can trigger continuous conversations between parents and children. The progressive and responsive principles make scaffolding and accountability procedures possible by encouraging children to learn more on the basis of their existing knowledge. Furthermore, stories in books are usually about a time and space remote from readers rather than the immediate environment, which can promote the acquisition of decontextualized language. Recall and distancing questions can also help children acquire this type of language by guiding them to retrieve decontextualized information from their memory and connect the story with their personal life. Thus, DR can theoretically enhance children's narrative competence.

As for empirical evidence, only a few studies have investigated the effects of DR on young children's narrative competence. For instance, Harkins et al. (1994) studied the effects of DR on 60 children's narrative skills and found that DR can significantly improve the number of clauses used and their use of evaluative devices. Further, they found that children could transfer their narrative skills into another story when they were told two different but related stories.

Also, in the study carried out by Zevenbergen et al. (2003), 71 four- to five-year-old children and their parents received DR interventions. The results revealed that a DR program can significantly improve one aspect of narrative competence—the use of evaluative devices. One limitation of this study is that only evaluation is involved and that other aspects of narrative competence are not considered. Moreover, the study used story retelling to collect children's narrative samples. As previously noted, story retelling is criticized by some researchers because it can be easily influenced by children's working memory (Feagans & Short, 1984; Hipfner-Boucher, 2011; O'Neill et al, 2004).

In another study, Reese et al. (2010) compared the effects of DR and *elaborative reminiscing* in daily speech—engaging children in talking about their past experiences—on four-year-old children's narrative skills. They found that the elaborative-reminiscing-speech group outperformed the DR group on narrative competence. A possible reason might be that elaborative reminiscing is a more explicit and direct narrative instruction than DR techniques. As suggested by Reese et al. (2010), elaborative reminiscing can be added into the DR program to make it more powerful in promoting children's language development. In this study, narrative competence was measured by answering comprehension questions and a story retelling task that was coded and measured by the number of propositions. However, comprehension questions and the number of propositions cannot capture the whole picture of children's narrative

competence, and working memory is still a possible confounding factor. In addition, only 33 children from low-income families participated in this study and were assigned to treatment and control groups, so the sample size is quite small, which might undermine the validity of the research.

Another example, from Lever and Sénéchal (2011), demonstrated only a modest effect for DR on children's narrative macrostructure and none on microstructure. This lack of significance results might stem from the low intensity of the intervention—only one hour of training for interventionists and two 20-minute reading sessions per week with children for eight weeks. Other researchers have usually suggested reading with children at least four times per week, and training of interventionists have tended to be longer or more frequent (Arnold et al., 1994; Arnold & Whitehurst, 1994; Valdez-Menchaca & Whitehurst, 1992; Whitehurst et al., 1994; Whitehurst et al., 1988). In addition, variability in SES status may have been another reason for the results of Lever and Sénéchal's study: most of the participants were from low-income families, which could have been a factor that prevented them from benefiting from the DR program. It should also be noted that the interventionists in this study were researchers not parents: the results might be different if the study were replicated with parents as interventionists.

A few DR studies did not focus specifically on children's narrative competence but did involve one or more narrative indicators as dependent variables (Dale et al., 1996;

Desmarais, Nadeau, Trudeau, Filiatrault-Veilleux, & Maxès-Fournier, 2013; Huennekens & Xu, 2010; Strouse et al., 2013). For example, Dale et al. (1996) reported that DR joint-book reading techniques significantly improved children's MLU and NDW in comparison with those of their peers in the control group. In this study, parents were told only to increase the interactivity level of their daily conversation with their child.

In another example, Huennekens and Xu (2010) reported that DR can significantly improve children's TNT and MLU-w. In addition, Desmarais et al. (2013) reported that the ability of 16 language-delayed children to answer inferential questions about the test story improved after an DR intervention. Given that inferential comprehension is an essential part of story comprehension, it can be said that DR can promote children's narrative comprehension. This result was confirmed by Strouse et al. (2013), whose video-story watching study showed that the level of story comprehension of the group that received a DR-questioning intervention was much higher than that of groups that received interventions involving pausing and commenting or questions from an actress halfway through the video, or no intervention at all.

In summary, narratives as a type of decontextualized language can be acquired by children through interaction with competent speakers, which might be the result of the semantic contingency, scaffolding processes, and accountability procedures embedded in adult-child interactions. Apart from these theoretical explanations, a few DR intervention

studies also provide empirical evidence that DR can enhance children's performance in some aspects of narrative competence. However, these studies all have their own limitations, and more empirical studies are required to support the causal link between DR and narrative competence.

2.6 DR With Mandarin-Speaking Children

2.6.1 DR Studies with Cantonese- and Mandarin-Speaking Children

A small body of Hongkong studies focuses on the effects of storybook reading using a DR approach on the receptive vocabulary of Cantonese-speaking children (Chow & McBride-Chang, 2003; Chow et al., 2008; Fung et al., 2005). Participants in these studies were either typically developing or hearing-impaired children, most of whom ranged from four to six years old. Significant differences were found between the DR and the CR group on receptive vocabulary, and the results favoured the former. One study (Chow et al., 2010) investigated the effects of DR in English on Cantonese-speaking children's Chinese and English literacy development. Three measures of Chinese literacy were involved in this study: Chinese character recognition, receptive vocabulary, and phonological awareness. English literacy was measured through English word reading, receptive vocabulary, and phonological awareness. The authors found that DR in English can promote Cantonese-speaking children's Chinese receptive vocabulary development. In the four studies, the causal relationship between the DR program and Chinese receptive

vocabulary was examined but not the effects of children's expressive vocabulary or narrative competence.

It should also be noted that children in these studies spoke Cantonese as their first language instead of Mandarin which differs substantially from Cantonese. Cantonese speakers use standard Mandarin as their written language while their oral language—Cantonese—is another variety of Chinese language. The difference between Mandarin and Cantonese might be compared with that between Spanish and Portuguese—both are Latin languages and so similar, but they differ in vocabulary and grammar. Thus, the effects of DR on Cantonese-speaking children's oral language might be different to those on Mandarin-speaking children's oral language. Moreover, the grammar, pronunciation, and vocabulary systems of oral Cantonese are very different from those of Mandarin although the two languages share the same written system.

The existence of these differences makes it important to investigate the impact of the DR program on Mandarin-speaking parents and their children. The mismatch between oral and written language might cause Cantonese-speaking children to benefit less from DR than Mandarin-speaking children. Furthermore, Mandarin-speaking children comprise a larger proportion of the Chinese population. By including Mandarin-speaking children, DR research is more meaningful because it will expand to include a larger group of children.

In addition to the above-mentioned research with Cantonese-speaking children, three Taiwanese studies have examined the effects of DR on Mandarin-speaking children's vocabulary or narrative skills. Lin and her colleagues (2005) carried out a DR intervention study with 21 five- to seven-year-old children who have autism or disabilities in intelligence, hearing, or language. They found that a home-plus-kindergarten DR intervention improved vocabulary development and parent-child reading motivation in the DR group. In this study, vocabulary was measured with a self-developed test including items for both receptive and expressive vocabulary. All target vocabulary items were selected from the books used during the intervention. It is impossible to distinguish the effects on the two types of vocabulary because only a single composite vocabulary score was obtained. Moreover, it is impossible to separate the effect of the home setting from that of the kindergarten setting. In addition, the sample size is small (11 in the DR group and 10 in the control group), which might undermine the validity of the study.

Another study (Lin & Peng, 2014) investigated the effects of a home-plus-kindergarten DR program on the receptive and expressive vocabulary of four-year-old Mandarin-speaking children from low-SES backgrounds. The children in this study were language-delayed children. The vocabulary was measured with the Receptive and Expressive Vocabulary Test [REVT, 华语儿童理解与表达词汇测验] for Chinese children, a diagnostic test developed by Huang, Jian, Zhu, and Lu (2010) to

identify children's language-delay problems. The results revealed that DR can enhance Mandarin-speaking children's expressive receptive vocabulary. In this study as well, it is impossible to separate the effects of home and kindergarten DR interventions, and the sample size is also very small (only 19 in the treatment group). Another limitation to this study is that the researchers used a diagnostic test to measure language-delayed children's vocabulary achievement, which is inappropriate and threatens their research validity.

The third study (Shiu & Tsai, 2016) examined the effects of DR on young children's language development using a qualitative approach. Five five-year-old language-delayed children participated in this study and participated in DR as a group with their teacher. The researchers analyzed the text and video materials collected during the reading process and found that DR group reading in kindergarten settings can enhance children's vocabulary, oral expression, semantics, grammar, and pragmatics. In addition, they also found that children can better describe the causal relations between story events and the problem-solving process in the story, which are signs of improvement in narrative competence. Despite these positive results, the study is descriptive in nature, and a causal link between vocabulary and DR or between narrative competence and DR cannot be confirmed.

2.6.2 Shared-Reading Studies with Chinese Children That Inform DR Research

Although only three studies have investigated the effects of a DR program on Mandarin-speaking children, quite a few shared-reading studies that did not use DR approach have been conducted with Mandarin-speaking children in mainland China. These studies, to some extent, offer guidance for the present study.

Some interventional studies found that shared reading in the school setting can exert positive effects on Mandarin-speaking children's reading and language skills. For example, Zhu (2005) found that classroom-based shared-reading practices had positive effects on kindergarteners' literacy development, characterized as character recognition, vocabulary, and story construction ability. The latter two factors are closely connected to the dependent variables of the present study.

Most of the other intervention studies (Anderson et al., 2002; Li, 2003 & 2008; Li, Wu, Zhang, Zheng, & Zhu, 2010a, 2010b; Wu, Li, Anderson, & Li, 2002) focused on the effects of classroom-based shared-reading practices on early-primary-school students' formal reading development, which was usually defined as reading comprehension and character recognition. It should be noted that the classroom-based shared-reading practices in those studies consisted not only of conventional shared-reading activities but also explicit instruction of print. These studies, which showed reading-comprehension gains, might be explained by the enhanced interaction between adults and students and

also by the explicit print instruction delivered by teachers. Given the close relationships between oral language and formal reading ability, the increase in formal reading ability demonstrated in these studies might be accompanied by an increase in oral language (Heath, 1982, 1986). Therefore, the positive results in these studies imply that reading with an adult might lead to an increase in a Mandarin-speaking child's oral-language skills.

Several other researchers have investigated the effect of home joint-reading activity on the literacy development of Mandarin-speaking children. Focusing on the home literacy environment, these studies have revealed positive relationships between the frequency of shared reading and primary-school students' general reading proficiency, word comprehension, passage reading comprehension, and long-term reading achievement (Meng, Zhou, & Kong, 2002; Shu, Li, Anderson, Ku, & Yue, 2002a; Shu, Li, Gu, Anderson, Wu, Zhang, & Xuan, 2002b; Wang, & Hu, 2005; Zhao, 2003). Moreover, an intervention study by Wu (2014) investigated the effects of interactive shared-reading practices between parents and four- to six-year-old children on children's reading motivation and narrative competence. The researcher found that a five-month interactive shared-reading intervention significantly increased children's reading motivation and narrative competence. It should be noted that the interactive shared reading in this study included only a few dialogic features—opening questions, such as, "How do you feel

when you see this picture?" and inspiring questions, such as, "Do you have the same feeling as Anna?"—and mainly focused on parents' reading style (comprehension-focused or with performance) and sharing of reflections between parent and child. The shared-reading approach used was interactive but much less dialogic and emphasized children's language less than a DR approach would have. The researcher, however, still posited that it was the interactive nature of shared reading that made it a rich context for learning vocabulary and narrative skills. If this is the case, it is possible that the DR approach, as a shared-reading method with a higher degree of interaction and more emphasis on language, can exert even more beneficial effects on children's vocabulary and narrative development. To summarize, the aforementioned correlational and intervention studies all indicate the possibility that a DR approach in the home setting can improve children's oral-language skills.

In summary, four Hong Kong studies provided evidence of causal relationships between DR and Cantonese children's receptive vocabulary in home settings. Three Taiwanese studies demonstrated the positive effects of a DR program on Mandarin-speaking children's receptive and expressive vocabulary and narrative ability in a home-plus-kindergarten setting or a kindergarten-only setting. In addition, the positive findings of several shared-reading studies in mainland China implied a possible causal link between the DR approach and Mandarin-speaking children's vocabulary

development and narrative competence. However, most shared-reading studies in mainland China have centred on the effects of shared reading on the formal reading skills of Mandarin-speaking students in early primary school.

The four Hong Kong DR studies focus on Cantonese-speaking children and their receptive vocabulary, and neither Mandarin-speaking children nor expressive vocabulary is included. The two DR Taiwan studies on Mandarin-speaking children are limited by small sample size, the lack of separation of the effects of home and kindergarten settings, and the absence of consideration of narrative competence. The single DR study that explored both vocabulary development and narrative competence was conducted in kindergarten only, was descriptive in nature, and could not provide evidence of a causal link between DR and vocabulary development or narrative competence. For these reasons, the effects of a DR program on Mandarin-speaking children' expressive vocabulary development and narrative competence in the home setting turns out to be a meaningful goal for future researchers.

2.7 The Present Study

The main purpose of the present study is to examine the causal relationships between a DR approach and Mandarin-speaking kindergarteners' expressive vocabulary and narrative competence. One secondary purpose is to explore the impact of a DR program on parental reading behaviours by observing parent-child reading sessions. The other

secondary purpose is to evaluate the extent to which the program was implemented as intended and how the DR program operates from the parents' perspective and also by analyzing observational records, reading frequency and duration checklists, parental self-reported information, and attendance checklists collected during the research process. The following questions are posed:

(1) What are the effects of a DR approach on Mandarin-speaking kindergartners' expressive vocabulary?
(2) What are the effects of a DR approach on Mandarin-speaking kindergartners' narrative competence?
(3) Are there/what are the changes in parental shared reading behaviour?
(4) To what extent is the program delivered as intended?
(5) What are parents' experiences / opinions of the DR program?

Based on the literature review, it is hypothesized that DR will exert positive effects on the expressive vocabulary and narrative competence of Mandarin-speaking kindergartners.

Chapter 3: Methodology

3.1 Participants

Eighty-one four- and five-year-old children and their parents were recruited from a kindergarten in a medium-sized city in southwest China. The city has a population of about five million. The kindergarten has a mainly middle-SES population. Participants were randomly assigned to one of two groups—the *DR gro*up, where parents received the DR intervention, or the *CR group*, where parents were encouraged to read at home with their children in their customary way. The criteria for selecting participants were as follows: (a) Mandarin was the primary language in the home, and (b) the child was either four or five years old. Due to children's illness, absence or unavailability for testing, and reluctance to be tested, the sample size was reduced to 64 during the research process.

3.2 Overall Procedures

In China, the academic year has two semesters—the fall semester from September to January and the spring semester from February to July. Young children usually start kindergarten when they are three years old and proceed through junior, middle, and senior classes over three years. At the beginning of the fall semester, parents of the children in the middle and senior classes of the target kindergarten were informed by their child's teacher via WeChat of a call for participants in a parent-child reading program. (WeChat is a Chinese mobile app featuring messaging, social communication, and payment

functions). Parents were also informed that an information session for the program was going to be held on a specific day of the next week in the largest conference room at the kindergarten.

Parents who were interested in the program participated in the information session. At the beginning of this session, basic information about the program was presented. Then, an envelope containing a letter with information about the study along with a consent form were distributed to them. The content of the information letter and the consent form was explained to parents before and as they read them, and questions raised by parents were immediately answered by the researcher. After this, parents signed the consent form if they agreed to participate, replaced it in the original envelope, and returned it b to the researcher; otherwise, they replaced the unsigned consent form in the envelope and returned it.

Some parents who were interested in the program but had no time to attend the information session requested the same information from their child's teacher, who had agreed to cooperate. If these parents had any concern or confusion about the program, the letter, or the consent form, they contacted the researcher via the contact options on the information letter, including at a cell phone number, WeChat number, and email address. Parents who agreed to participate signed the consent form, replaced it in the original

envelope, and then returned it to their child's teacher when they dropped their children off before school or picked them up afterwards.

The information letter provided parents with a general description of the research study. A brief introduction to the DR program was also included in the letter, though detailed DR procedures were not mentioned. The consent form contained the information about the program as outlined in the ethics application. Parents were informed that the program workshops would be offered in two stages, and that some of them would be assigned randomly to workshops in Stage one and the rest to Stage two, which would offer workshops several months later. They were also informed that the program was especially suitable for families for which parent-child reading was already a routine and that had access to at least several child picture books.

After the consent forms were received by the researcher, children were assigned randomly to the treatment group (DR) and the control group (CR). After randomization and two weeks before the first workshop, the parents in the DR group were telephoned to inform them of their group assignment and of the time and location of the first DR session. The parents in the CR group were also telephoned and told that they were in a group that would receive the workshops at later time. At the same time, the children in both groups were pretested on their verbal intelligence, expressive vocabulary, and narrative competence skills in the class's quiet sleeping room. (At this kindergarten, each class had

its own room for children to take a nap after lunch.) This room, which was connected with the classroom directly or through a washroom, was usually equipped with small tables and chairs in one corner. At the same time, each has two doors and multiple large windows so that the testing process was visible to parents and teachers.

During the pretest, children's parents were asked to fill out a family-information questionnaire, which was used to collect information about the family and its HLE (home literacy environment) information. The online link to the questionnaire was sent to parents of both groups via WeChat, and paper copies of the questionnaire were also provided as needed. All parents except one whose cellphone was broken at that time completed the questionnaire online. The parent who filled out the paper copy handed the questionnaire to the researcher, and their answers were input into the computer immediately.

After the pretest and within the first nine weeks of the 12-week intervention, parents in the DR group participated in three instructional workshops (at Weeks 1, 4, and 9). The workshops introduced DR, including its principles and underlying ideas, and over the three sessions gradually explained the PEER sequence, CROWD prompts, and connecting these to DR and to narrative competence. In the intervening weeks, parents, at home, worked on what they had learned during the workshops. Before each workshop, parents in the DR group received two text reminders via WeChat—the first reminder three

days before the workshop and the second one day before. Before the start of each workshop, an attendance checklist was used to record the presence of parents.

Parents in the DR group had assigned home activities on DR practices and were asked to read with their children at least four times per week. Parents in the CR group received no intervention but were also asked to read with their children at least four times per week. Parents who attended the workshops received two picture story books at the end of each session along with a reading checklist. Parents in the CR group received the same books and checklist from their children's teacher. The checklist was used to help parents track and record the frequency and daily duration of reading until they received a new checklist at the next workshop. Parents submitted a photo of their checklist directly to the researcher via WeChat. As well, parents in the DR group received additional DR guidelines and a list of suggested questions.

Beginning one week after the completion of the 12-week program, children were tested on their vocabulary and narrative competence. Four months later, a delayed posttest on children's vocabulary and narrative competence was re-administered.

To further understand the effects of the DR intervention and to monitor the intervention fidelity, four parent-child dyads—two from the DR group and two from the CR group—were selected into the present study, and their home reading sessions were videotaped on five occasions to examine parental reading behaviours over the duration of

the intervention. The five time points were just before the start of the first workshop, at Weeks 3, 8, and 12, and at the time of the delayed posttest.

After the 12-week intervention, parents from the DR group were asked to complete a follow-up questionnaire that requested information about the status of their HLE at that time and their feedback on the program. Their responses were intended to contribute to a fuller understanding of the findings obtained from the children's language assessments and provide information that could lead to further improvement of the program.

After the follow-up questionnaires were collected, eight parents from the DR group (not among those who had been observed on video in their homes) were selected through convenience sampling and interviewed to further explore their responses to the follow-up questionnaire.

After the delayed posttest, all the parents in the CR group were invited to participate in the same DR workshops. Workshops were also offered to kindergarten teachers in response to their request to learn more about this topic.

3.3 Measures

The systematic evaluation of a family-literacy program such as a parent-child reading program usually includes three aspects: assessing child outcomes, assessing parent outcomes, and evaluating the program process (Wasik, 2004). *Child outcomes* refer to

children's literacy achievement, which is usually the central focus of family-literacy programs. *Parent outcomes* are also very important and could consist of either their participation in the program or their own literacy. *Program progress* refers to the implementation or the operation of the family-literacy program. A systematic evaluation can facilitate the further improvement of a program (Pierre, Ricciuti, & Tao, 2004).

In the present study, several data-collection methods were employed to provide a comprehensive evaluation of the DR program. The measures of child outcomes, parental outcomes, and program process are summarized in three categories: (a) child measures, (b) parental measures, and (c) program fidelity and parental rating of the program.

3.3.1 Child Measures

3.3.1.1 Verbal Intelligence

The Chinese-Wechsler Young Children Scale of Intelligence, Vocabulary subtest (C- WYCSI; Gong & Dai, 1992) is a Chinese version of the Wechsler Preschool and Primary Scale of Intelligence, Vocabulary subtest (WPPSI; Wechsler,1967). In the present study, it was administered in the pretest phase to measure children's verbal intelligence, and scores were compared between groups in preliminary analyses to ensure group equivalence.

In the vocabulary subtest of C-WYCSI, the individual child is asked to select from four pictures the one that matches the word spoken by the tester. The test has different

items for children living in the city and those in the countryside. The subtest for city children, which consists of 44 items in total, was used in the present study. The Cronbach's alphas for C-WYCSI in the present study is .81.

3.3.1.2 Expressive Vocabulary

The Expressive Vocabulary Test – 2nd edition (EVT-2, Williams, 2008), is a measure of English expressive vocabulary for individuals aged from 30 months to 90 years that includes two tasks—producing labels and producing synonyms. The *labelling task* asks children to name the object in the picture. The *synonym task* asks children to produce a synonym for a word spoken by the tester. EVT-2 has two equivalent forms (A and B), and each form consists of 190 items arranged in order of increasing difficulty. Each form has eight starting points based on the examinee's age. Five consecutive correct answers are deemed as the basal. The testing continues until the child gives five consecutive incorrect answers. In the present study, children were tested individually at the pretest and delayed-posttest stages, using Form A for the pretest and Form B for the posttest. The Cronbach's α was 0.95 for EVT-2 Form A in the pretest and 0.89 for EVT-2 Form B in the posttest.

The test was translated into Chinese before its administration. Certain synonym items were modified in order to make them more applicable to Chinese language and to Chinese children. For example, one item asks children to provide the synonym of "cat." In

English, there are many synonyms like “kitten,” “kitty,” “wildcat,” etc. However, there is no synonym for “cat” (猫) in contemporary Chinese language, although ancient Chinese does have one, “狸奴,” which is overly difficult for four- and five-year-old children and so not at the same difficulty level as the original item. Thus, this item was adapted to ask, “What kind of cat is it?” Possible answers for this item are “greyish white cat (灰白猫),” “wildcat (野猫)”, “cat cub (幼猫/小奶猫),” etc., which are more similar in the difficulty level of the original item. The *Modern Chinese Dictionary* (7th ed., 2016) was consulted in the translation of this test.

3.3.1.3 Narrative Competence.

The Edmonton Narrative Norms Instrument (ENNI; Schneider et al., 2005) was adapted and used in the present study as the main instrument for eliciting and analyzing child narratives. ENNI uses a picture-aided story-production task and a set of comprehension questions about the story used in the production task to assess children’s narrative competence. These tasks tap two aspects of fictional narrative competence, story production and story comprehension. As suggested by Hayward (2003), story narrations and comprehension questions reflect different characteristics of children’s narrative competence, which makes it meaningful for researchers to use convergent methods (story production and comprehension questions) to measure narrative competence.

In the following paragraphs, the measurement of narrative production will be presented in sequence of the first three steps of language sample analysis procedure (LSA): elicitation, transcription, and coding and scoring. (The fourth step of LSA, analysis and interpretation of the narrative samples, will be described in chapter four and five of this dissertation). The measurement of story comprehension will then be presented, and the adaptations to the ENNI and the considerations taken into account when making them will also be explained.

Elicitation stories and procedures. Two stories from the ENNI were used as elicitation stories in the present study. The ENNI comprises two sets of stories (Set A and Set B) and a training story. Set A consists of three wordless picture stories ranging from simple to complex: story A1 includes five pictures, story A2 has eight pictures, and story A3 has 13 pictures. Set B also consists of three stories (B1, B2, B3) that each have the same number of pictures as the corresponding Set A stories. Set A stories and Set B stories are meant to be equivalent, and either set can be used in the test to assess young children's achievement in narrative competence. The training story consists of five wordless pictures and is used to familiarize children with the testing procedure at the beginning of the test.

In the present study, it was impossible to use a whole set of stories because of time constraints. Therefore, the practice adopted by Hayward et al. (2009) was followed: only

the longest stories, A3 and B3, were used. Story A3 tells a story in which an elephant and a giraffe are playing with a toy airplane by a pool when the airplane falls into the water. In Story B3, a rabbit releases a dog's balloon into the air accidentally. In the present study, A3 was used in the pretest and delayed-posttest stages, and B3 was used in the posttest. The longest stories were used instead of the shorter ones because they can test the highest level of the participating children's narrative achievement, this according to Hayward et al. (2009) and also observed among participants in the present study.

The narrative production of each child was assessed individually in two testing sessions. In the first session, the training story was shown to the child. The child was asked to look through the pictures associated with this story and then tell it to the examiner with the aid of the pictures. With this story, the examiner was able to use explicit prompts to help the child finish the story, such as "One day… a boy was…." In the second testing session, the child was asked to look through the pictures of the testing story (A3 or B3) and then tell the story to the examiner with the aid of the pictures. The examiner used only neutral responses, such as "Uh-huh" or "okay" in this part of the session. The second testing session was audio-recorded for further transcription and coding.

Transcription. All the narratives recorded were transcribed in Chinese in verbatim according to the Codes for the Human Analysis of Transcripts (CHAT;

MacWhinney, 2000). For convenience of analysis, the transcribed narratives were segmented into T-units or C-units using the standards of Hunt (1965).

Coding and Scoring. All three aspects of narrative competence—macrostructure, microstructure, and evaluation—were measured during the present study. The coding scheme for the ENNI narrative-production task focuses on narrative macrostructure while including a few indicators of microstructure (Schneider et al., 2005). To comprehensively evaluate children's narrative production, indicators of narrative microstructure and evaluation drawn from the literature, were applied in the present study.

In selecting these indicators, several criteria were applied in order to guarantee instrument reliability and validity. First, the selected indicators had appeared frequently in the literature and had strong theoretical and empirical bases. Second, due to the language differences between Chinese and English, the selected indicators, most of which were drawn from English-language research, can be applicable to the analysis of Chinese language. Third, the selected indicators were sensitive enough to capture individual differences between children and the changes in their narrative competence over the intervention process. To satisfy the three conditions, the indicators of narrative competence used in many studies, especially narrative studies and studies with experimental or causal-comparative designs conducted with speakers of a Chinese language, were examined (e.g., the 19 studies of narrative competence documented by

Wang, 2018). Commonly used indicators that were sensitive to capture the differences or changes were selected and included in the coding scheme of the present study.

Macrostructure. The ENNI coding and scoring scheme for story grammar, developed by Schneider et al. (2005) and adapted by Lever and Sénéchal (2011), was used to analyze the macrostructure of children's fictional narratives. In the coding scheme, macrostructure is evaluated by 12 story-grammar constituents:(a) formal beginning statement, (b) informal beginning statement, (c) characters, (d) setting, (e) initiating event, (f) internal response, (g) internal plan, (h) attempt, (i) outcome, (j) reaction of the character, (k) formal ending statement, and (l) informal ending statement. Story-grammar constituents (c) to (l) were originally used by Schneider et al. (2005) in ENNI. Lever and Sénéchal (2011) added four story-grammar elements—formal and informal beginning statements and formal and informal ending statements—because these are well documented in the research literature and are important to the macrostructure of a story. In scoring, the presence of certain elements (formal beginning, formal ending statement, initiating event, attempt, and outcome) were awarded 2 points each; other items received 1 point each. A summary of the coding scheme and scoring criteria for story grammar is presented in Table 3.1.

Table 3.1

Coding and Scoring Scheme of Story-Grammar Units

Story grammar	Descriptions
Formal beginning statement	The use of a very common or cliché opening, such as "One day," "Long, long ago," or "Once upon a time" (presence –2, absence–0)
Informal beginning statement	The use of a story opening that is not very common, such as "The elephant and the giraffe are good friends." (presence–1, absence–0)
Character	The first mention of a character in the story, such as "the elephant" or "the giraffe." Only one pronoun—"I" or "me"—is acceptable when the child puts themself into the story and uses the pronoun to refer to themself. (presence–1, absence–0)
Setting	Where the character is, what he / she / it is doing, and/or the habitual state or characteristic of the character, such as "She always wants to go shopping" or "He likes playing in the forest." (presence–1, absence–0)
Initiating event	The event that triggers the rest of the events in the story and will lead to the response of the central character, e.g., "The giraffe is flying his airplane" will stimulate the elephant's internal desire to have his own airplane. (presence-–2, absence–0)
Internal response	The reaction of the central character to the initiating event, such as "The elephant also wants to have an airplane." (presence–1, absence–0)
Internal plan	The central character's internal plan to cope with the initiating event, such as "The elephant decides to..." or "thinks he will...." (presence–1, absence–0)
Attempt	The central character's attempt to achieve the goal, such as "The big elephant tries to get the ball." (presence–2, absence–0)
Outcome	The consequence of the attempt, such as "The giraffe gets the ball." (presence–2, absence–0)
Reaction of the character	The character's feelings or thoughts about the outcome or their physical reaction, such as "The giraffe was happy again." (presence–1, absence–0)
Formal ending statement	The use of a very common story ending, such as "The end" or "And they lived happily ever after." (presence–2, absence–0)
Informal ending statement	The use of an uncommon story ending, such as "And they played together for the rest of the day." (presence–1, absence–0)

Note. Adapted from: Lever & Sénéchal, 2010; Schneider et al., 2005

Microstructure. After the indicator-selection criteria were applied, the following indicators of microstructure were included: (a) total number of words (TNW), (b) total number of different words (NDW) measured by one CLAN command Vocd, (c) one indicator of lexical density (LD), verb density (Vds), (d) another indicator of LD, adverb density (ADVds), (e) the mean length of the five longest utterances in words (MLU5), (f) the total number of T-units (TNT), (g) one indicator of cohesion, the total number of conjunctions (CONJ), and (h) another indicator of cohesion, the number of different conjunctions (DfCONJ). The eight indicators were analyzed using the Child Language Analysis (CLAN) program that had been downloaded from the Child Language Data Exchange System (CHILDES) website (MacWhinney, 2000). Unlike English, there are no spaces between words in Chinese languages. A JavaScript program, "Chinese segmenter" (Talk Bank, n.d.), was employed to segment the Mandarin sentences into words in the CHAT files. This program was available on the TalkBank website (https://talkbank.org/morgrams/), of which CHILDES is one component.

Evaluation. To measure narrative evaluation, two indicators were selected from the literature: character speech and references to internal states. As mentioned in Chapter 2, character speech (CS) refers to the dialogue between the characters, which can be direct, such as "giraffe said, 'I love you!'" or indirect, such as "Giraffe said he loved her." *References to internal states* (RIS) is defined as references to the internal states of the

story characters or the storyteller, which can be thoughts, emotions, or feelings, such as "happy," "anger," or "think." The presence of each CS item in the narrative sample, whether direct or indirect speech, was awarded one point. A tally of the CS items constitutes the total score for character speech. Similarly, participants were given one point for each RIS item in their narrative. A tally of the RIS items comprises the total score for children's references to internal states.

Story comprehension. Following the story-generation task, children were asked comprehension questions about the story they had just created. Twenty-one comprehension questions adapted by Hayward et al. (2009) from the original 29 ENNI story questions were used in the present study. Each story had two sets of questions, one set related to the story grammar and the other to problem-resolution. The set of story-grammar questions consisted of 19 questions, such as "Who is in the story?" or "What was the giraffe thinking?" The problem-resolution questions were concerned with the problem and resolution process in the story, such as "How did the problem get fixed in the story?" Children needed to integrate different information from the story to answer the question.

The story-comprehension test was also audio-recorded. The transcription techniques and the scoring criteria for the answers to these comprehension questions were consistent with those used by Hayward et al. (2009). Children's responses were

transcribed in verbatim and scored on a three-point scale. If the participant gave a fully correct response, they were awarded two points. One point was given for a partially correct answer, and zero was given for wrong answers or for a none response such as "I don't know." The score for answers to each set of questions and a composite score for both sets of questions were calculated.

Transcription Reliability. An innovative method introduced by Deuchar, Webb-Davies, and Donnelly (2018) was adapted and used to measure the transcription reliability in the present study. In their study, Deuchar et al. (2018) first randomly selected 10% of the recordings from their corpus, and then further selected one minute from the middle of each recording. Then each of these recordings was transcribed by two independent researchers. The transcripts from each researcher were then submitted to a commercial plagiarism detection website called Turnitin (www.turnitin.com) to compare the degree of similarity between them. The website calculated the similarity metrics and determined a similarity percentage, which was used as a quantitative indicator of their transcription reliability.

In the present study, 15% of recordings were selected using the stratified sampling method from pre-, post-, and delayed posttest recordings, nine from each test point and, 27 in total. The stratified sampling method was used because the narrative competence recordings were collected and transcribed by the researcher at three different time points

(pre-, post-, and delayed posttest). To counter the possibility that the differences in time might impact the researcher's transcription quality and thus ensure the sample recordings were truly representative, it was necessary to select an adequate number of recordings from each transcription wave.

After the recordings were selected, a second transcriber was trained and assigned to transcribe one part of each recording—children's story production. Different from Deuchar et al. (2018), 27 transcripts were generated in this process, one for each recording. These 27 transcripts then were compared with the 27 transcripts of story production generated by the researcher in terms of their similarity degree. Transcripts on the same recording were uploaded to a plagiarism detection website called Copyleaks (http://copyleaks.com) pair by pair. That is, two transcripts on the same recording were uploaded each time. Copyleaks gave each pair a similarity percentage, and then an average similarity percentage of 27 pairs was calculated and used as the indicator of transcription reliability, that is, 86.93%. Copyleaks was used because this website is accessible to anyone, while Turnitin requires course enrollment information for access. The transcription reliability (86.93%) in the present study compares favourably to other literature. For example, Deuchar et al. (2018) and Deuchar (n.d.) measured the transcription reliability of two corpora using Turnitin, 74% for one corpus, and 83% for another.

To measure transcription reliability, Hayward et al. (2009) used word by word agreement (WWA) and the WWA number for the story production and the question answering transcripts in their study were 96.5% and 97.2% respectively. Compare with these WWA numbers, the similarity percentage in the present study is somewhat lower. This is because the anti-plagiarism website used in the present study is very sensitive and strict with respect to slight disparities. For example, children sometimes repeated words when telling a story, such as, "he, he, he said." Researcher A transcribed the repetition as what it was while researcher B omitted one "he" and transcribed as "he, he said." When this difference was detected by Copyleaks, the whole utterance would be deemed as not identical or only meaning related. Deuchar et al. (2018) also reported that many differences detected by Turnitin were caused by slight disparities in the transcription instead of substantial differences. Thus, the similarity percentage generated by the plagiarism software is expected to be lower than WWA; however, this does not indicate lower transcription reliability.

Coding and Scoring Reliability. A second coder was assigned to code the 27 transcripts of story production generated by the researcher based on the coding scheme of story grammar (SG), references to internal states (RIS), character speech (CS), story grammar questions (SGQ), and problem and resolution questions (PRQ). This coder calculated a score for each indicator of each transcript. These scores were then compared

with those found by the researcher using a statistic called intra-class correlation coefficient (ICC). ICC can deal with the inter-rater agreement issues involving continuous variables (Ranganathan, Pramesh, & Aggarwal, 2017) while Cohen's Kappa is used for categorical variables only (Cyr & Francis, 1992). Transforming the scores of those indicators to binary variables could result in information loss; thus, ICC, instead of Cohen's Kappa, was utilized in the present study. According to Cicchetti (1994), an excellent ICC ranges from 0.75 to 1.00, good from 0.60 to 0.74, and fair from 0.40 to 0.59. Three ICCs in the present study reached excellent level—for RIS: ICC = 0.762 ($p < 0.001$), for CS: ICC = 0.932 ($p < 0.001$), and for PRQ: ICC = 0.769 ($p < 0.001$). One remaining ICC was good—for SG: ICC = 0.685 ($p < 0.001$), and for the final indicator, the ICC was fair—SGQ: ICC = 0.527 ($p < 0.001$). The indicators of narrative microstructure, in the present study, were automatically coded and scored by CLAN; thus, their coding and scoring reliability was not calculated.

3.3.2 Parental Measures

3.3.2.1 Family-Information Questionnaire

At the pretest stage, parents were asked to fill in a Chinese-language family-information questionnaire (see Appendix A) to answer questions about their family's demographic information and HLE. The demographic questions were adapted from demographic and background information questionnaires used in a series of studies

(Chow & McBride-Chang, 2003; Chow et al., 2008; Crain-Thoreson & Dale, 1999; Cronan et al., 1996; Doyle, 2009; Fielding-Barnsley & Purdie, 2003; Fung et al., 2005; Huebner, 2000; Valdez-Menchaca & Whitehurst, 1992; Whitehurst et al., 1988). Nine HLE questions were adapted from an HLE scale developed by Payne, Whitehurst, and Angell (1994) and based on the Stony Brook Family Reading Survey (Whitehurst, 1992). One of the nine HLE questions was about library visiting; however, in cities of China, many families use bookstores or book buildings instead of libraries as places to read or to obtain books. Thus, a new HLE question on bookstore visiting was added, which meant that there were 10 HLE questions. Several other HLE questions were also added into the questionnaire, and these questions were either on literacy activities other than reading or on the emotional and motivational literacy environment. Besides, the questionnaire also included questions on the dyad's DR reading behaviours and parental DR reading beliefs, and these questions, combined with certain questions out of the HLE question set, were also included in the parental follow-up questionnaire for the DR group administered at the posttest stage. The differences in parent answers to these two questionnaires were used to assess the influence of the program on child interest, family-literacy outings, child DR reading behaviours, parental interest in shared reading, parental confidence in shared reading, parental DR reading behaviours, and parental beliefs towards parent-child reading.

The questions in the family-information questionnaire administered at the pretest stage were organized into three sections: (1) questions about their child, (2) questions about their family, and (3) HLE questions and extended questions. The first section involved questions about the name, gender, date of birth, school, and class of their child. The second section asked questions about home-contact information; annual income; parents' name, education, occupation, parental expectation of the program, and other literacy programs they attend. The question format of first two sections were either filling in the blank questions or questions of multiple choices. The third section contained 10 HLE questions on home literacy materials, the age when their parent-child reading began, frequency and duration of parent-child reading, parents' own reading interests and duration, child's reading interests, and library-, bookstore-, and book building–visiting behaviour. The extended questions in the third section involved questions on parent-child reading behaviours, parental beliefs on parent-child reading, and parental efficacy in supporting their child to read. In this section, parents recorded their responses to the questions on a 5- or 7- or 10-point Likert-type scale.

3.3.2.2 Follow-Up Questionnaire

At the end of the intervention, parents in the DR group were asked to complete a follow-up questionnaire (see Appendix B). This questionnaire addressed two aspects: follow-up questions on HLE and DR reading beliefs / behaviours as well as parents'

perceptions of program participation and its perceived impact on them as participants. Questions 1–10, which addressed the first aspect, were taken from the pretest family-information questionnaire. The pretest and posttest answers to these questions were compared to evaluate the changes, if any, in HLE and DR reading beliefs / behaviours among DR group participants. Question 11–18 queried parents on their perceptions of program impact and operation and were devised after consulting interview protocols and questionnaires used in prior studies (Doyle, 2009; Payne et al., 1994; Seefeldt, 2004; Thomas & Fisher, 1996). The questions in this section asked parents about their perceptions on the impact of the DR program on their child and themselves, including changes in their beliefs about parent-child reading, the usefulness of different elements of the program, their suggestions for further improvement, and any difficulties they might have experienced in carrying out the recommended program techniques at home. Many of the answers were recorded on a 5- or 7-point Likert-type scale; others, as appropriate, were open questions.

3.3.2.3 Home Video Observation

The changes of parental reading behaviours over time were evaluated using home video observations. Two dyads from each group were selected using convenience sampling. The reading sessions of these four dyads were videotaped at five points in time: during the pretest stage; at weeks 3, 8, and 12, and during the delayed-posttest stage. In the

videotaped sessions, parents were asked to read a picture book with their child. The first 270 seconds of each reading session were transcribed, coded, and analyzed to determine the changes in parental DR behaviours. The consistency of the DR parents' reading behaviours with instruction given during workshops and the discrepancy in parental reading behaviours between the two groups were analyzed to evaluate program fidelity and program efficacy.

An observation coding scheme from Blom-Hoffman et al. (2007) was adapted and used in the coding and analysis of parental DR behaviours. All the DR behaviour codes used by these authors had been adapted from the DR techniques documented in Zevenbergen and Whitehurst (2003): (a) *page prompt*, a parental prompt that can initiate discussion on the objects on the page, such as open-ended or closed-ended questions about the book content, (b) *attending statement*, a parental prompt used to direct children's attention to the book, such as "Look at...," (c) *evaluating prompt*, feedback given by a parent on a child's response, such as praise, correction, or repetition of the child's words, (d) *expanding prompt*, parental expansion of a child's words, (e) *repeat prompt*, a parental request for the child to repeat what they have just said, (f) *completion prompt*, a parental requests that the child add words to their uncompleted sentences, (g) r*ecall prompt*, a parental prompt that requests the child to recall part of or all of the story,

and (h) *distancing prompt*, a request that the child connect book contents with their real life.

3.3.2.4 Semi-Structured Interviews with Parents

Eight parents from the DR group were selected by using a convenience sampling method to participate in a 30- to 45-minute semi-structured interview regarding their evaluation of the program. The interview questions were the same as those in the follow-up questionnaire; however, the interviewees were invited to elaborate on their responses. All the parental interviews were audio-recorded, and these were later transcribed into text.

The open-coding method (Given, 2008; Patton, 2002) was used to analyze the data. In the first step, the researcher segmented transcripts into segments. Back-and-forth dialogue related to the same theme was treated as a single segment. Then, the researcher read the transcripts word by word and line by line, by which process concepts or keywords related to the research purpose were extracted. In the third step, the concepts were categorized into groups according to their meaning. Next, categories of concepts were further interpreted and analyzed. The meaning of each concept was further explained and elaborated with examples, and implications were extracted.

3.3.2.5 Program Fidelity

Fidelity refers to the extent to which the program is conducted as intended (Powell, 2012). To support program fidelity, every week, parents in both groups received a WeChat

message that reminded them to read with their child at least four times per week. Parents in the DR group also received a calendar checklist to keep track of the frequency of parent-child reading at home and the duration of each reading session. WeChat reminders, frequency checklists, and records of session duration were used to remind parents to make reading with their child a routine. Moreover, before each workshop, parents in the DR group received two reminders of the workshop time and location, one three days before and another one day before the workshop. When they arrived for the workshop, they were asked to check off their name on the attendance list before its start. The workshop reminders and attendance list were used to promote participant attendance.

The data gathered with calendar checklists, records of the duration of reading sessions, and workshop attendance lists were used to evaluate and monitor program fidelity. As well, observations of parent-child home reading were also used as an evaluation measure of fidelity.

3.4 Intervention Program

Three DR workshops were provided for parents in the DR group. The DR techniques taught during the workshops were adapted from Zevenbergen and Whitehurst (2003). Several complementary tips to increase interest in reading were drawn from Ezell and Justice (2005). These tips were explained at the end of each instructional session and were listed on the handout distributed to the DR parents. The workshops occurred at

Weeks 1, 4, and 9, and each lasted about one hour. The workshops usually included the following elements: (1) explanation of the content, (2) viewing videos that modelled the techniques, (3) in-class activities and discussion, and (4) answering parents' questions.

The modelling videos used during the workshops were developed by the researcher based on the DR approach instructions. One mother and her five-year-old child who did not participate in the program and lived in another city were invited to model DR strategies in front of a video camera. First, the mother was taught the DR techniques. Next, the mother was asked to read with the child using DR techniques, and the process was videotaped. The video recordings were then used as instructional materials in the DR workshops. In keeping with ethical principles, the mother was provided with information about the study and both mother and child provided consent before videotaping proceeded. The content of the three workshops is outlined in Table 3.2.

Table 3.2

Main Content of Instructional Workshops for Parents

Workshop Session	Topics
Session 1	A brief introduction to the DR workshops and the objective of Session 1 What is shared reading, and what is its importance?
Introduction and PEER sequence	What is DR? Three DR principles (core topic) DR sequence: PEER with video modelling (core topic) Tips for enjoyment: – Remember the importance of enjoying parent-child reading – Give praise and encouragement not criticism – Judge and weigh the frequency and timing of asking questions in shared reading
Session 2	Review of Session 1 Objectives of Session 2
CROWD prompts and asking sophisticated questions	Vocabulary (definition and importance) and effects of DR on vocabulary Prompts: CROWD with video modelling (core topic) Discussion on the efficacy of different prompts (e.g., WH- question vs. yes/no question); the sophistication of the questions should increase with the age of the child Tips for enjoyment: – Allow the child to end reading when he or she is bored or tired – Be tolerant when the child's response is irrelevant to your question – Be a good listener
Session 3	Review Objectives of Session 3
Connecting DR with children's narrative competence	Narrative competence (definition and importance) and the effect of DR on narrative competence Using recall questions to urge the child to retell stories from pictures (with video modelling; core topic) Using distancing questions to ask the child to create their personal story (with video modelling; core topic) Tips for enjoyment: – Follow the child's interests by talking more about their interesting topic – Promote the child's control of the reading process: let the child hold the book, turn the pages, and direct the pace of reading

3.5 Materials

After each workshop, the participating parents in the DR group received two picture storybooks with a handout and a list of suggested questions for each book. The handout summarized the main ideas of the workshop. The lists of suggested questions were provided to help parents practice DR techniques during their own reading sessions. They were advised that they could select questions from the list or generate their own questions. They were advised to select questions that were appropriate to their child's level of language and familiarity with the book. Parents in the CR group received the same books at the same time via their child's teacher but without guidelines or the list of questions.

Parents in both groups were encouraged to read each book repeatedly, about one book a week. When they finished reading the assigned book, they could select their own book for reading. Given that one focus of the present study was fictional narrative, parents were encouraged to select picture storybooks to read instead of expository or other kinds of books.

Two teachers from the participating kindergarten who were familiar with storybooks, the participants, and the kindergarten library were involved in the book selection process. The picture books distributed to parents were selected according to the following criteria: (a) the books were commercially available, (b) they were storybooks

instead of expository or activity books, (c) they were suitable for four- and five-year old children, (d) the books were not common among the participating families as reported by the two teachers, and (e) they not included in the kindergarten library. Many well-known picture books, such as *Fredrick* or *Something from Nothing,* were excluded from the list because their fame made them prevalent in the city. Financial considerations were also taken into account; i.e., if two books were both of high quality, the cheaper one was selected to reduce researcher expense. The list of the selected books is presented in Appendix C.

Chapter 4: Results

This chapter is comprised of three main sections: preliminary analysis (Section 4.1), measures of program effects on children (Section 4.2), and measures of program effects on parents (Section 4.3). For the preliminary analysis (Section 4.1), the results of a series of descriptive and correlation analyses of the pretest data are presented, along with those of a series of *t*- tests that show the extent of differences between the experimental (DR) group and the control (CR) group with regard to control variables. To deepen understanding of the data, the results of generalized-linear-model (GLM) analysis of the relationship between the control variables and pretest data are also presented. The results for program effects on children (Section 4.2) stem from (a) GLM analysis of the immediate and delayed effects of the DR program on child outcomes and (b) HLM analysis of the growth curve of child outcomes over time.

The section on measures of program effects on parents (Section 4.3) covers the results of analyses of parental DR behaviours, program fidelity, and parents' ratings of the program. Parental DR behaviours were analyzed based on the data collected by videotaping parent-child DR sessions in the home. As well, the results of two other aspects of program evaluation, program fidelity and implementation, are presented. Program fidelity results were derived from the analysis of the data collected from the abovementioned videotaped DR sessions, checklists that recorded of the frequency and

duration of home DR sessions, and DR workshop attendance lists. Program implementation results are based on analysis of the data collected through the questionnaire completed by parents of children in the DR group at the posttest stage and interviews with selected parents in that group.

4.1 Preliminary Analyses

4.1.1 Descriptive Analysis and Group Differences

To ensure that sampling was unbiased, several parametric and non-parametric tests were performed to examine the differences between the DR and CR groups on measures of children's demographic information and HLE. The descriptive statistics for these variables are presented in Table 4.1. Several Chi-square tests were conducted to examine differences in gender ratio, family income, maternal education, and paternal education between the two groups. A series of independent-sample *t*-tests were carried out to compare the two groups with regard to children's age (in months), HLE resources, reading activities, other literacy activities, and the emotional and motivational atmosphere in the home. No significant differences were found in any of demographic and HLE measures; all p values are > 0.05.

Table 4.1

Descriptive Statistics: Demographic and HLE Factors

Item	Group	*N*	*M*	*SD*	Male	Female
Age	DR	32	58.75	8.51		
	CR	32	60.38	6.80		
	Total	64				
Income	DR	32	3.31	1.09		
	CR	32	3.19	1.09		
	Total	64				
Medu	DR	32	4.31	1.31		
	CR	32	4.47	.62		
	Total	64				
Pedu	DR	32	4.28	1.35		
	CR	32	4.28	1.28		
	Total	64				
Gender	DR	32			16	16
	CR	32			21	11
	Total	64			37	27
HLEResources	DR	32	0.67	0.31		
	CR	32	0.67	0.32		
	Total	64				
ReadAct	DR	32	0.49	0.24		
	CR	32	0.44	0.17		
	Total	64				
OthLiterAct	DR	32	0.52	0.25		
	CR	32	0.53	0.20		
	Total	64				
EmoMoti	DR	32	0.49	0.17		
	CR	32	0.57	0.16		
	Total	64				

Note. Medu = maternal education, Pedu = paternal education, ReadAct = reading activities, OthLiterAct = other literacy activities, EmoMoti = home emotional and motivational atmosphere, DR = dialogic reading group, CR = customary reading group

The DR training provided to parents in the DR group was not intended to change the frequency or duration of parent-child reading; rather, it was intended as a measure of implementation fidelity and a control variable for group equivalency. As expected, the data collected from weekly checklists show that reading frequency and duration did not increase over time. Parents were required to submit their weekly reading checklists at three time points—Week 4 (checklists for Phase 1 – Weeks 1–3), Week 9 (checklists for Phase 2 – Weeks 4–8), and Week 12 (checklists for Phase 3 – Weeks 9–11). The checklist for the frequency and duration of parent-child reading for Week 12 was not collected or included in the analyses because the third checklist was submitted that week. A series of independent-sample *t*-tests was conducted to determine the differences between the DR and CR groups with regard to average weekly frequency and duration of parent-child reading, and no significant differences were found,; all p values are > 0.05 (see Table 4.2).

Table 4.2

Descriptive Statistics and t*-Test Results: Reading Checklists*

Phase	Group	*n*	*M*	*SD*	*t*	*df*	*p*
Reading frequency							
1	DR	27	4.65	1.30	-0.52	53	0.61
	CR	28	4.85	1.44			
2	DR	26	5.11	1.24	-0.30	49	0.77
	CR	25	5.21	1.19			
3	DR	16	5.48	0.89	0.58	41	0.57
	CR	28	5.29	1.34			
Reading duration							
1	DR	25	104.20	60.80	0.37	43	0.71
	CR	27	98.72	43.01			
2	DR	26	118.33	55.14	0.33	46	0.74
	CR	22	113.22	51.68			
3	DR	16	136.33	80.04	1.46	42	0.15
	CR	28	107.13	52.84			

Note. DR = dialogic reading group, CR = customary reading group

The descriptive statistics on the language measures conducted at the pretest, posttest, and delayed posttest stages are presented in Table 4.3. A series of independent-sample *t*-tests was conducted to compare scores on C-WYCSI, EVT pretest scores, and narrative competence pretest measures between two groups to test whether children's verbal intelligence and the baselines of dependent variables differ. There was no statistically significant difference on children's scores in these measures, all *p* values are > 0.05.

Table 4.3

Descriptive Statistics: Language Assessments at the Pretest, Posttest, and Delayed-Posttest Stages

Task	Group	Pretest (N = 32)	Posttest (N= 32)	Delayed posttest (N = 32)
		M(*SD*)	*M*(*SD*)	*M*(*SD*)
C-WYCSI, vocab	DR	30.47 (6.21)		
	CR	29.97 (5.55)		
EVT	DR	63.31 (13.12)	66.72 (7.99)	77.56 (13.17)
	CR	68.94 (14.61)	67.09 (8.84)	77.00 (12.11)
SG	DR	23.34 (3.37)	22.13 (5.06)	26.50 (3.95)
	CR	22.25 (3.90)	20.84 (5.69)	25.75 (2.78)
TNW	DR	130.78 (51.59)	147.53 (41.72)	140.63 (51.06)
	CR	114.72 (40.90)	143.06 (51.33)	121.16 (39.49)
Vocd	DR	27.53 (7.23)	27.51 (8.70)	29.96 (8.31)
	CR	27.52 (8.84)	27.52 (12.30)	29.14 (7.72)
Vds	DR	0.19 (0.03)	0.18 (0.03)	0.19 (0.04)
	CR	0.19 (0.04)	0.19 (0.03)	0.19 (0.04)
ADVds	DR	0.11 (0.04)	0.09 (0.03)	0.12 (0.03)
	CR	0.11 (0.03)	0.09 (0.04)	0.10 (0.03)
TNT	DR	17.53 (4.70)	19.09 (4.28)	20.03 (6.06)
	CR	16.03 (3.34)	18.72 (5.18)	16.94 (4.04)
MLU5	DR	10.61 (2.34)	11.34 (2.11)	10.61 (2.06)
	CR	10.27 (2.67)	11.19 (2.42)	10.28 (2.16)
CONJ	DR	9.19 (7.13)	9.72 (5.06)	9.25 (5.76)
	CR	7.59 (4.91)	7.25 (4.78)	7.22 (4.65)
DfCONJ	DR	1.94 (1.01)	2.16 (0.92)	2.06 (1.01)
	CR	1.81 (0.97)	2.06 (1.24)	1.81 (0.78)
CS	DR	2.69 (4.07)	2.56 (2.02)	1.72 (1.63)
	CR	2.00 (2.38)	2.63 (2.95)	2.03 (1.94)
RIS	DR	4.00 (2.29)	4.28 (2.26)	5.47 (2.89)
	CR	3.44 (1.78)	3.25 (2.17)	4.19 (1.58)
SGQ	DR	24.16 (5.98)	26.75 (7.03)	29.63 (3.99)
	CR	25.50 (4.06)	27.13 (7.00)	30.13 (4.46)
PRQ	DR	2.91 (1.25)	3.09 (1.06)	3.19 (1.12)
	CR	2.75 (1.19)	2.38 (1.34)	2.84 (1.14)

Note. C-WYCSI, vocab = Chinese-Wechsler Young Children Scale of Intelligence, Vocabulary subtest, EVT = Expressive Vocabulary Test score, SG = story grammar, TNW = total number of words, Vocd = proportion of novel words, Vds = verb density, ADVds = adverb density, TNT = total number of T-units, MLU5 = mean length of five longest utterances, CONJ = number of conjunctions, DfCONJ = number of different conjunctions, CS = character speech, RIS = references to internal states, SGQ = story-grammar questions, PRQ = problem-resolution questions, DR = dialogic reading group, CR = customary reading group

4.1.2 Correlation Analysis

The correlation coefficients (Pearson's *r*) were computed for the 14 dependent variables from the pretest, and the results are presented in Table 4.3. The results show that 37 of 91 correlations were statistically significant. Most of the significant correlations were weak or moderate in strength. One indicator of narrative microstructure (PreTNW) was strongly and positively related to another two indicators of narrative microstructure (PreTNT and PreMLU5) and one indicator of evaluation (PreCS). A strong and positively significant relationship was also found between two indicators of narrative microstructure (PreTNT and PreCS). These strong significant correlations suggest two obvious facts: children who told longer stories produced more T-units, longer utterances, and more direct and indirect speech in their story telling; also, children who produced more T-units used more direct or indirect speech (back-and-forth dialogue- with parents) in their stories.

Except for some indicators of narrative microstructure and narrative evaluation, most of the independent variables were positively correlated with each other at the pretest stage, though not all of the correlations were significant; i.e., if a child scored high on one task or indicator, they tended to have high scores on most other tasks or indicators. For example, children's pretest EVT scores was positively and significantly related to their indicator of narrative macrostructure (PreSG), two indicators of narrative microstructure (PreVocd and PreMLU5), and two indicators of narrative comprehension (PreSGQ and

PrePRQ). The indicator of narrative macrostructure (PreSG) was positively and significantly correlated with four indicators of narrative macrostructure (PreTNW, PreVocd, PreTNT, and PreMLU5), both two indicators of narrative evaluation (PreRIS and PreCS), and one indicator of narrative comprehension (PreSGQ). One indicator of narrative microstructure (PreTNW) was positively and significantly correlated to almost all the indicators of narrative microstructure and narrative evaluation except for the two indicators of lexical density (Vds and ADVds). Two indicators of narrative evaluation (PreRIS and PreCS) were moderately positively correlated. The correlation between one indicator of narrative evaluation (PreRIS) and indicators of narrative comprehension (PreSGQ and PrePRQ) was also positive and significant, indicating that children who used more references to internal states comprehended the story better.

The negative correlations between some specific indicators of narrative competence and other tasks, especially the moderately negatively significant correlation between two indicators of narrative microstructure (PreVds and PreDfCONJ), suggest that the relationships among them are not unidirectional: a child who performed well on one indicator might do poorly in another.

Table 4.4
Correlation Matrix of Pretest Language Measures

	PreEVT	PreSG	PreTNW	PreVocd	PreVds	PreADVds	PreCONJ	PreDfCONJ	PreTNT	PreMLU5	PreRIS	PreCS	PreSGQ	PrePRQ
PreEVT	—													
PreSG	0.364**	—												
PreTNW	0.112	0.546***	—											
PreVocd	0.296*	0.413***	0.289*	—										
PreVds	0.051	0.003	0.051	0.054	—									
PreADVds	0.107	0.068	0.097	0.025	-0.020	—								
PreCONJ	-0.182	0.032	0.522***	-0.229	0.004	0.084	—							
PreDfCONJ	0.153	0.205	0.343**	0.124	-0.358**	-0.005	0.332**	—						
PreTNT	-0.139	0.313*	0.800***	0.256*	0.117	-0.052	0.469***	0.275*	—					
PreMLU5	0.409***	0.672***	0.756***	0.437***	-0.053	0.146	0.200	0.261*	0.370**	—				
PreRIS	0.074	0.363**	0.538***	0.171	0.091	-0.077	0.148	0.203	0.572***	0.325**	—			
PreCS	-0.133	0.271*	0.724***	0.340**	0.109	-0.065	0.418***	0.193	0.698***	0.349**	0.385**	—		
PreSGQ	0.355**	0.277*	0.066	0.124	0.042	-0.176	-0.144	-0.143	0.086	0.189	0.312*	-0.094	—	
PrePRQ	0.253*	0.096	0.125	0.146	-0.152	0.088	0.026	0.194	0.126	0.140	0.311*	-0.048	0.411***	—

Note. * p <.05, ** p <.01, *** p <.001, Pre- = pretest, EVT = Expressive Vocabulary Test score, SG = story grammar, TNW = total number of words, Vocd = proportion of novel words, Vds = verb density, ADVds = adverb density, TNT = total number of T-units, MLU5 = mean length of five longest utterances, CONJ = number of conjunctions, DfCONJ = number of different conjunctions, CS = character speech, RIS = references to internal states, SGQ = story-grammar questions, PRQ = problem-resolution questions

4.1.3 The Relationship Between Control Variables and Child Outcomes at the Pretest Stage

Two sets of GLM analyses were conducted on each child outcome at the pretest stage, one to examine the relationship between demographic factors and the outcome and the other to examine the relationship between HLE factors and the outcome. For this purpose, experimental and control groups were combined. In other words, all child participants were treated as one group. GLM analysis was employed because the model is robust in nature and able to account for both categorical and continuous variables.

The set of demographic variables includes gender, age, mother's education, father's education, and family annual income. The set of HLE variables includes home literacy resources, reading activities, other literacy activities, and emotional and motivational atmosphere. The pretest variables include expressive vocabulary (EVT) and narrative competence. Narrative competence was measured through two aspects—narrative production and narrative comprehension. Three sub-facets of narrative production were measured: (a) macrostructure, measured by story grammar (SG); (b) microstructure, measured by total number of words (TNW), number of different words (Vocd), verb density (Vds), adverb density (ADVds), total number of conjunctions (CONJ), number of different conjunctions (DfCONJ), total number of T-units (TNT), and mean length of the five longest utterances (MLU5); and (c) evaluation, measured by

references to internal states (RIS) and character speech (CS). Narrative comprehension was measured through story-grammar questions (SGQ) and problem-resolution questions (PRQ). The results of the two-wave GLM analyses were summarized in Table D1-D14 (see Appendix D).

4.1.3.1 Expressive Vocabulary

Applying GLM analysis to the set of HLE variables revealed that no HLE factor significantly predicted children's expressive vocabulary. Analysis of the set of demographic variables showed that age was a significant predictor of children's expressive vocabulary, $\beta = 1.053$, $z = 5.303$, $p < 0.001$.

4.1.3.2 Story Grammar

Similarly, GLM analysis showed that no HLE factor significantly predicted children's story grammar. For the set of demographic variables, age was again found to be highly predictive of children's story grammar, $\beta = 0.225$, $z = 4.158$, $p < 0.001$.

4.1.3.3 Narrative Microstructure

The different indicators of narrative microstructure were analyzed individually. Gender was found to predict children's total number of words: boys produced fewer words in their stories than girls, $\beta = -0.234$, $z = -2.382$, $p < 0.021$. Similarly, boys obtained lower scores on MLU5, $\beta = -1.4505$, $z = -2.339$, $p < 0.024$; however, they produced more T-units than girls, $\beta = 0.008$, $z = 2.237$, $p < 0.030$. In addition, maternal education

negatively predicted children's verb density, β = -0.015, z = -2.727, $p < 0.009$. Age was found to predictive of children's MLU5, β = 0.116, z = 3.082, $p < 0.004$.

The results showed a trend towards significance in the predictive power of two factors: Family annual income was found to negatively predict the number of different conjunctions with a nearly significant *p*-value, 0.058. Also, the emotional and motivational atmosphere of the home negatively predicted the measure of the total number of different words with a nearly significant *p*-value, 0.059.

4.1.3.4 Narrative Evaluation

Neither HLE nor demographic factors were found to be predictive of the two indicators of narrative evaluation—references to internal states and character speech.

4.1.3.5 Narrative Comprehension

Applying GLM analysis to demographic data set showed that age significantly predicted one indicator of narrative comprehension—story-grammar questions, β = 0.270, z = 3.342, $p < 0.001$. No significant predictors were found for the other indicator, problem-resolution questions.

In summary, for the sample in the present study, no significant group differences were found at the pretest stage. When all the participants were treated as a single group, age was found to be a significant predictor of children's expressive vocabulary, story grammar, MLU5, and story-grammar questions. Gender was found predictive of

children's total number of words, total number of T-units, and MLU5. In addition, children's verb density was negatively predicted by mother's education.

4.2 Measures of Program Effects on Children

4.2.1 Immediate and Delayed Program Effects

Two series of GLM analyses were conducted to test the immediate and delayed effects of the DR program. For these analyses, group (treatment) assignment was the independent variable, and children's expressive vocabulary and narrative competence were the dependent variables. Four covariates were inserted into the analyses based on the literature review and the results of the univariate analyses presented in Section 4.1: age, gender, maternal education, and the pretest score of the dependent variable at the pretest stage. The results of these two waves of GLM analyses on expressive vocabulary and each indicator of narrative competence are summarized in Tables 4.5–4.18.

4.2.1.1 Expressive Vocabulary

At the posttest stage, results of GLM analyses revealed no significant differences in expressive vocabulary between groups; however, in the delayed posttest, the experimental group significantly outperformed the control group after controlling for children's baseline data: EVT achievement, age, gender, and maternal education, $\beta = 4.245$, $z = 2.078$, $p < 0.043$.

4.2.1.2 Story Grammar

Children's story grammar was the dependent variable. The results of GLM analyses results showed no difference between groups in children's story grammar at the posttest and delayed-posttest stages

4.2.1.3 Narrative Microstructure

Each indicator of narrative microstructure was entered into the model individually as a dependent variable. At the posttest stage, no significant differences were found between groups on any indicator of narrative microstructure. In the delayed posttest, however, the experimental group significantly outperformed the control group on adverb density, $\beta = 0.01964$, $z = 2.374$, $p < 0.022$.

4.2.1.4 Narrative Evaluation

Two indicators of narrative evaluation, references to internal states and character speech, were dependent variables. GLM analysis revealed no group differences at the posttest and delayed-posttest stages.

4.2.1.5 Narrative Comprehension

The score for story-grammar questions and problem-resolution questions were the dependent variables. At the posttest stage, children in the experimental group scored significantly lower than those in the control group on problem-resolution questions, $\beta = -0.2530$, $z = -2.522$, $p < 0.014$. As for performance on story-grammar questions, there

was no difference in children's scores between the two groups. At the delayed-posttest stage, while the difference between the two groups on performance on problem-resolution questions diminished, group differences were present for both indicators.

In summary, the posttest stage showed no significant differences between the two groups on almost all measures except for performance on problem-resolution questions, which surprisingly favoured the control group. At the delayed-posttest stage, the experimental group outperformed the control group on expressive vocabulary and adverb density.

Table 4.5

Effects of the DR Program on Expressive Vocabulary at the Posttest and Delayed Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 415.691			$R^2 = 0.548$
Group[a]	2.140	1.5148	1.412	0.163	1.995
Age	0.395	0.1183	3.340	0.001	11.158
Gender[b]	1.233	1.5129	0.815	0.419	0.664
Medu	0.758	0.7320	1.035	0.305	1.071
PreEVT	0.278	0.0657	4.229	< 0.001	17.883
Delayed Posttest Model Information		AIC = 453.953			$R^2 = 0.635$
Group[a]	4.245	2.0426	2.078	0.042	4.319
Age	0.398	0.1595	2.498	0.015	6.242
Gender[b]	-0.710	2.0400	-0.348	0.729	0.121
Medu	-0.579	0.9870	-0.587	0.559	0.345
PreEVT	0.575	0.0885	6.499	< 0.001	42.241

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, PreEVT = Results of expressive vocabulary test at pretest stage

Table 4.6

Effects of the DR Program on Story Grammar at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 379.896			R^2 = 0.375
Group[a]	0.676	1.1538	0.586	0.560	0.3436
Age	0.148	0.0840	1.760	0.084	3.0987
Gender[b]	-0.404	1.1701	-0.345	0.731	0.1193
Medu	-0.171	0.5551	-0.308	0.760	0.0946
PreSG	0.691	0.1805	3.826	< 0.001	14.6394
Delayed Posttest Model Information		AIC = 336.628			R^2 = 0.209
Group[a]	0.864	0.8229	1.050	0.298	1.1030
Age	0.151	0.0599	2.529	0.014	6.3963
Gender[b]	-0.251	0.8345	-0.301	0.764	0.0907
Medu	0.370	0.3959	0.936	0.353	0.8752
PreSG	0.138	0.1287	1.069	0.283	1.1433

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, PreSG = Pretest story grammar

Table 4.7

Effects of the DR Program on Total Number of Words *at the Posttest and Delayed-Posttest Stages*

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 656.097		R^2 = 0.373	
Group[a]	-2.820	9.887	-0.285	0.776	0.0814
Age	1.037	0.642	1.615	0.112	2.6070
Gender[b]	-16.846	10.147	-1.660	0.102	2.7561
Medu	11.100	4.910	2.261	0.028	5.1111
PreTNW	0.503	0.112	4.484	< 0.001	20.1053
Delayed Posttest Model Information		AIC = 34.832		R^2 = 0.243	
Group[a]	0.10437	0.07712	1.353	0.181	1.832
Age	0.00489	0.00501	0.976	0.333	0.952
Gender[b]	-0.03966	0.07915	-0.501	0.618	0.251
Medu	0.05400	0.03829	1.410	0.164	1.988
PreTNW	0.00287	8.75e-4	3.286	0.002	10.796

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, Pre- = Pretest, TNW = Total number of words

Table 4.8

Effects of the DR Program on Proportion of Novel Words (Vocd) at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 35.525		R^2 = 0.331	
Group[a]	0.04304	0.07760	0.555	0.581	0.308
Age	0.02389	0.00500	4.778	< 0.001	22.826
Gender[b]	-0.14761	0.07910	-1.866	0.067	3.482
Medu	0.01809	0.03936	0.460	0.648	0.211
PreVocd	0.00617	0.00497	1.242	0.219	1.543
Delayed Posttest Model Information		AIC = 436.285		R^2 = 0.248	
Group[a]	1.139	1.867	0.610	0.544	0.372
Age	0.387	0.120	3.215	0.002	10.337
Gender[b]	-1.740	1.903	-0.914	0.364	0.836
Medu	-0.368	0.947	-0.389	0.699	0.151
PreVocd	0.279	0.120	2.332	0.023	5.438

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, Pre- = Pretest

Table 4.9

Effects of the DR Program on Verb Density at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = -264.3330		R^2 = 0.0801	
Group[a]	-0.01325	0.00738	-1.796	0.078	3.224
Age	-3.99e-4	4.82e-4	-0.827	0.411	0.684
Gender[b]	0.00509	0.00760	0.670	0.506	0.448
Medu	-0.00201	0.00375	-0.537	0.594	0.288
PreVds	0.05906	0.11517	0.513	0.610	0.263
Delayed Posttest Model Information		AIC = -236.0230		R^2 = 0.0374	
Group[a]	-7.54e-4	0.00920	-0.0819	0.935	0.00671
Age	1.73e-4	6.01e-4	0.2879	0.774	0.08287
Gender[b]	0.00138	0.00949	0.1451	0.885	0.02107
Medu	-0.00211	0.00467	-0.4524	0.653	0.20465
PreVds	0.16375	0.14368	1.1397	0.259	1.29895

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, PreVds = Pretest verb density

Table 4.10

Effects of the DR Program on Adverb Density at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = -235.9530		R^2 = 0.0795	
Group[a]	0.00203	0.00920	0.221	0.826	0.0487
Age	-9.43e-5	5.95e-4	-0.158	0.875	0.0251
Gender[b]	-0.01428	0.00931	-1.533	0.131	2.3495
Medu	0.00390	0.00452	0.862	0.392	0.7425
PreADVds	0.20497	0.13109	1.564	0.123	2.4448
Delayed Posttest Model Information		AIC = -249.5350		R^2 = 0.1410	
Group[a]	0.01964	0.00827	2.374	0.021	5.637
Age	-3.52e-4	5.35e-4	-0.657	0.513	0.432
Gender[b]	0.00952	0.00838	1.137	0.260	1.293
Medu	-0.00132	0.00407	-0.324	0.747	0.105
PreADVds	0.17436	0.11789	1.479	0.145	2.187

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, PreADVds = Pretest adverb density

Table 4.11

The Effects of the DR Program on Total Number of T-Units at Posttest and Delayed Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = -11.274		R^2 = 0.313	
Group[a]	-0.00197	0.05368	-0.0366	0.971	0.00134
Age	0.00562	0.00345	1.6287	0.109	2.65265
Gender[b]	-0.10535	0.05519	-1.9090	0.061	3.64424
Medu	0.06382	0.02648	2.4097	0.019	5.80685
PreTNT	0.02571	0.00682	3.7673	< 0.001	14.19238
Delayed Posttest Model Information		AIC = -1.985		R^2 = 0.319	
Group[a]	0.11327	0.05772	1.962	0.055	3.850
Age	0.00219	0.00371	0.591	0.557	0.349
Gender[b]	-0.05025	0.05934	-0.847	0.401	0.717
Medu	0.04539	0.02848	1.594	0.116	2.541
PreTNT	0.02916	0.00734	3.974	< 0.001	15.795

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, Pre- = Pretest, TNT = Total number of T-units

Table 4.12

The Effects of DR Program on Mean Length of Five Longest Utterances at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 266.744		$R^2 = 0.391$	
Group[a]	0.1164	0.4678	0.2489	0.804	0.06196
Age	0.0668	0.0329	2.0305	0.047	4.12294
Gender[b]	-0.0468	0.4944	-0.0946	0.925	0.00894
Medu	0.1663	0.2315	0.7185	0.475	0.51624
PreMLU5	0.4580	0.1062	4.3109	< 0.001	18.58391
Delayed Posttest Model Information		AIC = 276.979		$R^2 = 0.180$	
Group[a]	0.4370	0.5068	0.862	0.392	0.7437
Age	0.0856	0.0356	2.403	0.019	5.7732
Gender[b]	0.0673	0.5356	0.126	0.900	0.0158
Medu	0.0613	0.2507	0.244	0.808	0.0597
PreMLU5	0.1585	0.1151	1.377	0.174	1.8963

Note. a = DR compared with CR, b = Male compared with female, Medu = Maternal education, Pre- = Pretest, MLU5 = Mean length of five longest utterances

Table 4.13

Effects of the DR Program on Number of Conjunctions at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 375.291		$R^2 = 0.337$	
Group[a]	1.472	1.0977	1.34	0.185	1.80
Age	-0.151	0.0712	-2.11	0.039	4.47
Gender[b]	-1.839	1.1041	-1.67	0.101	2.77
Medu	0.593	0.5414	1.10	0.278	1.20
PreCONJ	0.350	0.0908	3.85	< 0.001	14.83
Delayed Posttest Model Information		AIC = 390.192		$R^2 = 0.242$	
Group[a]	1.645	1.2332	1.334	0.188	1.7783
Age	-0.126	0.0800	-1.577	0.120	2.4879
Gender[b]	1.837	1.2404	1.481	0.144	2.1923
Medu	0.117	0.6082	0.192	0.848	0.0370
PreCONJ	0.306	0.1020	2.994	0.004	8.9668

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, Pre- = Pretest, CONJ = Number of conjunctions

Table 4.14

Effects of the DR Program on Number of Different Conjunctions at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 196.062		$R^2 = 0.132$	
Group[a]	0.0460	0.2692	0.171	0.865	0.0292
Age	0.0201	0.0174	1.156	0.253	1.3355
Gender[b]	-0.0919	0.2732	-0.336	0.738	0.1131
Medu	-0.1808	0.1318	-1.372	0.175	1.8827
PreDfCONJ	0.3027	0.1363	2.221	0.030	4.9341
Delayed Posttest Model Information		AIC = 175.106		$R^2 = 0.102$	
Group[a]	0.1583	0.2286	0.6924	0.491	0.47944
Age	6.86e-4	0.0148	0.0464	0.963	0.00215
Gender[b]	-0.4266	0.2319	-1.8396	0.071	3.38428
Medu	-0.1026	0.1119	-0.9175	0.363	0.84188
PreDfCONJ	0.0811	0.1157	0.7012	0.486	0.49163

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, Pre- = Pretest, DfCONJ = The number of different conjunctions

Table 4.15

Effects of the DR Program on Character Speech at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 137.075		$R^2 = 0.301$	
Group[a]	0.0123	0.1700	0.0722	0.943	0.00521
Age	0.0262	0.0110	2.3827	0.020	5.67748
Gender[b]	-0.2617	0.1717	-1.5240	0.133	2.32264
Medu	0.0674	0.0845	0.7978	0.428	0.63645
PreCS	0.1074	0.0259	4.1391	< 0.001	17.13174
Delayed Posttest Model Information		AIC = 213.2200		$R^2 = 0.0936$	
Group[a]	-0.3171	0.3082	-1.029	0.308	1.059
Age	0.0285	0.0199	1.430	0.158	2.046
Gender[b]	-0.4216	0.3113	-1.354	0.181	1.834
Medu	0.1471	0.1532	0.960	0.341	0.922
PreCS	0.0427	0.0470	0.908	0.368	0.824

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, PreCS = Pretest character speech

Table 4.16

Effects of the DR Program on References to Internal States at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 116.786			R^2 = 0.120
Group[a]	0.26243	0.14613	1.796	0.078	3.2253
Age	-0.00110	0.00954	-0.116	0.908	0.0134
Gender[b]	-0.02153	0.14803	-0.145	0.885	0.0212
Medu	-0.02875	0.07092	-0.405	0.687	0.1644
PreRIS	0.06142	0.03629	1.692	0.096	2.8642
Delayed Posttest Model Information		AIC = 55.898			R^2 = 0.155
Group[a]	0.16120	0.09081	1.775	0.081	3.151
Age	0.00273	0.00593	0.461	0.647	0.212
Gender[b]	-0.05883	0.09199	-0.639	0.525	0.409
Medu	0.03771	0.04407	0.856	0.396	0.732
PreRIS	0.04441	0.02255	1.969	0.054	3.878

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, Pre- = Pretest, RIS = References to internal states

Table 4.17

Effects of the DR Program on Story-Grammar Questions at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 26.363			R^2 = 0.244
Group[a]	-0.01611	0.07175	-0.225	0.823	0.0504
Age	-0.01684	0.00494	-3.408	0.001	11.6151
Gender[b]	-0.07000	0.07214	-0.970	0.336	0.9414
Medu	0.01945	0.03495	0.556	0.580	0.3097
PreSGQ	-0.00731	0.00744	-0.982	0.330	0.9643
Delayed Posttest Model Information		AIC = -16.575			R^2 = 0.230
Group[a]	0.01848	0.05130	0.360	0.720	0.1298
Age	-0.01203	0.00353	-3.406	0.001	11.5999
Gender[b]	0.03484	0.05158	0.675	0.502	0.4561
Medu	-0.00478	0.02499	-0.191	0.849	0.0365
PreSGQ	-0.00441	0.00532	-0.829	0.411	0.6866

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, Pre- = Pretest, SGQ = Story grammar questions

Table 4.18

Effects of the DR Program on Problem-Solution Questions at the Posttest and Delayed-Posttest Stages

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Posttest Model Information		AIC = 69.378		R^2 = 0.185	
Group[a]	-0.2530	0.10031	-2.522	0.014	6.3601
Age	-0.0161	0.00656	-2.460	0.017	6.0528
Gender[b]	-0.0316	0.10128	-0.312	0.756	0.0974
Medu	0.0354	0.04910	0.722	0.473	0.5211
PrePRQ	-0.0201	0.04155	-0.483	0.631	0.2332
Delayed Posttest Model Information		AIC = 110.3560		R^2 = 0.0935	
Group[a]	-0.1949	0.13816	-1.411	0.164	1.991
Age	-0.0120	0.00903	-1.326	0.190	1.759
Gender[b]	0.1590	0.13949	1.140	0.259	1.298
Medu	-0.0543	0.06762	-0.803	0.425	0.644
PrePRQ	-0.0297	0.05723	-0.519	0.606	0.269

Note. a = DR compared with CR, b = Male compared with Female, Medu = Maternal education, Pre- = Pretest, PRQ = Problem-resolution questions

4.2.2 Growth Curve Analysis

To examine the changes in children's outcomes over time, a series of HLM analyses (Raudenbush & Bryk, 2002) were used. HLM takes the hierarchical structure of outcomes into consideration, allows for the control of the non-independence of observations, and assesses both fixed and random effects. In the present study, three-time measures were nested within each individual, and all individuals were nested within groups, which constituted the hierarchical data structure. With this structure, the non-independence of observation caused by repeated measures and group assignment was taken into account. In the present study, the *fixed effects* were the constant effects across individuals while

random effects referred to the variance in the intercepts and the slopes of individual growth curves and the covariance between them. In the present study, HLM analyses included two levels—within-subject (Level 1) and between-subject (Level 2). The two-level null model was written as follows:

Level 1: $Y_{ti} = \pi_{0i} + e_{ti}$

Level 2: $\pi_{0i} = \beta_{00} + r_{0j}$

In the present study, the Level 1 model depicted the pattern of changes in child outcomes over three time points: pretest, posttest, and delayed posttest. Predictors included in Level 1 were time and/or quadratic time. The inclusion of quadratic time as one predictor was appropriate because the changes of child outcomes over three time points might be best depicted by a quadratic model instead of a linear one. Also, quadratic change can be modelled in HLM with three points of data.

When including the time predictor only, the two-level null model is a linear model and written as follows:

Level 1:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME_{ti}) + e_{ti}$$

Level 2:

$$\pi_{0i} = \beta_{00} + r_{0j}$$

$$\pi_{1i} = \beta_{10} + r_{1j}$$

When including both time and quadratic time into Level 1, the model becomes a quadratic model and is written as follows:

Level 1:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME_{ti}) + \pi_{2i}(QUADRTIME_{i}) + e_{ti}$$

Level 2:

$$\pi_{0i} = \beta_{00} + r_{0j}$$

$$\pi_{1i} = \beta_{10} + r_{1j}$$

$$\pi_{2i} = \beta_{20} + r_{2j}$$

The two-level model with only Level 1 predictors included is also called an *unconditional-model* or a null-growth model. The decision to use a linear or quadratic model was taken after using a Chi-square test of deviance (Singer & Willett, 2003). If the deviance of the quadratic model was smaller than that of the linear time model and the difference between their deviance was significant in the Chi-square test, then the quadratic model was employed. Otherwise, the linear time model was retained. The unconditional-model analyses provide information about the linear and/or quadratic growth slopes when all the participants are treated as a single group. Moreover, this information can help reveal whether these slopes varied across individuals significantly.

The Level 2 model in the present study also included two predictors—group assignment and age—the functions of which depicted how these two factors predict the

Level 1 coefficients. The group assignment variable was included because the purpose of the HLM analyses was to detect differences between the two groups with regard to growth trajectories. Age was included because the preliminary analyses revealed that it significantly predicted almost all measures. While maternal education significantly and negatively predicted verb density and gender significantly predicted a few pretest outcomes in the preliminary analyses (see Section 4.1), these indicators were not included in the model because they were less influential than age. In addition, a larger sample size would be required if more predictors were to be included in the HLM analysis. Given the relatively small sample size in the present study, including gender and maternal education could decrease the model fit. Thus, in the Level 2 model, only group assignment and age were included, and the two-level linear model was built and written as follows:

Level 1:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME_{ti}) + e_{ti}$$

Level 2:

$$\pi_{0i} = \beta_{00} + \beta_{01}(GROUP_i) + \beta_{02}(AGE_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(GROUP_i) + \beta_{12}(AGE_i) + r_{1i}$$

For the quadratic model, the above functions were extended as follows:

Level 1

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME_{ti}) + \pi_{2i}(QUADRTIME_{ti}) + e_{ti}$$

Level 2:

$$\pi_{0i} = \beta_{00} + \beta_{01}(GROUP_i) + \beta_{02}(AGE_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(GROUP_i) + \beta_{12}(AGE_i) + r_{1i}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(GROUP_i) + \beta_{22}(AGE_i) + r_{2i}$$

In the data coding process, time was coded as 0, 1, and 2, and correspondingly, quadratic time was coded 0, 1, and 4. Following the analytical plan for the present study, the null model was run first to compute the intraclass correlation coefficient (ICC). According to Cohen (1988), HLM analysis is suggested for use when the ICC of the null model is larger than 0.059. Second, the two Level 1 predictors were added to the model one at a time. In this step, the fit of the linear and quadratic models was compared to determine which model should be used. This step also showed whether there was significant variability in the growth slope across individuals. Third, both Level 2 predictors were added to the model, which further explained whether there were group assignment and age differences in the growth slopes.

When the predictors were entered into the model, time and quadratic time were group mean centred while age was grand-mean centred. Group assignment was kept uncentred because this predictor is a binary variable. In addition, for the sake of conciseness, the indicators of narrative microstructure and narrative evaluation were normalized into [0, 1] and then amalgamated into one score in the curve analyses below,

labelled as MIC and NEV respectively. The scores on problem-resolution questions and story-grammar questions were combined into a composite score, labelled as narrative comprehension or COMQ.

4.2.2.1 HLM Results: Expressive Vocabulary.

The ICC of the null expressive-vocabulary model was 0.495 > 0.059, and thus it was suitable for HLM analysis. When comparing the linear with the quadratic model, the latter outperformed the former in model fitness, $\chi(3) = 56.01$, $p < 0.001$. Thus, the quadratic model was employed.

The results of the HLM analysis are shown in Table 4.19. They reveal no significant differences between the two groups in their pretest outcomes after controlling for age. However, older children did outperform younger children in expressive vocabulary at the pretest stage, with a mean difference of 0.904, $t = 7.388$, $p < 0.001$. The results also showed no differences in growth rate between the two groups, but younger children had a significantly higher growth rate, $t = -2.159$, $p < 0.036$. Older children had significantly higher acceleration with a mean difference of 0.334, $t = 2.355$, $p < 0.023$.

Table 4.19

HLM Results: Expressive Vocabulary

Fixed effects	Coefficients	*SE*	*t*-ratio	*p*-value
For EVT intercept, π_{0i}				
Intercept, β_{00}	70.275	1.349	52.086	<0.001
Group, β_{01}	-0.342	1.907	-0.179	0.858
Age, β_{02}	0.904	0.122	7.388	<0.001
For time slope, π_{1i}				
Intercept, β_{10}	-7.159	2.811	-2.546	0.013
Group, β_{11}	6.286	4.323	1.454	0.151
Age, β_{12}	-0.689	0.319	-2.159	0.035
For quadtime slope, π_{2i}				
Intercept, β_{20}	5.603	1.252	4.477	<0.001
Group, β_{21}	-1.613	1.983	-0.813	0.419
Age, β_{22}	0.334	0.142	2.355	0.022
Random effects	**Variance component**	***df***	χ^2	***p*-value**
EVT intercept, r_{0i}	56.602	61	506.840	<0.001
Time slope, r_{1i}	160.470	61	125.859	<0.001
Quadtime slope, r_{2i}	30.074	61	113.639	<0.001
Level 1 error, e_{1i}	23.234			
Deviance = 1319.297 with 7 *df*				

Note. EVT=Expressive Vocabulary Test

4.2.2.2 HLM Results: Story Grammar.

The ICC of the null story-grammar model was 0.211 > 0.059, and thus suitable for HLM analysis. The quadratic model had better fit than the linear model, χ (3) = 57.841, p < 0.001. Thus, the quadratic model was used.

The results of the HLM analysis are shown in Table 4.20. The results show a significant difference between DR and CR groups at the pretest stage with regard to story grammar favouring the experimental group, with a mean difference of 0.237, $t = 5.617$, $p < 0.001$, which is different from the results of the preliminary analyses. Moreover, older children outperformed their younger peers at the pretest stage, with a mean difference of 1.427, $t = 2.148$, $p < 0.037$. There were no significant group and age differences in the growth rate, and there were no group and age differences in acceleration.

Table 4.20

HLM Results: Story Grammar

Fixed effects	Coefficients	*SE*	*t*-ratio	*p*-value
For SG intercept, π_{0i}				
Intercept, β_{00}	22.755	0.496	45.875	<0.001
Group, β_{01}	0.237	0.042	5.617	<0.001
Age, β_{02}	1.427	0.664	2.148	0.036
For time slope, π_{1i}				
Intercept, β_{10}	-4.707	1.503	-3.131	0.003
Group, β_{11}	0.178	0.149	1.199	0.235
Age, β_{12}	0.837	2.118	0.395	0.694
For quadtime slope, π_{2i}				
Intercept, β_{20}	3.238	0.776	4.172	<0.001
Group, β_{21}	-0.100	0.076	-1.325	0.190
Age, β_{22}	-0.523	1.114	-0.469	0.641
Random effects	Variance component	*df*	χ^2	*p*-value
SG intercept, r_{0i}	4.953	61	213.119	<0.001
Time slope, r_{1i}	40.673	61	125.111	<0.001
Quadtime slope, r_{2i}	13.228	61	151.328	<0.001
Level 1 error, e_{1i}	5.959			
Deviance = 1026.833 with 7 *df*				

Note. SG=Story grammar

4.2.2.3 HLM Results: Narrative Microstructure.

The ICC of the null MIC model was 0.420 > 0.059, and therefore HLM analysis was appropriate. Different from Sections 4.2.2.1 and 4.2.2.2, the quadratic model did not have better fit than the linear model, and thus the linear model was retained. Table 4.21

presents the results of the HLM analysis. The results suggest that neither group assignment nor age accounted for the significant differences between outcomes at the pretest stage and in the growth rate.

Table 4.21

HLM Results: Narrative Microstructure

Fixed effects	Coefficients	*SE*	*t*-ratio	*p-value*
For MIC intercept, π_{0i}				
Intercept, β_{00}	3.036	0.114	26.701	<0.001
Group, β_{01}	0.327	0.180	1.814	0.075
Age, β_{02}	0.023	0.012	1.825	0.073
For time slope, π_{1i}				
Intercept, β_{10}	-0.182	0.098	-1.870	0.066
Group, β_{11}	0.193	0.124	1.543	0.128
Age, β_{12}	0.004	0.008	0.523	0.603
Random effects	**Variance component**	***df***	**χ^2**	***p*-value**
MIC intercept, r_{0i}	0.342	61	184.942	<0.001
Time slope, r_{1i}	0.007	61	60.410	>0.200
Level 1 error, e_{1i}	0.504			
Deviance = 500.258 with 4 *df*				

Note. MIC=Narrative microstructure

4.2.2.4 HLM Results: Narrative Evaluation.

The ICC for the null narrative evaluation (NEV) model was 0.336 > 0.059, and thus HLM analysis was applicable to this data set. The quadratic model yielded a better fit when compared to the linear model, χ (3) = 10.128, $p < 0.018$, and therefore was used in the analysis. The results of the HLM analysis are presented in Table 4.22. They indicate no significant differences between narrative-evaluation scores at the pretest stage, or in growth rate or acceleration, that can be attributed to group assignment or age.

Table 4.22

HLM Results: Narrative Evaluation

Fixed effects	Coefficients	*SE*	*t*-ratio	*p*-value
For NEV intercept, π_{0i}				
Intercept, β_{00}	0.431	0.039	11.065	<0.001
Group, β_{01}	0.090	0.058	1.535	0.130
Age, β_{02}	0.003	0.004	0.869	0.388
For time slope, π_{1i}				
Intercept, β_{10}	0.212	0.099	2.151	0.035
Group, β_{11}	0.034	0.141	0.240	0.811
Age, β_{12}	0.001	0.011	0.069	0.945
For quadtime slope, π_{2i}				
Intercept, β_{20}	-0.118	0.053	-2.229	0.030
Group, β_{21}	-0.022	0.070	-0.315	0.754
Age, β_{22}	-0.000	0.005	-0.068	0.946
Random effects	**Variance component**	***df***	**χ^2**	***p*-value**
NEV intercept, r_{0i}	0.037	61	202.509	<0.001
Time slope, r_{1i}	0.036	61	68.149	0.247
Quadtime slope, r_{2i}	0.009	61	68.827	0.229
Level 1 error, e_{1i}	0.048			
Deviance = 110.946 with 7 *df*				

Note. NEV=Narrative evaluation

4.2.2.5 HLM Results: Narrative Comprehension.

The ICC for the null narrative-comprehension model was 0.309 > 0.059, and thus HLM analysis was suitable. The quadratic model outperformed the linear model in fit, $\chi(3) = 28.931$, $p < 0.001$, and was employed in the ensuing analysis.

Table 4.23 summarizes the results of the HLM analysis. They show that older children outperformed their younger peers at the pretest stage, $t = 4.484$, $p < 0.001$, and had a significantly higher growth rate, $t = 2.224$, $p < 0.031$. However, older children had significantly lower acceleration than younger children with a mean difference -0.249, $t = -2.298$, $p < 0.026$. There were no significant differences between groups in the results for narrative comprehension at the pretest stage, or in growth rate or acceleration.

Table 4.23

HLM Results: Narrative Comprehension

Fixed effects	Coefficients	*SE*	*t*-ratio	*p*-value
For COMQ intercept, π_{0i}				
Intercept, β_{00}	29.947	0.575	52.089	<0.001
Group, β_{01}	0.254	0.911	0.279	0.781
Age, β_{02}	0.360	0.080	4.484	<0.001
For time slope, π_{1i}				
Intercept, β_{10}	-0.276	2.00	-0.138	0.891
Group, β_{11}	3.378	3.203	1.055	0.296
Age, β_{12}	0.513	0.231	2.224	0.030
For quadtime slope, π_{2i}				
Intercept, β_{20}	1.312	0.932	1.407	0.164
Group, β_{21}	-1.418	1.478	-0.959	0.341
Age, β_{22}	-0.249	0.108	-2.298	0.025
Random effects	**Variance component**	***df***	χ^2	***p*-value**
COMQ intercept, r_{0i}	11.324	61	248.464	<0.001
Time slope, r_{1i}	110.724	61	115.091	<0.001
Quadtime slope, r_{2i}	22.517	61	144.017	<0.001
Level 1 error, e_{1i}	11.058			
Deviance = 1151.781 with 7 *df*				

Note. COMQ=Narrative comprehension questions

In summary, the results of HLM analyses show that the experimental group was almost equivalent to the control group at the pretest stage except for the results on story grammar, which favoured the experimental group. Further, older children outperformed their younger peers at the pretest stage on scores for expressive vocabulary, story grammar, and narrative comprehension. There were no group differences in the growth rate for any measure. With respect to age, younger children had a higher growth rate than older children for expressive vocabulary, while the growth rate of older children in narrative comprehension was higher than that of younger children. In addition, older children had significantly higher acceleration in expressive vocabulary but lower acceleration in narrative comprehension.

4.3 Measures of Program Effects on Parents

4.3.1 Parental DR Reading Behaviours

The parent-child reading of four participants (two from DR group and two from CR group) was observed at five time points: during the pretest stage (T1), at Week 3 (T2), Week 8 (T3), Week 12 (T4), and during the delayed-posttest stage (T5). For each individual from the same group, the frequency of each type of DR reading behaviour was tallied. Then a group average was produced by adding up the frequencies of that DR behaviour and then dividing the sum by the number of individuals in the group. An average composite score for each group for each of the eight DR behaviours was also produced by calculating the

average of each DR behaviour for that group. The changes in the average composite score and the average frequency of each DR behaviour of each group across five time points are displayed in Figure 4.1.

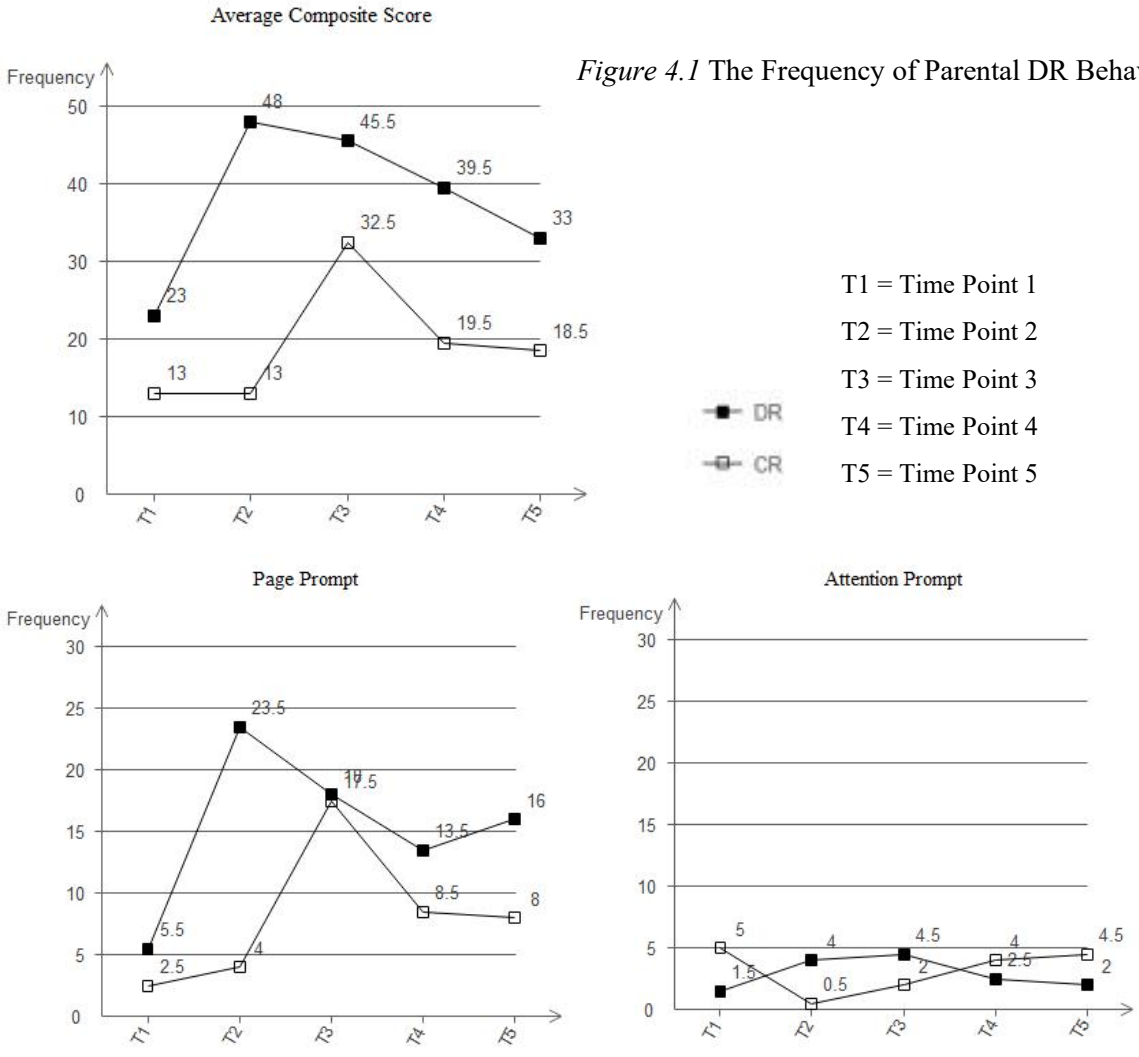
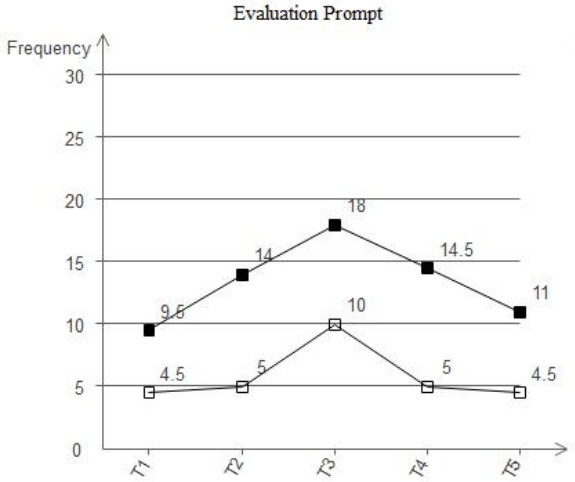
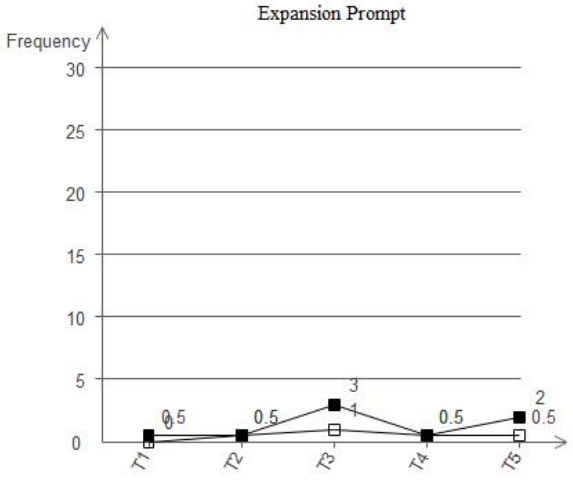
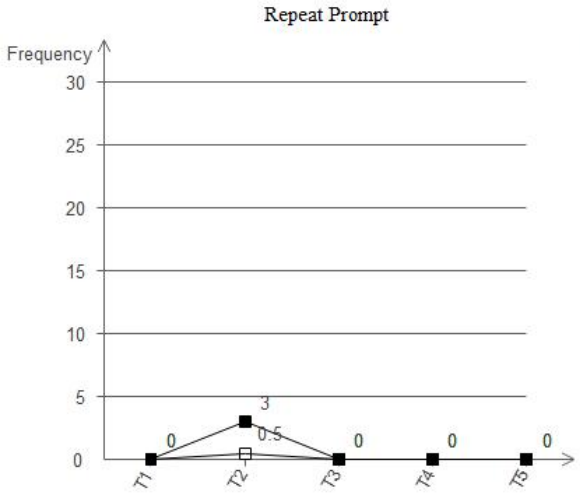
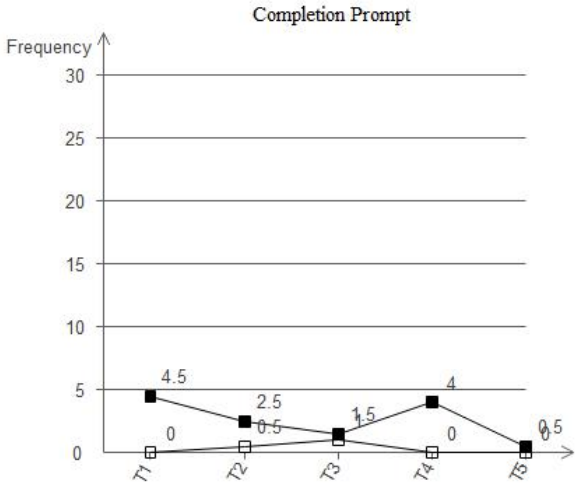


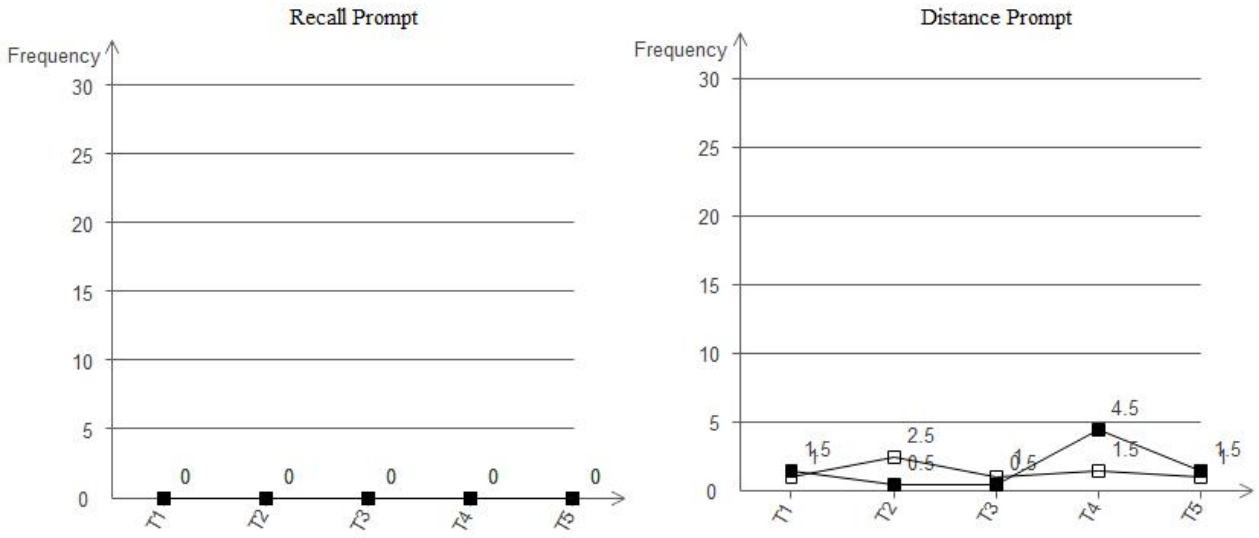
Figure 4.1 The Frequency of Parental DR Behaviours


Evaluation Prompt
Frequency
30
25
20
15
10
5
0
T1
T2
T3
T4
T5
9.5
14
18
14.5
11
4.5
5
10
5
4.5


Expansion Prompt
Frequency
30
25
20
15
10
5
0
T1
T2
T3
T4
T5
0.5
0.5
3
1
0.5
2
0.5


Repeat Prompt
Frequency
30
25
20
15
10
5
0
T1
T2
T3
T4
T5
0
3
0.5
0
0
0


Completion Prompt
Frequency
30
25
20
15
10
5
0
T1
T2
T3
T4
T5
4.5
2.5
0.5
1.5
4
0
0
0.5



The first graph in Figure 4.1 displays the changes in the average composite score for all types of DR behaviour by group. At the baseline time point (T1), the difference between the two groups was small, but this dramatically increased at T2. The average composite score of the DR group decreased slightly while that of the CR group increased greatly for some reason at T3. Therefore, the differences between the two groups narrowed a great deal at T3. After T3, the average composite scores of both groups decreased, with that of the CR group falling more rapidly. The differences between the two groups increased again after T3 but did not approach that at T2. Generally speaking, the DR behaviours of both groups increased first and then decreased; however, the participants from DR group experienced faster and earlier growth and a slower decrease, which suggests an advantage to the DR group.

The remaining eight graphs in Figure 4.1 present the changes in the average frequency of each DR behaviour type by group. A comparison of the eight graphs reveals that parents of both groups used page prompts and evaluation more often than the six other DR behaviour types. For example, the average frequency for the recall prompt by group was zero across five time points, and that for the repeat prompt was above zero only at T2. Except for the page and evaluation prompts, the average frequency of DR prompt types fell between zero and five. The differences between the two groups are much higher with regard to the frequency of page and evaluation prompts than for other DR behaviour types, indicating that those two types of prompts were more easily influenced by the DR program.

4.3.2 Program Fidelity

Parents were required to read at least four times per week with their child. Table 4.2 shows that the average frequency of shared reading per week across the whole program period was higher than four times per week, and this indicates that parents from both groups generally complied with the researcher's request. Of the 32 DR group participants, 29, 28, and 22 persons participated in the first, second, and third workshop respectively. The corresponding participation rates were 90.63%, 87.50%, and 68.75%. As shown in Figure 4.1, the increased frequency of some DR reading behaviours of parents in the DR group, such as page prompt, evaluation prompt, and the increased composite frequency of

all DR prompts, occurred more quickly than for the CR group. This suggests that participants from the DR group successfully learned the information presented in the workshops and applied this knowledge in their shared-reading sessions.

4.3.3 Follow-Up Questions About HLE and Parents' Rating of the Program

Twenty-nine out of 32 participants from the DR group (90.63%) returned their follow-up questionnaire. Several questions in the pretest family-information questionnaire completed at the pretest stage were also included in the follow-up questionnaire. Theoretically, the questions were grouped into seven categories based on HLE category groupings in the literature. Some categories included a single question, but most of the categories included more than one.

Three questions measured the child's interest in reading: questions on the child asking to be read to, their independent reading, and their pretending to read. Questions on visits to the library and to the bookstore building visiting were used to measure literacy outings. Questions on the child's DR reading behaviours, parental DR reading behaviours, parental DR reading beliefs, parental interest in shared reading, and parental confidence in shared reading were also re-assessed with participants in the DR group at the posttest stage. Seven paired-sample *t*-tests were carried out to detect the differences in parents' answers to these question before and after the program. The results showed that only the

scores on parental interest in shared reading with their child were significantly higher at the posttest stage than at the pretest stage, $t(28) = 3.684$, $p < 0.001$.

Many other questions on parents' rating of the DR program were also included in the follow-up questionnaire, including questions on child benefits from the DR program, parental benefits, parental changes in parent-child reading beliefs, the favourite and most useful part of the program, suggested changes to the program, barriers to participating in the program, and barriers to applying the DR techniques at home. These same questions were also used in individual interviews with seven parents from the DR group. In the following, information from parents' evaluation and rating of the program, drawn from the questionnaire and the interviews, is summarized and synthesized below.

4.3.3.1 Parental Benefits.

Of the respondents, 69% reported that the DR program was extremely beneficial to their shared-reading skills, and 27.6% rated it as very beneficial. One comment was, "Now I have learned how to read with my child, and during this process I can also inspire their thinking." Another participant said, "The program enhanced my shared-reading skills and my child's comprehension ability."

Regarding frequency of parent-child reading frequency, 62.1% of the respondents reported that the program was extremely beneficial in maintaining a satisfactory frequency of parent-child reading, and 24.1% rated it as very beneficial. One respondent

said that she was reading more frequently with her child. Another parent said that the program helped their family establish parent-child reading habits.

With respect to parental interest in shared reading, 62.1% of respondents reported that the program was extremely beneficial in cultivating their interest, and 27.6% reported it as very beneficial. Some participants also mentioned in the questionnaire and in the interview that the program improved their parent-child relationships. One parent said, "The shared- reading activity gives me a good opportunity to communicate more with my child." Several other parents reported that the program gave them more opportunity to connect with their child and therefore enhanced the feeling of attachment between them and their child. Two parents also mentioned that the books provided in the program improved their knowledge about children's books, and that they learned more about how to select high-quality books.

4.3.3.2 Child Benefits.

A set of questions was used to measure the benefits of the program for children. Sixteen parents (55.2%) reported that the program was extremely beneficial to their child's language development, 12 (41.38%) rated it as very beneficial, and one (3.4%) as "moderately beneficial." Many parents reported that their child gained considerable vocabulary knowledge in the 12-week program. One parent said that the program

improved their child's verbal ability. Another parent said, "My child learned to speak using correct sentence order."

Although the results of the paired-sample *t*-tests examining pre–post-program changes in the DR group did not find significant differences, parents' reports of change in their children suggest otherwise. For example, 58.7% of respondents reported that the program was extremely beneficial to their child's reading behaviour, and 37.9% rated it as very beneficial. One parent reported that their child loved interacting with them during the reading process more than before.

Regarding frequency of reading, 55.2% of the respondents reported that the program was extremely beneficial in improving their children's reading frequency, and 34.5% rated it as very beneficial.

With regard to their child's interest in reading, 55.2% of respondents reported that the program was very beneficial in increasing this, and 37.9% rated it as very beneficial. One participant said, "She now often reads on her own initiative, reads the familiar books again and again, has more enthusiasm for character recognition, and desperately wants to learn all the characters and then read to us." In addition, some participants mentioned that the reading, writing, and character recognition ability of their child had improved. Some other parents reported that the program was giving their children opportunities to practice their thinking and observation skills. One parent noted, "Before, my child tended to raise

irrelevant questions, but now many of her questions are about what she saw." Another parent said, "My child used to have difficulty in expressing his own feelings, but now he does better."

4.3.3.3 Useful and Favourite Features

Two groups of questions in the follow-up questionnaire contained sets of questions about the useful and favourite features of the program, as perceived by parents. The most useful item was the picture books provided. Twenty-two (75.9%) out of 29 respondents believed that the picture books were extremely useful, and the rest reported that they were fairly useful. Their second choice for usefulness was workshop instruction, and question lists and workshop outlines distributed to them along with the books: 72.4% reported each as extremely useful and 20–25% rated them as fairly useful. In the open-ended questions and interviews, however, the features most frequently mentioned as being useful were workshop instruction and program content. One parent mentioned that the CROWD questioning technique learned in the program helped her raise questions efficiently in the reading process. Some other features of the program, such as the checklist for reading frequency and duration and video modelling, were rated highly and positively.

As for favourite features of the program, nine parents (31.03%) chose the DR techniques taught in the program. One noted, "We learned how to ask our children questions [which is my favourite]." Six parents (20.69%) identified the interactions with

their child during the reading process as their favourite part of the program. One said, "Reading is very enjoyable. It can also promote parent-child interactions and our child's reading development, which makes it more enjoyable." Also, three parents (10.34%) regarded the picture books provided by the program as their favourite, while two other parents (6.90%) mentioned the question lists distributed to them along with the books.

4.3.3.4 Barriers to Participating and Suggested Changes to the Program.

Two key words contained in parents' responses on this topic were "work" and "busy." The workshops were scheduled on three Wednesdays in the late afternoons for the convenience of the kindergarten and most parents. Despite this, some people complained that the schedule was not friendly to them. One parent said, "We are very busy with our work and exhausted. We also have to send our child to other extra classes and keep pushing him to practise the skills every day. We really do not have a lot of time for this program." Another parent said, "I have regular meetings on Wednesday in my company [so that I cannot participate]." Other parents also reported unexpected events in their family or the sickness of their child. Generally, however, he biggest barrier that prevented parents from participating in the program was that they were so busy with their work that they could not take time off. Correspondingly, when parents were asked about their recommendations for future programs, the major change they suggested was changing to a time convenient to them.

4.3.3.5 Barriers to Applying DR Techniques at Home.

Parents reported a few different types of barriers. First, several parents reported that their children were not accustomed to the new techniques (DR). Some parents noted that their children did not like being questioned frequently during the reading process. Two parents noted that their child refused to answer questions. In contrast, one parent said, “It took time for my child to get used to the new method. At the beginning, he felt discomfort, and then he got used to it gradually.”

Second, some parents noted that they often made mistakes in applying the techniques or forgot to use them during the reading process. One parent noted, “Sometimes I used the techniques in a wrong way [at the beginning]. Later, I decided to prepare questions before the reading session and ask the questions directly during the reading process.” Two parents reported that their children had difficulty in focusing on the book. One parent stated that their child was deeply obsessed with giving correct answers and thus they asked the parent to read the same book repeatedly in order to be correct every time. Another parent noted that they had little time to read with their child.

In summary, the comparison between the pre- and posttest questionnaires showed that only parents’ interest in shared reading increased after the program among DR group participants. Yet the results obtained from parents’ ratings of the program in the follow-up questionnaire and interviews were generally high and positive.

Chapter 5: Discussion and Conclusion

The present study focused on the effects of a parent-child DR program on four- and five-year-old Mandarin-speaking children's expressive vocabulary and narrative competence. Further, changes in parental reading behaviours, as well as parents' perception of program impact and program participation were also explored. Randomized and controlled pre-, post-, delayed posttest trial was the predominant design of this study. In addition to individual assessments of children's expressive vocabulary and narrative competence, other measures including questionnaires, home video observations, and semi-structured interviews were also employed as data collections methods. The results showed that the DR program significantly improved the DR group children's expressive vocabulary and one indicator of narrative competence—adverb density—at the delayed posttest. No significant group differences were found across most indicators of narrative competence. Also, there were no significant group differences in the growth trajectories of the two dependent variables. For parental measures, the present study revealed positive changes in reading behaviours among DR participants. In addition, parents in the DR group gave high rating of the DR program in terms of program impacts and participation experience. In this chapter, the results of the present study, within the context of the entire research process, are further discussed.

5.1 Dialogic Reading and Expressive Vocabulary

In the present study, the positive effects of DR on the expressive vocabulary of Mandarin-speaking children's expressive vocabulary were significant even after controlling for influential demographic factors, HLE elements, the baseline level of dependent variables, and the frequency and duration of parent-child reading over the intervention period, and even when DR practices were used at home and not in a laboratory. This suggests that the effects of DR on young children's expressive vocabulary were robust, which is typical of many DR studies of English-speaking children (Hargrave & Sénéchal, 2000; Lonigan & Whitehurst, 1998; Mol et al., 2009; Sim et al., 2014; Whitehurst, Arnold, et al., 1994; Whitehurst, Epstein, et al., 1994).

It is interesting in that the DR group did not significantly outperform the CR group at the posttest stage but did do so at the delayed-posttest stage. In other words, it took longer than the 12-week duration of the program for the differences in expressive vocabulary between the two groups to reach a statistically significant level. There are two possible explanations for this result. First, it takes longer than 12 weeks for a DR program to exert significant effects on Mandarin-speaking children's expressive vocabulary. Another possible explanation is that the complex experimental factors—such as the sensitivity of the EVT-2 test to program effects, the books a parent chose to read with their child, and frequency and duration of reading at home, the rate of development of the

participants' expressive vocabulary and other language capacities—led to results not being significant at the posttest stage but becoming significant at the delayed-posttest stage. To be clearer, the translated EVT-2 test may not have been as culturally pertinent as the original English-language version, and so differences in expressive vocabulary between groups needed to be larger in order to be captured by the test. The quality of books owned by DR participants themselves might have been questionable: books that were beyond the limit of children's language ability, such as text books, or had little to do with language development, such as game books). For some DR participants, parent-child reading was relatively insufficient in frequency and duration. These factors, if present, might have led to the slow development of children's expressive vocabulary and the lack of significant results at the posttest stage. Also, it is possible that the expressive vocabulary of children participants in the present study developed so gradually that it took longer for EVT-2 to capture the differences between the two groups.

In contrast to the present study, Whitehurst et al. (1988) found that the immediate effects of DR on children's expressive vocabulary were significant at the posttest stage while the significance diminished in the follow-up test nine months later. The probable reason for the differences in the effects curve between the two studies is their complex discrepancies in the methodology and research design. For example, the two studies used participants who spoke different languages—English and Mandarin. The possibly

differing effects of DR for the two languages lead to further discrepancies in significance onset point—when significant effects appear—and in the duration of effect—when the effects of the DR program diminished. In the absence of other studies of DR on Mandarin-speaking children, however, such an explanation cannot be confirmed. Also, the child participants in Whitehurst et al. (1988) were between 21 and 35 months in age while the children in the present study were much older, between 37 and 71 months. Age has been found to have important effects of a DR program. One meta-analysis showed that younger children tend to benefit more from DR programming than their older peers (Mol et al., 2008). Children in Whitehurst et al.'s (1988) study were much younger and therefore possibly gained more and progressed faster from the DR program; thus, the significant effects of DR appeared after a short-term program (four weeks). The children in the present study, due to their older age, were less sensitive to the DR program, and thus, significant effects may not have manifested until the delayed-posttest stage (four months after the 12-week program). The late appearance of the effects may also indicate that time had a multiplicative effect for the older children. Another methodological disparity—treatment intensity—might also explain the difference in the effects curve of a DR program between the two studies. In the present study, the DR program lasted for 12 weeks, much longer than for Whitehurst et al.'s study (1988), which was four weeks long. For that reason, the effects of DR program in the present study were still highly

significant at the delayed-posttest stage (four months after the posttest stage and about seven months after the baseline stage), even though child participants were older and less sensitive to the DR program. The briefer DR program in the Whitehurst et al.'s study (1988) was less intense. Thus, although their program exerted significant immediate effects at the posttest stage, the effects had diminished at the delayed-posttest stage (nine months after the posttest stage and about 10 months after the baseline stage). Other discrepancies, such as the different measures used in the research, the different demographic backgrounds of participants, and the different DR training plan, might also have led to the differences in the results of the two studies.

5.2 DR and Narrative Competence

The effects of DR on Mandarin-speaking children's narrative competence were not significant on almost all aspects of macrostructure, microstructure, evaluation of narrative production, and narrative comprehension, which was unexpected. The experimental group outperformed the control group on only one dimension of microstructure—adverb density, at the delayed-posttest stage.

5.2.1 Possible Explanations for Non-Significant Results

One explanation for the lack of significant results is that narrative development is multifaceted, and progression in narrative competence is associated with children's cognitive maturity that develops over several years. Since the study was conducted within

a one-year time frame, it may be that the change in children's cognitive maturity was not substantive enough to impact their narrative skills to a measurable degree. Many studies have shown that the extent to which children's narrative competence can benefit from parent-child reading activities is dependent on their stage of cognitive and language development (Crain-Thoreson & Dale, 1992; Dale et al., 1996; Zevenbergen et al., 2003). For example, Dale et al. (1996) reported that children who performed less well on a test of cognitive abilities—including verbal, perceptual-performance, and quantitative abilities—acquired more vocabulary and verbal engagement through reading activities while children who scored higher on these tests showed greater increase in mean length of utterances (MLU). Another example, Zevenbergen et al. (2003) examined the effects of DR on a series of children's evaluative devices and found that children benefitted for only two indicators—references to internal states and character speech—and that the child participants were not able to add other devices, such as qualifying comments and causal statements, to their stories. Child participants in the present study, who fell in the similar age group as those participating in Zevenbergen et al. (2003) and Dale et al. (1996), possibly could not produce or comprehend some targeted narrative elements or indicators at this stage of development and within one-year program time frame.

Another reason for the lack of significant results might be the complexity of narrative competence, as suggested by its rich indicators. This complexity makes it harder

for children to improve their narrative ability from a single literacy activity such as DR. Possibly, DR cannot meet all the requirements for improving children's narrative competence, or it takes longer than the program lasted and requires more story input (in both type and quantity) to bring about significant changes.

The third possible explanation for the lack of significant results is that the DR program devised from the literature focused more on vocabulary than on narrative competence. The PEER procedure is very suitable for the correction and acquisition of vocabulary or short expressions instead of long narrative descriptions. Among the CROWD prompts, only recall questions and distancing questions ask children to retell the story or recall their personal experiences, and so were closely related to the narrative competence. Most prompts were more suitable for the acquisition of vocabulary or short expressions rather than for developing narrative language.

In addition, when parents applied the DR techniques, prompts that are closely connected with children's narrative competence—recall and distance prompts—were barely used, according to the observation results. This might have further decreased the effects of the DR program on children's narrative competence.

The fifth explanation might be that story-book reading is only part of the solution to enhancing narrative competence. For example, studies have shown that *narrative play*, also called pretend play or the story reenactment, can support story production and

comprehension (Baumer et al., 2005; Lillard et al., 2013; Roskos & Christie, 2003). To effect a significant improvement in children's narrative competence, it may be necessary to incorporate complementary techniques/strategies into the story-book reading process.

Last but not least, much is still unknown about narrative competence, and the measures of narrative competence in the literature including the one used in the present study are so far not well established. It is possible that the measures of narrative competence used in the present study could not capture the changes to this variable over the three time points (pretest, posttest, and delayed posttest).

5.2.2 Preliminary Analyses Results on Narrative Competence

In the preliminary analyses, family annual income of participants was found to negatively and significantly predict children's total number of conjunctions at the pretest stage; i.e., the higher the family annual income, the fewer conjunctions used by the child from that family. The negative correlation does not mean that the children from richer families had lower narrative competence. The children from lower-income families may have used "and" to join fragments because they had less sophisticated syntactic knowledge, which would have increased the number of conjunctions they used.

Also, in the preliminary analyses, maternal education was negatively correlated with children's verb density; i.e., the higher a mother's education, the lower the verb density of her child. This negative correlation does not mean that a child whose mother

had a higher level of education used fewer verbs in story production than their peers whose mothers had a less education, but rather that a child whose mother had higher education produced longer stories, which decreased the proportion of verbs in a story.

When program effects were analyzed, the children in the CR group significantly outperformed their peers in the DR group on problem-resolution questions, though no significant group differences were found on this measure at the delayed-posttest stage. One possible explanation is that the CR group increased their comprehension of story problems and their resolution over the 12-week period for some reason, e.g., their self-selected books contained more problem-resolution elements.

At the delayed-posttest stage, the DR group was also found to have higher adverb density than the CR group. A previous study (Kong & Fu, 2004) showed that most Chinese adverbs start to appear in children's oral languages between the ages of 18 and 54 months and that only a few were picked up after that point in time. The age of the participants in the present study ranged from 37 to 71 months, which overlaps the appearance age range of Chinese adverbs. It is possible that the DR program enhanced the adverb acquisition of the DR group during the program, especially for younger participants who were still in adverb appearance age range. This result also indicates that the DR program may have increased the proportion of content words in children's narrative production, of which adverb density is one index measure (Johansson, 2009).

5.2.3 Implications Despite the Lack of Significance

Although the DR program in the present study exerted significant effects on one indicator of narrative microstructure and a lack of significance was found on most narrative competence indicators, it extends the research literature on the potential benefits of the DR approach by examining a linguistic group not found in the extant literature—Mandarin-speaking children living in mainland China. It provides evidence for the beneficial effects of DR on Mandarin-speaking children's language development, especially their expressive vocabulary. It also demonstrates that children and their parents living in a city in a developing country can benefit from the DR approach. In particular, the DR program in the present study is family-focused as well as school-based. The positive effects of DR on children's vocabulary and parents' reports of beneficial effects on themselves and their children suggest that a school-based family-literacy program is worthwhile.

5.3 The Growth Curve of Child Outcomes

The HLM analyses revealed a difference between the two groups in children's story grammar at the pretest stage, in contrast to the results of the preliminary analyses, which found no group differences with regard to story grammar. The reason is that the HLM analyses took more factors into consideration—such as group, age, the hierarchical data structure, and random effects—while independent-sample *t*-tests in the preliminary

analyses only compared the mean of two groups without controlling for other factors. The absence of group differences on other measures at the pretest stage shows that the two groups were almost equivalent on most aspects of child outcomes, which is consistent with the results of the preliminary analyses.

The HLM analyses also revealed no group differences in growth rate and acceleration, indicating that child outcomes for both groups increased at the same speed and rate. In other words, the program did not make the child outcomes of the DR group grow faster. One possible reason for this is that age is the most important influence on four-to-five-year-old children's language outcomes, which can explain most of the change in growth rate and acceleration. At the same time, the results suggest that age differences existed in the growth rate and the acceleration of expressive vocabulary and narrative comprehension among child participants. This is consistent with the results of the preliminary analyses, according to which age strongly influenced children's language outcomes.

5.4 The DR Program and Parental DR Behaviours

The results of the observations carried out in participants' homes suggest that the DR behaviours of parents in the DR group increased faster and earlier compared to their peers in the control group. Further, the DR group parents showed a slower decrease in use of DR reading behaviours over time. These results indicate that participating parents

acquired DR techniques from the DR program, which is typical of many studies (Blom-Hoffman, 2007; Brannon & Dauksas, 2012; Beschorner & Hutchison, 2016; Whitehurst et al., 1988). For example, Beschorner and Hutchison (2016) found that parents can acquire DR techniques from both online and face-to-face programs. Brannon and Dauksas (2012) demonstrated that DR training was effective in enhancing parent-child reading interactions.

5.4.1 The Developmental Curve for the Usage of DR Techniques

When Blom-Hoffman's (2007) observation results are compared with that in the present study, the developmental curve for DR-group parents' usage of DR techniques are similar. In both studies, the use of DR techniques by parents in the DR group increased rapidly at first and then decreased slowly. The possible reason might be that at the beginning parents were affected by the novelty of the program and therefore were more motivated to utilize the newly learned techniques in parent-child reading. As time went by, the novelty effects decreased and so did their motivation to use the DR techniques. Or parents may have fallen back into their former habits after their initial heightened awareness and practice. This indicates the necessity to employ some techniques to maintain learning effects in a program. In the present study, the researcher sent weekly reminders via WeChat that asked parents to read at least four times per week. Individuals implementing a program in the future could take similar measures to remind parents to use the techniques. The use of

techniques for maintaining learning effects are also recommended for future programs that do not have a research purpose, such as online-contact video tips and ideas or a blog or social media group to keep the group connected. Such measures are not recommended for programs associated with literacy research because they increase the risk of design contamination, i.e., the treatment content is more easily obtained by control-group participants when it is talked about or posted online, a possible threat to research validity.

In the present study, the developmental curve for the use of DR techniques by parents in the CR group increased at T3 (Week 8) and T4 (at posttest), while in Blom-Hoffman's study (2007) the DR behaviours of the control group decreased slowly and consistently. The difference between the findings of the two studies might be explained by the possible contact between control- and experimental-group participants in the present study: participants from the CR group may have picked up some DR techniques through their interactions with parents from the DR group. This is possible since all child participants attended the same kindergarten. Also, other circumstances may have added to the usage of DR techniques by parents in the CR group. One possibility is that the reading materials used in the homes where observation occurred were not the same as those used by the other group and that the books read by CR parents in some observation sessions were more provocative than those read by the DR parents. Parents were expected to read the books distributed to them during the program; however, in

some home observation sessions, the assigned book was not available for some reason, such as the book having been left in another house out of town. During the third home observation (T3, right before the third workshop), #6 dyad from the DR group read a story whose text employed many long narrative sentences, while #18 dyad read a story in which the text was more conversational and included many situations suitable for questioning and labelling activities. During the fourth session of home observation (T4, during the posttest stage), the child in #55 dyad from the CR group was particularly familiar with and interested in the book being read, and he initiated many discussions during the session. In addition, the small sample size of home observations made the data set easily affected by the irregularities in story books and parent-child interactions. All these factors increased the frequency of use of DR techniques by parents in the CR group during the third and fourth home observations, explaining the increased use of DR techniques by CR participants at T3 and T4.

5.4.2 The Disparity in Usage of Prompts

Another finding of the present study is that frequency of use differed according to DR technique. It was obvious that parents from both groups used page and evaluation prompts much more frequently than other prompt types. The recall prompt was not used during any observation at any of the five time points when home visits took place. This result is similar to the findings in Blom-Hoffman's study (2007). Using the same coding

scheme as the present study, Blom-Hoffman also found that page and evaluation prompts were frequently used while others were not.

One possible explanation for this phenomenon is that page and evaluation prompts were already frequently used in daily conversation and sometimes in traditional storybook reading, and therefore were easily increased after DR instruction. Other types of prompts, however, such as recall, expansion, completion, repeat, and distance prompts, were rarely used in conversation. Thus, more efforts are needed in program design to feature the unfamiliar techniques more prominently and incorporate practice opportunities into the workshop to increase their frequency and flexibility of use in the home.

Another possible explanation for the limited use of some prompts is that the less frequently used prompts were harder for parents and children to master than the frequently used ones. Both page and evaluation prompts were part of the immediate reading environment, either related to the objects or events in the book or to the child's responses generated during the reading process. For this reason, they were less demanding of the parent's and children's language and cognitive capacity and thus were easier for them to generate and apply. In contrast, less frequently used prompts, such as recall, expansion, completion, repeat, and distance prompts, had higher language or cognitive demands and were more abstract or decontextualized, and thus harder for

parents and children to master. For example, one parent noted that recall questions were hard for their child to answer because their child could not remember the story plot.

5.5 Changes in HLE and Parents' Rating of the Program

The comparison between pretest and posttest questionnaires found that, among parents in the DR group, there were no differences on most follow-up HLE measures including child interest in reading, family-literacy outings, parental confidence about shared-reading, child DR reading behaviours, parental DR reading behaviours, and parental beliefs about DR reading. Significant difference was found for parental interest in shared reading. This indicates that DR parents became more interested in reading with their child after the DR program.

One possible explanation for the lack of significant results for the other measures is that parents tended to give positive answers in both questionnaires due to social desirability, which minimized the differences between questionnaires. *Social desirability* refers to participants' tendency to answer questions in a socially accepted manner (Bugental, Johnston, New, & Silvester, 1998). In contrast, many parents reported positive changes in answers to open-ended questions and during interviews that were not captured in the section of the follow-up questionnaire that measured HLE. As suggested by McMillan (2012), qualitative methods can provide rich descriptions of the phenomenon

being studied, contribute to a deeper understanding, and capture more information than quantitative methods.

Another possible explanation for the lack of significant results for some measures is that the follow-up questionnaire was answered at the end of the year—a time when everyone was busy completing job-related duties before Spring Festival holidays that usually begins in late January. Parents were likely too busy to spend much time on parent-child reading and practicing the new storybook reading techniques or visiting libraries or /bookstores. This interruption in practice may have eroded parents' confidence and children's interest in shared reading. This might explain the diminished effects of the DR program on participants' HLE. In the answers to the interview questions and the open questions of the follow-up questionnaire, many parents noted that they had less time than the beginning of this program to focus on reading activities with their child, because of laborious job-related duties before Spring Festival. For example, one parent said, "I used to go to the provincial library with my child...but now I barely have time to do so."

In their rating of the program, some parents reported that their child initially felt uncomfortable about entering into a dialogue during the reading process but then gradually got used to the new approach. Also, some parents reported that they often forgot to use the techniques and fell back to their former habits. These comments suggest that, when providing information and instruction, it is important for facilitators to tell parents

that they need to be patient with themselves and their children in learning and using DR techniques. One parent said that their child became better than before at expressing his emotions and feelings during the program. One possible reason is that children acquired words that express during the DR program. As suggested by Barrett (2019), the more emotion words an individual has, the better ability they have to perceive and describe their own emotions and those of others.

5.6 Program Fidelity

The fidelity of the DR program was of particular interest in the present study because it has the potential to be a threat to the research validity. *Program fidelity*, also called program integrity, refers to the extent to which the program is implemented as intended or the extent to which parents used the targeted program techniques (Blom-Hoffman, 2007; Powell & Carey, 2012). Some researchers pointed out that the small effects or the absence of effects found by many studies of family-literacy programs might indicate a low level of program fidelity (Powell & Carey, 2012), and DR programs are one type of family-literacy program. In other words, poor program fidelity might reduce the effects of the program and increase the risk of type II error.

Fidelity of program implementation is the result of multiple factors that affect not only the quantifiable aspects of program implementation but also the quality of the implementation. In their discussion of the issues associated with program fidelity, Powell

and Carey (2012) noted the myriad factors that influence the achievement of fidelity and the quality of its assessment. These include the complexity and specificity of the program, staff training and educational background, the characteristics of the organization implementing the program, the degree of impartiality of the sources of data used to assess fidelity, and the measures and methods of data collection. These factors comprise three critical dimensions of fidelity—the delivery of program components to parents, the receipt or understanding and involvement in programs by parents, and parents' enactment of intended behaviours in the home setting.

The present study took all the three dimensions of fidelity into consideration. First, the researcher took the responsibility for workshop delivery, and the DR workshops were devised based on the DR research literature. This ensured that the DR techniques delivered to parents were authentic, which strengthened the first dimension of fidelity—the delivery of program components to parents. Second, the present study collected data about workshop attendance using checklists, which was intended to monitor the second dimension of fidelity—the involvement of parents in the program. The relatively high attendance ratios reflect that the majority of the parents were involved in the workshop, which was a precursor of good understanding. Third, observations of parent-child reading in the home, weekly WeChat reminders about reading, and checklists for recording reading frequency and duration were utilized to collect information about

the third dimension of fidelity—parents' enactment of intended behaviours in the home setting. The obvious positive changes in the reading behaviours of parents in the DR groups indicated that they had indeed applied the techniques they learned in the workshops to their daily reading with their children. The consistent and stable reading frequency and duration of both groups were also satisfactory, showing that participants from both groups, at least on this aspect, followed the researcher's instructions. Using these data collection methods, the fidelity of the present study was monitored in a relatively holistic way.

One thing that is worth bearing in mind is that the inconsistent usage of DR techniques during parent-child reading by the DR group might have affected program fidelity in the present study. Participants from the DR group were expected to use the CROWD prompts and PEER procedures at a relatively high but appropriate frequency. However, only a few types of prompts were used frequently, and in this sense, DR techniques were not applied as intended. Yet, no study including the present one can achieve perfect fidelity in a real-world, family-focused context. Considering the multiple sources and positive evidence of program fidelity, it is reasonable to conclude that the program fidelity in the present study was very good, overall.

5.7 Other Possible Threats to Research Validity

One possible threat to the present study is the non-equivalency between the experimental (DR) and the control (CR) group. Many actions were taken to ensure group equivalency including randomization, control of demographic and HLE factors, control of discrepancies in dependent variables at the baseline, entering important influencing factors into the regression model as covariates, providing the same books to the two groups, and control of reading frequency and duration of parent-child reading over the course of the program. In one sense, the sampling bias resulting from group assignment was minimized to the extent possible. Although it is still hard to claim that the two groups were fully or perfectly equivalent, it is reasonable to conclude that the non-equivalency between groups was highly controlled. Further, the DR program largely accounted for any significant discrepancies with dependent variables at the posttest and delayed-posttest stages.

The second possible threat to the research validity is the *Hawthorne effect*, which refers to how participants modify their behaviour because they know their group is receiving special treatment and, further, that the modifications in their behaviours lead to discrepancies between them and the baseline or the control group (Adair, 1984; McCarney et al., 2007). To minimize the impact of the Hawthorne effect, participants in the present study were told that all would receive the same training but that, due to the

limited number of slots in the program, only a certain number of persons could attend the training at one time. They were also told that, to ensure everyone received high-quality training, they would attend the workshop series in two groups—one in the first semester (DR group) and the other in the second semester (CR group)—one academic year consists of two semesters in China. In addition, families in both groups received the same books and checklists for reading frequency; thus, advising parents they would receive the same experiences was expected to ease the feeling of being specially treated among families in the DR group. Although it is impossible to say that no one in the experimental group knew that they were in a specially treated group, based on the experiences of the researcher during home visits and workshop instruction, no parents in the DR group expressed the idea of being in a special group. Only a few parents in the CR group showed sadness that they could not attend the workshop earlier. Thus, there is good reason to believe that the Hawthorne effect did not impact the present study.

Another possible threat to the research validity is *design contamination*; i.e., participants in the control group in some way obtain access to treatment content that should only be available to the experimental group (Keogh-Brown et al., 2007; Torgerson, 2001). Design contamination is harmful because it decreases the effects of the treatment and increases the risk of type II error—an effective treatment is rejected because of non-significant results (Keogh-Brown et al., 2007). To reduce the potential impact of

design contamination, participants in the control group (CR) who received the training in the second semester were told that receiving the training a little bit later would not do harm to their child and that the contents would be the same as for the other group that would receive the training earlier. Participants in the experimental group (DR) were not given a copy of the Powerpoint presentation and were asked not to take photos of it. A paper copy of the content summary was distributed to them at the end of each workshop. Also, the researcher tried to keep group assignments and the list of participant names confidential, not only for ethical reasons but also in order to eliminate the opportunity for cross-group communications about program content. In addition, for the convenience of participants, the researcher contacted them using WeChat, but any requests from the participants to build a WeChat group for the workshop class or for the two training groups were rejected. This is because information control becomes harder when people gather together, especially online. These actions were taken to reduce the risk of design contamination. Yet, given that all the participants in the present study came from the same kindergarten, it might be hard to completely avoid the risk of design contamination.

As suggested by some researchers, self-report has been commonly used in parenting research (Manz Hughes, Barnabas, Bracaliello, & Ginsburg-Block, 2010; Morsbach & Prinz, 2006). *Social desirability*—participants' inclination to answer questions in a socially accepted manner (Bugental et al., 1998)—is another possible threat

to research validity resulting from the self-report method. In the present study, the reading frequency and duration checklist, home interview and observation, and pre- and posttest questionnaire are all self-report to some extent. The use of these data collection methods might bring the threat of social desirability to the present study.

As suggested above, parents' responses to the HLE questions in the questionnaires they completed at the pretest and posttest stages were likely influenced by social desirability so that the comparison of the answers to the two questionnaires revealed no significant differences. When completing the checklists for the reading frequency and duration, both groups were likely again influenced by social desirability, which impacted their responses and the interpretation of program effects. The degree of social desirability on interviews and observations in the home, and on parents' ratings of the program in the follow-up questionnaire can be evaluated by the consistency in the results obtained through these measures and child measures. Certain studies show that self-report by parents can have high correlations with parent and child outcomes and therefore good instrumental validity (DeBaryshe & Binder, 1994; Johnston, Scoular, & Ohan, 2004). The present study collected data on both parents and child outcomes and employed different data collection methods such as questionnaires, video observation, interviews, and standard assessments. The triangulation of both data sources and data collection methods were expected to monitor the risk of social desirability—i.e., by evaluating the

consistency of research results. The results for both parent and child outcomes obtained through these data collection methods were consistent to some extent (except for the HLE questions in the follow-up questionnaire), indicating that the threat of social desirability was still under control.

5.8 Limitations and Future Directions

The first limitation of the present study is that DR techniques were provided as a whole package so that the unique effects of each technique on children's expressive vocabulary and narrative competence are still unknown. It is possible that some techniques are more effective than others and also that some techniques are redundant. In the future, researchers interested in this topic might consider more nuance in testing the efficacy of different techniques.

The second limitation stems from the program content. In the analyses of the video recordings of observations in the home, it is apparent that parents used some techniques more frequently than others and that the differences in frequency were quite large. To exploit the potential of the DR approach to the fullest, in future research it might be good practice to design the DR training more carefully so that different techniques that fall under the DR umbrella can be enhanced together in the program. For example, one possible solution can be to add to the instructional plan for the program intensive training on the techniques that have been shown in this and previous studies to be hard for parents

to master. Also, it might be interesting for some DR researchers to explore the underlying reasons for the varying difficulty of different techniques.

The third limitation lies in the measures used in the present study. The high Cronbach's α of the translated EVT-2 test (0.95 for form A and 0.89 for form B) in the present study indicates that the test had good reliability when used with participants involved in the present study. Also, the test captured the discrepancies between groups at the delayed-posttest stage. However, given that the test was translated, it is not possible to use its original standard score and norms to extract more information from the data, e.g., comparing children's growth in each group with the norm, or, with other studies using the same measure and norms. Also, it is possible that, if using a test developed specifically for Mandarin-speaking children in mainland China, more discrepancies between groups can be captured. In terms of the measure of narrative competence, the adapted ENNI instrument used in the present study has been a valuable tool and is used by many researchers because of its value to the research community. Nonetheless, the state of the research on narrative competence is still evolving; therefore, existing measures such as ENNI may not yet be capturing the nuanced dimensions of narrative. In the future, greater efforts are needed to increase the knowledge of Mandarin-speaking children's language development and to develop accurate measures for Mandarin-speaking children based on that knowledge.

Purcell-Gates (2000) developed a framework for categorizing family literacy programs. Her work was informed by Nickse (1991) who noted that family-literacy program usually have multiple objectives and use different types of methods to achieve these goals, and Hayes (1996) who contended that a comprehensive program evaluation should be connected to these objectives and goals. Purcell-Gates framework of program goals and evaluation can be divided into three categories: (a) child behavioural, cognitive, and emotional outcomes; (b) parental cognitive, behavioural, and emotional outcomes, and (c) literacy interactions between parents and child. In the present study, both parent and child outcomes were studied while parent-child interactions were not. In future studies or in the further exploration of the data collected for the present study, parent-child interactions might be a topic of interest.

Other limitations and possible future directions are also noteworthy. First, as stated by Whitehurst et al. (1988, p. 558), developmental psycholinguistics tends to define its subjects as the "lowest common denominator," which might lead to the neglect of social, cultural, and individual differences in language acquisition. In the future, researching language from different perspectives in a DR program might be a topic of interest for some researchers. Second, to avoid the impact of design contamination, the DR program can be conducted in two or more homogeneous institutions instead of in a single one. Third, for certain of the five observations of parent-child reading observations,

the reading materials for each group were not the same. This circumstance might have influenced the results of the observations, posing a risk to research validity, given that only four parent-child dyads participated in the observations. In future studies, it should be ensured that equivalent reading materials is used for each observation of each parent-child reading observation. Fourth, the present study suggests that DR might have a limited effect on all aspects of narrative development. Reese et al. (2010) found that training parents to use *elaborative reminiscing*—encouraging child to talk about past events—had a stronger effect on children's narrative skills than DR and suggested that elaborative reminiscing can supplement a DR program in preschool classrooms. Surely, if the program is implemented for research purposes, the role of child's memory of past events in such a mixed program should be carefully considered.

5.9 Conclusion

Although the effects of DR on English-speaking children's expressive vocabulary have been well researched, only a few studies focus on its effects on children's narrative competence and Mandarin-speaking children's expressive vocabulary. Yet, due to the methodological or theoretical limitations, the DR effects on these two language variables are still in need of further exploration. To examine this matter, the present study examined the efficacy of a family-focused DR program with 81 four- to-five-year-old kindergartners in a medium-sized city in southwest China. The findings of this study indicate that

Mandarin-speaking children's expressive vocabulary was significantly improved by a 12-week DR program. The use of DR techniques also significantly improved one aspect of children's narrative competence—adverb density. However, no significant treatment effects were found on other aspects of narrative competence. In addition, there was no evidence of group differences in the growth rate of child outcomes over the three measurement points, indicating that the DR treatment did not have a significant effect on the growth curve of children's expressive vocabulary and narrative competence. The present study also found that parents can acquire DR techniques through attendance at three workshops focused on DR and narrative skills. Future studies should aim to further refine measures suitable to the Mandarin-speaking population. The refined measures can be further used to explore the effects of individual DR techniques or the impact of a DR approach that incorporates supplementary programs that develop narrative competence or that add special intensive training on infrequently used DR techniques.

References

Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of Applied Psychology*, *69*(2), 334–345. doi:10.1037/0021-9010.69.2.334

Agdam, S. J., & Sadeghi, K. (2014). Two formats of word association Tasks: A study of depth of word knowledge. *English Language Teaching, 7*(10), 1–12. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1075954.pdf

Anderson, R. C., Gaffney, J. S., Wu, X., Wang, C., Li, W., Shu, H., Nagy, W. E., & Ming, X. (2002). Shared book reading in China. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp.131–155). Norwell, MA: Kluwer Academic Publishers Group.

Applebee, A. N. (1978). *The child's concept of story: Ages two to seventeen* (Midway reprinted.). Chicago, IL: University of Chicago Press.

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86*(2), 235–243. doi:10.1037//0022-0663.86.2.235

Arnold, D. S., & Whitehurst, G. J. (1994). Accelerating language development through picture book reading: A summary of Dialogic Reading and its effect. In D. K. Dickinson (Eds.) *Bridges to literacy: Approaches to supporting child and family literacy* (pp. 103-128). Cambridge, MA: Blackwell Publishing.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority and youth.* Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, S., Simmons, D. S., & Kame'enui, E. J. (1998). Vocabulary acquisition: Research bases. In D. S. Simmons & E. J. Kame'enui (Eds.), *What reading research tells us about children with diverse learning needs* (pp. 183–210). Mahwah, NJ: Erlbaum.

Bamberg, M. G. (1997). *Narrative development: Six approaches*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Bamberg, M. G. W. (1987). *The acquisition of narratives: Learning to use language*. Berlin: Mouton de Gruyter.

Bamberg, M., & Damrad-Frye, R. (1991). On the ability to provide evaluative comments: Further explorations of children's narrative competencies. *Journal of Child Language*, *18*(3), 689–710. doi:10.1017/S0305000900011314

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Bartlett, F. C., Bartlett, F. C., & Kintsch, W. (1995). *Remembering: A study in experimental and social psychology* (Vol. 14). New York, NY: Cambridge University Press.

Baumann, J. F. (2009). Vocabulary and reading comprehension: The nexus of meaning. In S. E. Iseal & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. NY & London: Routledge.

Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (Vol.2, pp.752–785). Mahwah, NJ: Lawrence Erlbaum Associates.

Baumer, S., Ferholt, B., & Lecusay, R. (2005). Promoting narrative competence through adult–child joint pretense: Lessons from the Scandinavian educational practice of playworld. *Cognitive Development*, *20*(4), 576–590. doi:10.1016/j.cogdev.2005.08.003

Brannon, D., & Dauksas, L. (2012). Studying the effect dialogic reading has on family members' verbal interactions during shared reading. *SRATE Journal*, *21*(2), 9–20. Retrieved from: https://files.eric.ed.gov/fulltext/EJ990631.pdf

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251–271. doi:10.1086/511706

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506–521. doi:10.1037/0022-0663.74.4.506

Beck, R. J., & Clarke-Stewart, K. A. (1998). Improving 5-year-olds' narrative recall and comprehension. *Journal of Applied Developmental Psychology*, *19*(4), 543–569. doi:10.1016/S0193-3973(99)80055-0

Beschorner, B., & Hutchison, A. (2016). Parent education for dialogic reading: Online and face-to-face delivery methods. *Journal of Research in Childhood Education*, *30*(3), 374–388. doi:10.1080/02568543.2016.1178197

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44–62. doi:10.1037/0022-0663.98.1.44

Blachowicz, C. L. Z., & Fisher, P. (2000). *Vocabulary instruction*. In M. L. Kamil, , P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol.3, pp. 503–524). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Blom-Hoffman, J., O'Neil-Pirozzi, T., Volpe, R., Cutting, J., & Bissinger, E. (2007). Instructing parents to use Dialogic Reading strategies with preschool children: Impact of a video-based training program on caregiver reading behaviors and children's related verbalizations. *Journal of Applied School Psychology, 23*(1), 117–131. doi:10.1300/J370v23n01_06

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Buckner, J. P., & Fivush, R. (1998). Gender and self in children's autobiographical narratives. *Applied Cognitive Psychology, 12*(4), 407–429. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.7787&rep=rep1&type=pdf

Bugental, D. B., Johnston, C., New, M., & Silvester, J. (1998). Measuring parental attributions: Conceptual and methodological issues. *Journal of Family Psychology*, *12*(4), 459-480. Retrieved from: https://www.researchgate.net/profile/Charlotte_Johnston/publication/232538475_Measuring_parental_attributions_Conceptual_and_methodological_issues/links/5c630d40299bf1d14cc1eaa7/Measuring-parental-attributions-Conceptual-and-methodological-issues.pdf

Burger, A., & Chong, I. (2011). Receptive vocabulary. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (pp. 1231–1231). Springer US. doi:10.1007/978-0-387-79061-9_2359

Bus, A., van Ijzendoorn, M., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65*(1), 1–21. doi:10.3102/00346543065001001

Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology*, *21*(3), 335–351. doi:10.1348/026151003322277739

Cain, K., & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology*, *14*(2), 187–201. doi:10.1111/j.2044-835X.1996.tb00701.x

Cameron, C. A., Hunt, A. K., & Linton, M. J. (1988). Medium effects on children's story rewriting and story retelling. *First Language, 8*(22), 3–18. doi:10.1177/014272378800802201

Carlisle, F. J., & Katz, L. A. (2005). Word learning and vocabulary instruction. In J. R. Birsh (Eds.), *Multisensory teaching of basic language skills* (pp. 345–375). Baltimore, MD: Paul H. Brookes Publishing Company.

Caspe, M., & Melzi, G. (2008). Cultural variations in mother-child narrative discourse style. In A. McCabe, A. L. Bailey, & G. Melzi (Eds.), *Spanish-language narration and literacy: Culture, cognition, and emotion* (pp. 6–33). New York, NY, US: Cambridge University Press. doi:10.1017/CBO9780511815669.004

Chang, C., & McCabe, A. (2013). Evaluation in Mandarin Chinese children's personal narratives. In A. McCabe, & C. Chang (Eds.), *Chinese language narration: culture, cognition, and emotion* (pp.33–56). Concentric: Studies in Linguistics. Retrieved from http://ebookcentral.proquest.com

Cheung, H. T. (1998). 平均語句長度在中文的應用 [The use of the mean length of utterance in Chinese language]. *The journal of Speech-language-hearing association/聽語會刊*, (13), 36–48. doi:10.6143/JSLHAT.1998.06.04

Chow, B. W. Y., & McBride-Chang, C. (2003). Promoting language and literacy development through parent-child reading in Hong Kong preschoolers. *Early Education and Development, 14*(2), 233–248. doi:10.1207/s15566935eed1402_6

Chow, B. W. Y., McBride-Chang, C., & Cheung, H. (2010). Parent-child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. *Journal of Research in Reading, 33*(3), 284–301. doi:10.1111/j.1467-9817.2009.01414.x

Chow, B. W. Y., McBride-Chang, C., Cheung, H., & Chow, C. (2008). Dialogic Reading and morphology training in Chinese children: Effects on language and literacy. *Developmental Psychology, 44*(1), 233–244. doi:10.1037/0012-1649.44.1.233

Christiansen, J. A. (1995). Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, *51*(2), 291–317. doi:10.1006/brln.1995.1062

Clarke-Stewart, K. A., & Beck, R. J. (1999). Maternal scaffolding and children's narrative retelling of a movie story. *Early Childhood Research Quarterly*, *14*(3), 409–434. doi:10.1016/S0885-2006(99)00018-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Crain-Thoreson, C., & Dale, P. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19*(1), 28–39. doi:10.1177/027112149901900103

Cronan, T. A., Cruz, S. G., Arriaga, R. I., & Sarkin, A. J. (1996). The effects of a community-based literacy program on young children's language and conceptual development. *American Journal of Community Psychology, 24*(2), 251–272. doi:10.1007/BF02510401

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945. Retrieved from: https://www.researchgate.net/profile/Anne_Cunningham2/publication/13845653_Early_Reading_Acquisition_and_Its_Relation_to_Reading_Experience_and_Ability_10_Years_Later/links/02bfe510ff38d02c58000000/Early-Reading-Acquisition-and-Its-Relation-to-Reading-Experience-and-Ability-10-Years-Later.pdf

Curenton, S. M., & Justice, L. M. (2004). African American and Caucasian preschoolers' use of decontextualized language: Literate language features in oral narratives. *Language, Speech, and Hearing Services in Schools*, *35*(3), 240–253. doi:10.1044/0161-1461(2004/023)

Cyr, L., & Francis, K. (1992). Measures of clinical agreement for nominal and categorical data: The kappa coefficient. *Computers in Biology and Medicine, 22*(4), 239–246. doi:10.1016/0010-4825(92)90063-S

Dale, P. S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education*, *16*(2), 213–235. doi:10.1177/027112149601600206

Davies, P., Shanks, B., & Davies, K. (2004). Improving narrative skills in young children with delayed language development. *Educational Review*, *56*(3), 271–286. doi:10.1080/0013191042000201181

Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika, 9*(3), 185-197. Retrieved from: http://iapsych.com/wmfhcaarchive/LinkedDocuments/DAVI11.pdf

Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, *3*(4), 499-545. Retrieved from: http://www.jstor.org/stable/747153

DeBaryshe, B. D., & Binder, J. C. (1994). Development of an instrument for measuring parental beliefs about reading aloud to young children. *Perceptual and Motor Skills*, *78*(3_suppl), 1303–1311. doi:10.2466/pms.1994.78.3c.1303

DeLoache, J. S., & DeMendoza, O. A. (1987). Joint picture book interactions of mothers and 1-year-old children. *British Journal of Developmental Psychology*, *5*(2), 111–123. doi:10.1111/j.2044-835X.1987.tb01047.x

Desmarais, C., Nadeau, L., Trudeau, N., Filiatrault-Veilleux, P., & Maxès-Fournier, C. (2013). Intervention for improving comprehension in 4-6 year old children with specific language impairment: Practicing inferencing is a good thing. *Clinical Linguistics & Phonetics*, *27*(6–7), 540–552. doi:10.3109/02699206.2013.791880

Deuchar, M. (n.d.). The Miami corpus: Documentation file. Bangor. Retrieved from: http://bangortalk.org.uk/docs/Miami_doc.pdf

Deuchar, M., Webb-Davies, P., & Donnelly, K. (2018). *Building and using the Siarad corpus: Bilingual conversations in Welsh and English*. Holland: John Benjamins Publishing Company.

Dickinson, D. K., & Snow, C. E. (1987). Interrelationships among prereading and oral language skills in kindergartners from two social classes. *Early Childhood Research Quarterly, 2*(1), 1–25. doi:10.1016/0885-2006(87)90010-X

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, MD: Paul H. Brookes Publishing.

Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology*, *34*(1), 83–98. doi:10.1007/s10802-005-9003-x

Doyle, A. R. L. (2009). *A family literacy program with varying participation structures: Effects on parent and child learning* [Doctoral dissertation, University of Toronto].

Draper, A. G., & Moeller, G. H. (1971). We think with words (therefore, to improve thinking, teach vocabulary). *The Phi Delta Kappan*, *52*(8), 482–84. Retrieved from: http://www.jstor.org/stable/20372973

Eisenberg, S. L., Fersko, T. M., & Lundgren, C. (2001). The use of MLU for identifying language impairment in preschool children: A review. *American Journal of Speech-Language Pathology, 10*(4), 323–342. doi:10.1044/1058-0360(2001/028)

Ezell, H. K., & Justice, L. M. (2005). *Shared storybook reading: Building young children's language & emergent literacy skills*. Baltimore, MD: Brookes.

Farrell M. L. & White N. C. (2018). Structured literacy instruction. In J. R. Birsh & S. Carreker (Eds.), *Multisensory teaching of basic language skills* (4th ed.). Baltimore, Maryland: Paul H. Brookes Publishing Company.

Feagans, L., & Appelbaum, M. I. (1986). Validation of language subtypes in learning disabled children. *Journal of Educational Psychology*, *78*(5), 358–364. doi:10.1037/0022-0663.78.5.358

Feagans, L., & Short, E. J. (1984). Developmental differences in the comprehension and production of narratives by reading-disabled and normally achieving children. *Child Development, 55*(5), 1727–1736. Retrieved from: http://www.jstor.org/stable/1129920

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i+iii-v+1-185. doi:10.2307/1166093

Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning, 18*(2), 77–82. doi:10.1111/1467-9604.00284

Fleury, V. P., & Schwartz, I. S. (2017). A modified dialogic reading intervention for preschool children with autism spectrum disorder. *Topics in Early Childhood Special Education*, *37*(1), 16–28. doi:10.1177/0271121416637597

Frank, C. K., Baron-Cohen, S., & Ganzel, B. L. (2015). Sex differences in the neural basis of false-belief and pragmatic language comprehension. *NeuroImage*, *105*, 300–311. doi:10.1016/j.neuroimage.2014.09.041

Fung, P. C., Chow, B. W. Y., & McBride-Chang, C. (2005). The impact of a Dialogic Reading program on deaf and hard-of-hearing kindergarten and early primary school–aged students in Hong Kong. *Journal of Deaf Studies and Deaf Education, 10*(1), 82–95.doi:10.1093/deafed/eni005

Given, L. (Eds.). (2008). Open coding. In *The SAGE encyclopedia of qualitative research methods*. SAGE. doi:10.4135/9781412963909.n299

Geva, E., & Olson, D. (1983). Children's story-retelling. *First Language, 4*(11), 85–109. doi:10.1177/014272378300401102

Glosser, G., & Deser, T. (1991). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, *40*(1), 67–88. doi:10.1016/0093-934X(91)90117-J

Gong, Y. X., Dai, X. Y. (1992). *中国-韦氏幼儿智力量表（C-WYCSI）手册 [Manual of Chinese-Weschler young children scale of intelligence (C-WYCSI)]*. Changsha, PRC: Hu'nan Map Press/长沙，中国：湖南地图出版社.

Gough, P. (1996). How children learn to read and why they fail. *Annals of Dyslexia, 46*(1), 1–20. doi:10.1007/BF02648168

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. doi:10.1177/074193258600700104

Graesser, A., Golding, J. M., & Long, D. L. (1996). Narrative presentation and comprehension. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 191–205). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Graves, M. F., & Slater, W. H. (1987, April). *The development of reading vocabularies in rural disadvantaged students, inner-city disadvantaged students, and middle-class suburban students*. Annual Meeting of the American Educational Research Association, Washington, DC.

Graves, M. F., Brunetti, G. J., & Slater, W. H. (1982). The reading vocabularies of primary grade children of varying geographic and social backgrounds. In J. A. Harris & L. A. Harris (Eds.), *New inquiries in reading research and instruction* (pp. 99–104). Rochester, NY: National Reading Conference.

Greenhalgh, K. S., & Strong, C. J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools*, *32*(2), 114–125. doi:10.1044/0161-1461(2001/010)

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language, 24*(2), 123–147. doi:10.1177/0142723704042369

Haden, C. A., Haine, R. A., & Fivush, R. (1997). Developing narrative structure in parent-child reminiscing across the preschool years. *Developmental Psychology*, *32*(2), 295–367. Retrieved from: https://s3.amazonaws.com/academia.edu.documents/31344750/Haden__Haine___Fivush_%281997%29.pdf?response-content-disposition=inline%3B%20filename%3DHaden_C._A._Haine_R._A._and_Fivush_R._19.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F20191101%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20191101T003141Z&X-Amz-Expires=3600&X-Amz-SignedHeaders=host&X-Amz-Signature=9eb06a9cd82b01e7af2f02dac4823310ec478ea1fb2b0e252638468fa5602c6d

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and Dialogic Reading. *Early Childhood Research Quarterly, 15*(1), 75-90. doi:10.1016/S0885-2006(99)00038-1

Harkins, D. A., Koch, P. E., & Michel, G. F. (1994). Listening to maternal story telling affects narrative skill of 5-year-old children. *Journal of Genetic Psychology, 155*(2), 247–257. doi:10.1080/00221325.1994.9914775

Harris, T. L., & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes Publishing.

Hayes, A. (1996). Longitudinal study of family literacy program outcomes. In L. A. Benjamin & J. Lord (Eds.), *Family literacy: Directions in research and implications for practice* (pp. 4554). Washington, DC: U.S. Government Printing Office.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language, 15*(2), 395–410. doi:10.1017/S0305000900012411

Hayward, D. V. (2003). *What's in a question? Measuring narrative comprehension using a picture story in typically developing children aged 4-8* (305254496) [Doctoral dissertation, University of Alberta]. ProQuest Dissertations & Theses Global.

Hayward, D., & Schneider, P. (2000). Effectiveness of teaching story grammar knowledge to pre-school children with language impairment. *Child Language Teaching and Therapy, 16*(3), 255–284. doi:10.1177/026565900001600303

Hayward, D. V., Schneider, P., & Gillam, R. B. (2009). Age and task-related effects on young children's understanding of a complex picture story. *Alberta Journal of Educational Research*, *55*(1), 54–72. Retrieved from: https://jmss.org/index.php/ajer/article/view/55274

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11*(1), 49–76. doi:10.1017/S0047404500009039

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge, UK: Cambridge University Press.

Heath, S. B. (1986). Sociocultural contexts of language development. In H. T. Trueba (Eds.), *Beyond language: Social and cultural factors in schooling language minority students* (pp. 143–186). Sacramento, CA: Bilingual Education Office, California State Department of Education.

Heath, S. B., & Branscombe, A. M. (1986). The book as narrative prop in language acquisition. In B. B. Schieffelin & P. Gilmore (Eds.), *The Acquisition of Literacy: Ethnographic Perspectives* (Vol. 11, pp. 16–34). Norwood, NJ: Ablex.

Hedberg, N. L., & Stoel-Gammon, C. (1986). Narrative analysis: Clinical procedures. *Topics in Language Disorders*, *7*(1), 58–69. doi:10.1097/00011363-198612000-00008

Hedberg, N. L., & Westby, C. E. (1993). *Analyzing storytelling skills: Theory to practice*. Tucson, AZ: Communication Skill Builders.

Hipfner-Boucher, K. (2011). *A comparative study of ELL and EL1 narrative competence during the kindergarten years* (Doctoral dissertation, University of Toronto). Retrieved from: https://tspace.library.utoronto.ca/bitstream/1807/31781/6/Hipfner-Boucher_Kathleen_A_201111_PhD_thesis.pdf

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378. doi:10.1111/1467-8624.00612

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. doi:10.1016/j.dr.2005.11.002

Hoff, E. (2013). *Language development*. Cengage Learning.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing*, *2*(2), 127–160. doi:10.1007/BF00401799

Hudson, J. A., & Shapiro, L. R. (1991). From knowing to telling: Children's scripts, stories, and personal narratives. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure*. Hillsdale, NJ: Erlbaum.

Huebner, C. E. (2000). Promoting toddlers' language development through community-based intervention. *Journal of Applied Developmental Psychology, 21*(5), 513–535. doi:10.1016/S0193-3973(00)00052-6

Huennekens, M. E., & Xu, Y. (2010). Effects of a cross-linguistic storybook intervention on the second language development of two preschool English language learners. *Early Childhood Education Journal*, *38*(1), 19–26. doi:10.1007/s10643-010-0385-1

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.

Hunt, K. W. (1965). *Grammatical structure written at three grade levels*. Champaign, IL: National Council of Teachers of English.

Hutson-Nechkash, P. (2001). *Narrative toolbox: Blueprints for story building*. Thinking Publications.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & et al. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. doi:10.1037/0012-1649.27.2.236

Jackson, J. R., & Dizney, H. (1963). Intensive vocabulary training. *Journal of Developmental Reading*, *6*(4), 221–229. Retrieved from http://www.jstor.org/stable/40015940

Jenkins, J. R., Pany, D., & Scheck, J. (1978). *Vocabulary and reading comprehension: Instructional effects*. Champaign, IL: University of Illinois at Urbana-Champaign, Center for the Reading Study. Retrieved from: https://www.ideals.illinois.edu/bitstream/handle/2142/18005/ctrstreadtechrepv01978i00100_opt.pdf

Johnston, C., Scoular, D. J., & Ohan, J. L. (2004). Mothers' reports of parenting in families of children with symptoms of attention-deficit/hyperactivity disorder: Relations to Impression Management. *Child & Family Behavior Therapy*, *26*(1), 45–61. doi:10.1300/J019v26n01_04

Justice, L., Bowles, R., Kaderavek, J., Ukrainetz, T., Eisenberg, S., Gillam, R. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology,*

15(2), 177–191. doi:10.1044/1058-0360(2006/017)

Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly*, *25*(2), 218–234. doi:10.1016/j.ecresq.2009.11.002

Kame'enui, E. J., Carnine, D. W., & Freschi, R. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly, 17*(3), 367–388. Retrieved from: http://www.jstor.org/stable/747525

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes*, *1*(1), 61–85. doi:10.1080/01690968508402071

Keogh-Brown, M., Bachmann, M., Shepstone, L., Hewitt, C., Howe, A., Ramsay, C. R., Song, F., Miles, J., Torgerson, D., Miles, S., & others. (2007). *Contamination in trials of educational interventions*. Health Technology Assessment Report, Vol. 2, No. 43. Available from: https://aura.abdn.ac.uk/bitstream/handle/2164/184/Keogh.pdf?sequence=1

Klecan-Aker, J. S., & Kelty, K. R. (1990). An investigation of the oral narratives of normal and language-learning disabled children. *Journal of Childhool Communication Disorders*, *13*(2), 207–215. doi:10.1177/152574019001300207

Klecan-Aker, J. S., & Swank, P. R. (1987). *The narrative styles of normal first and third grade*. London, UK: SAGE Publications.

Klop, D. (2011). *The relationship between narrative skills and reading comprehension: when mainstream learners show signs of specific language impairment* [Doctoral dissertation, University of Stellenbosch]. Retrieved from: https://scholar.sun.ac.za/bitstream/handle/10019.1/6755/klop_relationship_2011.pdf?sequence=1

Kong, L.-D. & Fu, M.-Y. (2004). 儿童语言中副词的发展 [Development of adverbs in children language]. *Journal of Anhui Normal University (Hum .& Soc.Sci .) /安徽师范大学学报(人文社会科学版), 32* (5), 591–596. doi:10.14182/j.cnki.j.anu.2004.05.020

Labov, W. (1972). The transformation of experience in narrative syntax. *Language in the Inner City: Studies In The Black English Vernacular* (pp. 354–396). Philadelphia, PA: University of Pennsylvania Press.

Labov, W., Cohen, P., Robins, C., & Lewis, S. (1968). *A study of non-standard English of Negro and Puerto Rican speakers in New York City*. New York, NY: Columbia University.

Labov, W., & Waletzky, J. (1997). Narrative analysis: Oral versions of personal experience. *Journal of Narrative & Life History*, *7*(1), 1–4.

Lai, W.-F. (2013). Socioeconomic differences in Taiwanese children's personal narratives: Conjunctions, internal state terms, and narrative structures. In A. McCabe, & C. Chang (Eds.), *Chinese language narration: culture, cognition, and emotion* (pp. 115–142). Amsterdam, Netherlands: John Benjamins Publishing Company.

Leadholm, B. J., & Miller, J. F. (1994). *Language sample analysis: The Wisconsin guide.* Madison, WI: Wisconsin Department of Public Instruction.

Leung, C. B. (1992). Effects of word-related variables on vocabulary growth repeated read-aloud events. In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views from many perspectives: Forty-first Yearbook of the National Reading Conference* (pp. 491–498). Chicago, IL: The National Reading Conference.

Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a Dialogic Reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology, 108*(1), 1–24. doi:10.1016/j.jecp.2010.07.002

Li, H. (2003). *The efficacy of shared reading instruction and its influencing factors* [分享阅读有效性及其影响因素的研究] [Unpublished master's thesis]. Beijing Normal University, Beijing/北京师范大学，北京.

Li, H. (2008). 整合分享阅读与词素意识教学加速培养幼儿自主阅读能力[Integrating morphological awareness instruction into shared reading to facilitate young children's independent reading ability]. *Early Education/幼儿教育·教育教学*, *9*, 32–33. Retrieved from: http://vpnlib.imu.edu.cn/rwt/6/http/MRYGPLUYMFYGMZLPM7TGC7DBF3SX85JPMNYA/Periodical_yejy200809014.aspx

Li., H., Wu, X., Zhang, J., Zheng, Q, Zhu, J. (2010a). 不同形式分享阅读对儿童字词学习和阅读动机的长期影响 [The long-term effects of different shared book reading programs on children's character learning and reading motivation]. *Journal of Educational Studies/教育学报, 6*(5), 89–95. Retrieved from: http://kns.cnki.net/KCMS/detail/detail.aspx?filename=xkjy201005018&dbname=CJFD&dbcode=CJFQ

Li., H., Wu, X., Zhang, J., Zheng, Q, Zhu, J. (2010b). 不同形式的阅读干预在儿童读写能力发展中的作用 [The role of different reading interventions in the development of reading and writing]. *Studies of Psychology and Behavior/心理行为与研究, 8*(4), 263–267. Retrieved from: http://kns.cnki.net/KCMS/detail/detail.aspx?filename=clxw201004006&dbname=CJFD&dbcode=CJFQ

Lin, P.-Y., & Peng, H.-N. (2014). 提升台灣低社經幼兒語言發展的「對話式閱讀」延伸實驗 [Accelerating language development for Taiwan low SES preschoolers through the Dialogic Reading program extended with different strategies]. *Journal of Early Childhood Education/幼兒教育年刊, 25*, 185–204. Retrieved from: http://ntcuir.ntcu.edu.tw/bitstream/987654321/6489/2/190-209.pdf

Lin, Y.-H., Wu, Y.-Y., & Su, C.-Y. (2005). 對話式閱讀對學前身心障礙兒童詞彙能力之影響 [The effects of Dialogic Reading on vocabulary ability of preschool aged children with disabilities]. *Bulletin of Special Education/特殊教育研究學刊, 29*, 49–72. Retrieved from: http://bse.spe.ntnu.edu.tw/upload/journal/prog/C06_14X0_12OW_0OR1023.pdf

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin*, *139*(1), 1–34. doi:10.1037/a0029321

Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention*, *22*(4), 306–322. doi:10.1177/105381519902200406

Lonigan, C. J., Purpura, D. J., Wilson, S. B., Walker, P. M., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of Experimental Child Psychology*, *114*(1), 111–130. doi:10.1016/j.jecp.2012.08.010

Lonigan, C. A. & Shanahan, T. (2008). Executive summary. In National Center for Family Literacy (Eds.), *Developing early literacy: Report of the national early literacy panel*. National Center for Family Literacy. Retrieved from http://files.eric.ed.gov/fulltext/ED508381.pdf

Lonigan, C.J., Shanahan, T., Cunningham, A. (2008). Impact of shared reading interventions on young children's early literacy skills. In T. Shanahan, & C. A. Lonigan (Eds.), *Developing early literacy: Report of the national early literacy panel* (pp. 153–171). Washington, DC: U.S Department of Education. Retrieved from http://lincs.ed.gov/publications/pdf/NELPReport09.pdf

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, *13*(2), 263–290. doi:10.1016/S0885-2006(99)80038-6

MacWhinney, B. (2000). *The CHILDES project* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*(1), 111–151. doi:10.1016/0010-0285(77)90006-8

Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., & Ginsburg-Block, M. (2010). A descriptive review and meta-analysis of family-based emergent literacy interventions: To what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? *Early Childhood Research Quarterly*, *25*(4), 409–431. doi:10.1016/j.ecresq.2010.03.002

Margosein, C. M., Pascarella, E. T., & Pflaum, S. W. (1982). The effects of instruction using semantic mapping on vocabulary and comprehension. *The Journal Of Early Adolescence*, *2*(2), 185–194. doi:10.1177/027243168200200210

McCabe, A., & Peterson, C. (1991). Getting the story: A longitudinal study of parental styles in eliciting narratives and developing narrative skill. In A. McCabe & C. Peterson (Eds.)*, Developing narrative structure* (p. 217–253). Mahwah, NJ: Lawrence Erlbaum Associates.

McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne effect: A randomised, controlled trial. *BMC Medical Research Methodology*, *7*, 30. doi:10.1186/1471-2288-7-30

McCormick, C., & Mason, J. M. (1984). *Intervention procedures for increasing preschool children's interest in and knowledge about reading*. Retrieved from: https://www.ideals.illinois.edu/bitstream/handle/2142/18043/ctrstreadtechrepv01984i00312_opt.pdf?sequence=1

McMillan, J. H. (2012). *Educational research: Fundamentals for the consumer* (6th ed.). Pearson.

Meng, X., Zhou, X., & Kong, R. (2002). 中文读写能力及其相关因素研究 [A study of reading and writing ability and related factors]. *Science of Psycholog/心理科学, 25*(5), 544–547, 572.

Merritt, D. D., & Liles, B. Z. (1987). Story grammar ability in children with and without language disorder: Story generation, story retelling, and story comprehension. *Journal of Speech, Language, and Hearing Research*, *30*(4), 539–552. doi:10.1044/jshr.3004.539

Moerk, E. L. (1985). Picture-book reading by mothers and young children and its impact upon language development. *Journal of Pragmatics, 9*(4), 547–566. doi:10.1016/0378-2166(85)90021-9

Moerk, E. L., & Moerk, C. (1979). Quotations, imitations, and generalizations: Factual and methodological analyses. *International Journal of Behavioral Development, 2*(1), 43–72. doi:10.1177/016502547900200103

Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*(2), 979–1007. doi:10.3102/0034654309332561

Mol, S., Bus, A., de Jong, M., & Smeets, D. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education & Development, 19*(1), 7–26. doi:10.1080/10409280701838603

Morsbach, S. K., & Prinz, R. J. (2006). Understanding and improving the validity of self-report of parenting. *Clinical Child and Family Psychology Review*, *9*(1), 1–21. doi:10.1007/s10567-006-0001-5

Muñoz, M. L., Gillam, R. B., Peña, E. D., & Gulley-Faehnle, A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools*, *34*(4), 332–342. doi:10.1044/0161-1461(2003/027)

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development:

Evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665–681. doi:10.1037/0012-1649.40.5.665

Nagy, W. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 27–44). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Nagy, W. E. & Scott J. A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Muhwah, NJ: Lawrence Erlbaum Associates. Reprint version.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356. doi:10.1111/j.1467-9817.2004.00238.x

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

National Institute of Child Health and Human Development [NICHD]. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: US Government Printing Office. Retrieved from: https://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Nelson, N. W. &Van Meter, A. M. (2007). Measuring written language ability in narrative samples. *Reading & Writing Quarterly, 23*(3), 287–309. doi:10.1080/10573560701277807

Nickse, R. S. (1991). *A typology of family and intergenerational literacy programs: Implications for evaluation*. Paper presented at the Annual Meeting of the American Educational Research Association. Chicago.

Niklas, F., Cohrssen, C., & Tayler, C. (2016). Home learning environment and concept formation: A family intervention study with kindergarten Children. *Early Childhood Education Journal*, *44*(5), 419–427. doi:10.1007/s10643-015-0726-1

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development, 51*(2), 587–590. doi:10.2307/1129299

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5*(01), 1–15. doi:10.1017/S0305000090000189

Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and young adults*. Austin, TX: PRO-ED.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, *54*(1), 31–39. doi:10.1111/j.2044-8279.1984.tb00842.x

O'Neill, D., Pearce, M., & Pick, J. (2004). Preschool children's narratives and performance on the Peabody Individualized Achievement Test – Revised: Evidence of a relation between early narrative and later mathematical ability. *First Language, 24*(2), 149–183. doi:10.1177/0142723704043529

Pan, B. A. (1994). Basic measures of child language. In *Handbook of research in language development using CHILDES* (pp. 26–49). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*(4), 763–782. doi:10.1111/j.1467-8624.2005.00876.x

Parker, M. D., & Brorson, K. (2005). A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language*, *25*(3), 365–376. doi:10.1177/0142723705059114

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.

Paul, R. (2001). *Language disorders from infancy through adolescence* (2nd ed.). St Louis, MO: Mosby.

Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research*, *36*(3), 592–598. doi:10.1044/jshr.3603.592

Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly*, *9*(3–4), 427–440. doi:10.1016/0885-2006(94)90018-3

Pellegrini, A. D., Brody, G. H., & Sigel, I. E. (1985). Parents' teaching strategies with their children: The effects of parental and child status variables. *Journal of Psycholinguistic Research, 14*(6), 509–521. doi:10.1007/BF01067382

Pellegrini, A. D., & Galda, L. (1993). Ten years after: A reexamination of symbolic play and literacy research. *Reading Research Quarterly, 28*(2), 163–175. Retrieved from: http://www.jstor.org/stable/747887

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*(1), 23–33. doi:10.1037/0022-0663.94.1.23

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. doi:10.1080/10888430701530730

Perfetti, C., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. In J. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 3-20). Polymouth, UK: Rowman & Littlefield Education.

Peterson, C., & Biggs, M. (2001). “I was really, really, really mad!” Children’s use of evaluative devices in narratives about emotional events. *Sex Roles*, *45*(11–12), 801–825. doi:10.1023/A:1015692403932

Peterson, C., Jesso, B., & McCabe, A. (1999). Encouraging narratives in preschoolers: An intervention study. *Journal of Child Language, 26*(1), 49–67. doi:10.1017/S0305000998003651

Peterson C., & McCabe A. (1983). *Developmental psycholinguistics: three ways of looking at a child's narrative*. New York, NY: Plenum Press.

Peterson, C., & McCabe, A. (1992). Parental styles of narrative elicitation: Effect on children's narrative structure and content. *First Language, 12*(36), 299–321. doi:10.1177/014272379201203606

Peterson, C., & McCabe, A. (1994). A social interactionist account of developing decontextualized narrative skill. *Developmental Psychology, 30*(6), 937–948. doi:10.1037/0012-1649.30.6.937

Peterson, C., & McCabe, A. (2004). Echoing our parents: Parental influences on children's narration. In M. W. Pratt & B. H. Fiese (Eds.), *Family stories and the life course: Across time and generations* (pp. 27–54). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Pierre R. G., Ricciuti A. E., & Tao F. (2004). Continuous improvement in family literacy programs. In B. Wasik (Eds.), *Handbook of family literacy* (Vol. 2, pp. 587–600). Mahwah, NJ: Lawrence Erlbaum Associates.

Pillinger, C., & Wood, C. (2014). Pilot study evaluating the impact of dialogic reading and shared reading at transition to primary school: Early literacy skills and parental attitudes: Dialogic reading and early literacy. *Literacy*, *48*(3), 155–163. doi:10.1111/lit.12018

Powell, D., & Carey, A. (2012) Approaches to program fidelity in family literacy research. In B. Wasik (Eds.), *Handbook of family literacy* (Vol. 2, pp. 87–400). Mahwah, NJ: Lawrence Erlbaum Associates.

Propp, V. (1958). *Morphology of the folktale*. Bloomington, Indiana: Indiana University Research Center in Anthropology.

Purcell-Gates, V. (2000). Family literacy. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 853–870). Mahwah, NJ: Lawrence Erlbaum Associates.

Rahn, N. L., Coogle, C. G., & Storie, S. (2016). Preschool children's use of thematic vocabulary during Dialogic Reading and activity-based intervention. *The Journal of Special Education*, *50*(2), 98–108. doi:10.1177/0022466915622202

Ranganathan, P., Pramesh, C. S., & Aggarwal, R. (2017). Common pitfalls in statistical analysis: Measures of agreement. *Perspectives in Clinical Research*, *8*(4), 187–191. doi:10.4103/picr.PICR_123_17

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. SAGE.

Reese, E., Leyva, D., Sparks, A., & Grolnick, W. (2010). Maternal elaborative reminiscing increases low-income children's narrative skills relative to Dialogic Reading. *Early Education and Development, 21*(3), 318–342. doi:10.1080/10409289.2010.481552

Richardson, & L. (1990). Narrative and sociology. *Journal of Contemporary Ethnography*, *19*(1), 116–135. doi:10.1177/089124190019001006

Roskos, K., & Christie, J. (2011). The play-literacy nexus and the importance of evidence-based techniques in the classroom. *American Journal of Play*, *4*(2), 204–224. Retrieved from: https://files.eric.ed.gov/fulltext/EJ985588.pdf

Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research*, *95*(5), 259–272. doi:10.1080/00220670209596600

Roth, F. P., & Spekman, N. J. (1986). Narrative discourse: Spontaneously generated stories of learning-disabled and normally achieving students. *Journal of Speech and Hearing Disorders*, *51*(1), 8–23. doi:10.1044/jshd.5101.08

Roth, F. P., Spekman, N. J., & Fye, E. C. (1995). Reference cohesion in the oral narratives of students with learning disabilities and normally achieving students. *Learning Disability Quarterly*, *18*(1), 25–40. doi:10.2307/1511363

Rudd, L. C., & Kelley, H. M. (2011). Language development. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (p. 865). Springer US. doi:10.1007/978-0-387-79061-9_1609

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding* (pp. 211–236). doi:10.1016/B978-0-12-108550-6.50013-6

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, *14*(3), 245–302. doi:10.1006/drev.1994.1010

Scherer, N. J., & Olswang, L. B. (1984). Role of mothers' expansions in stimulating children's language production. *Journal of Speech, Language, and Hearing Research, 27*(3), 387–396. doi:10.1044/jshr.2703.387

Schmitt, N. (2000). *Vocabulary in language teaching*. UK: Cambridge University Press.

Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton narrative norms instrument*. University of Alberta Faculty of Rehabilitation Medicine website. Retrieved from: http://www.rehabresearch.ualberta.ca/enni

Schwanenflugel, P. J., Stahl, S. A., & Mcfalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research, 29*(4), 531–553. doi:10.1080/10862969709547973

Seefeldt, C. (2004). *Family literacy survey*. Baltimore, MD: Johns Hopkins University Center for the Social Organization of Schools.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24*(1), 123–138. doi:10.1017/S0305000996003005

Sénéchal, M. (2011). A model of concurrent and longitudinal relations between home literacy and child outcomes. In S. B. Neuman and D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol.3, pp.175–188). New York, NY: Guilford.

Sénéchal, M., & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, *73*(2), 445–460. doi:10.1111/1467-8624.00417

Sénéchal, M., LeFevre, J., Hudson, E., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology, 88*(3), 520–536. doi:10.1037/0022-0663.88.3.520

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. *Handbook of early literacy research* (Vol.2, pp.173–182). New York, NY: Guilford Press.

Shapiro, L. R. & Hudson, J. A. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental Psychology, 27*(6), 960–974. Retrieved from: https://www.researchgate.net/profile/Judith_Hudson6/publication/232418074_Tell_Me_a_Make-Believe_Story_Coherence_and_Cohesion_in_Young_Children's_Picture-Elicited_Narratives/links/56ba05d708ae7e3a0fa09f4a.pdf

Shiu, S.-P., & Tsai, Y.-C. (2016). 對話式閱讀對於弱勢幼兒的語言能力之影響 [The Effects of a Dialogic Reading intervention on the language ability of disadvantaged children]. *Educational Journal of NHCUE/新竹教育大學教育學報, 33*(2), 61–93. Retrieved from: http://edujou.web.nthu.edu.tw/ezfiles/989/1989/img/1172/33-2-3.pdf

Shu, H., Li, W., Anderson, R., Ku, Y.-M., & Yue, X. (2002a). The role of home-literacy environment in learning to read Chinese. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 207–224). Boston: Kluwer.

Shu, H., Li, W., Gu, Y., Anderson, R., Wu, X., Zhang, H., & Xuan Y. (2002b). 家庭文化背景儿童阅读发展中的作用 [Role of family cultural background in children's reading development]. *Science of Psychology/心理科学, 25*(2), 136–139. Retrieved from: http://www.cqvip.com/qk/95682a/200202/12303354.html

Sim, S. S. H., Berthelsen, D., Walker, S., Nicholson, J. M., & Fielding-Barnsley, R. (2014). A shared reading intervention with parents to enhance young children's early literacy skills. *Early Child Development and Care*, *184*(11), 1531–1549. doi:10.1080/03004430.2013.862532

Singer, H. (1964). Substrata-factor patterns accompanying development in power of reading, elementary through college level. In E. Thurston and L. Hafner (Eds.), *Fourteenth Yearbook of the National Conference on Reading*, (pp. 41–56). Milwaukee, WI: National Conference on Reading.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Snow, C. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53*(2), 165–189. Retrieved from https://www.researchgate.net/profile/Catherine_Snow/publication/232581419_Literacy_and_Language_Relationships_During_the_Preschool_Years/links/54807b120cf2ccc7f8bd0664.pdf

Snow, C. E. (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education*, *6*(1), 5–10. doi:10.1080/02568549109594817

Snow, C. E., & Dickinson, D. K. (1991). Skills that aren't basic in a new conception of literacy. In *Literate systems and individual lives: Perspectives on literacy and schooling* (pp. 179–191). Albany, NY: State of New York Press.

Snow, C. E., & Goldfield, B. A. (1983). Turn the page please: Situation-specific language acquisition. *Journal of Child Language, 10*(3), 551–569. doi:10.1017/S0305000900005365

Snow, C. E., Tabors, P. O., Nicholson, P. A., & Kurland, B. F. (1995). SHELL: Oral language and early literacy skills in kindergarten and first-grade children. *Journal of Research in Childhood Education*, *10*(1), 37–48. doi:10.1080/02568549509594686

Snyder, L. S., & Downey, D. M. (1991). The language-reading relationship in normal and reading-disabled children. *Journal of Speech, Language, and Hearing Research*, *34*(1), 129–140. doi:10.1044/jshr.3401.129

Soifer, L. H. (2011). Development of oral language and its relationship to literacy. In J. R. Birsh. (Eds.), *Multisensory teaching of basic language skills* (3rd ed., pp. 49–91). Baltimore, MD: Brookes.

Sperry, L. L., & Sperry, D. E. (1996). Early development of narrative skills. *Cognitive development, 11*(3), 443–465. doi:10.1016/S0885-2014(96)90013-1

Stein, N. L. (1978). *How children understand stories: A developmental analysis*. Center for the Study of Reading, University of Illinois at Urbana-Champaign. Retrieved from: https://www.ideals.illinois.edu/bitstream/handle/2142/17672/ctrstreadtechrepv01978i00069_opt.pdf?sequence=1

Stein, N. L., & Glenn, C. G. (1975). *An analysis of story comprehension in elementary school children: A test of a schema*. Retrieved from: https://eric.ed.gov/?id=ED121474

Stein, N., & Albro, E. R. (1997). Building complexity and coherence: Children's use of goal structured knowledge in telling stories. In M. G. W. Bamberg (Eds.), *Narrative*

Development: Six Approaches (pp. 5–44). Mahwah, NJ: Lawrence Erlbaum Associates Publishers

Stein, N., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. *New Directions in Discourse Processing*, *2*, pp. 53–120. Retrieved from: https://www.researchgate.net/publication/243501171_An_Analysis_of_Story_Comprehension_in_Elementary_School_Children

Stephens, I. M. (1988). Pragmatics. In M. A. Nippold (Ed.), *Later language development: The school-age and adolescent years* (pp. 247–262). Boston, MA: Little, Brown, and Company.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934–947. doi:10.1037//0012-1649.38.6.934

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology*, *49*(12), 2368–2382. doi:10.1037/a0032463

Thomas, A., & Fisher, B. (1996). *Assessment and evaluation strategies in family literacy program development*. Guide prepared for National Literacy Secretariat of Canada and the Ontario Training and Adjustment Board. Retrieved from:

http://en.copian.ca/library/learning/aestrat/aestrat.pdf

Torgerson, D. J. (2001). Contamination in trials: Is cluster randomisation the answer? *BMJ*, *322*(7282), 355–357. doi:10.1136/bmj.322.7282.355

Towson, J. A., & Gallagher, P. A. (2014). Training Head Start parents in dialogic reading to improve outcomes for children. *International Journal of Child Health and Human Development*, *7*(3), 287–296. Retrieved from: https://search.proquest.com/openview/c902fb719a5b714f63b5faedeb1cb31f/1?pq-origsite=gscholar&cbl=2034871

Towson, J. A., Fettig, A., Fleury, V. P., & Abarca, D. L. (2017). Dialogic Reading in early childhood settings: A summary of the evidence base. *Topics in Early Childhood Special Education*, *37*(3), 132–146. doi:10.1177/0271121417724875

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15*(3), 249–275. doi:10.1080/01638539209544812

Tuinman, J. J., & Brady, M. E. (1974). How does vocabulary account for variance on reading comprehension tests? A preliminary instructional analysis. In *Twenty-third yearbook of the National Reading Conference* (pp. 176–184). Clemson, SC: National Reading Conference.

Uccelli, P., Hemphill, L., Pan, B.A., & Snow, C. (1999). Telling two kinds of stories: Sources of narrative skill. In L. Balter & C.S. Tamis-Lemonde (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 215–233). Philadelphia, PA: University Press.

Uccelli, P., & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*, *38*(3), 225–236. doi:10.1044/0161-1461(2007/024)

Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology, 28*(6), 1106–1114. doi:10.1037/0012-1649.28.6.1106

Vernon-Feagans, L., Hammer, C. S., Miccio, A. D. E. L. E., & Manlove, E. L. I. Z. A. B. E. T. H. (2001). Early language and literacy skills in low-income African American and Hispanic children. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol.1, pp.192–210). New York, NY: Guilford Press.

Vinco, M. H. (2013). *Assessment of preschool vocabulary: Expressive and receptive knowledge of word meanings* (Doctoral dissertation, The Florida State University). Retrieved from: http://diginole.lib.fsu.edu/islandora/object/fsu:185179/datastream/PDF/view

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

Wallach, G. P., & Butler, K. G. (1994). Creating communication, literacy, and academic success. In G. P. Wallach & K. G. Butler (Eds.), *Language learning disabilities in school-age children and adolescents* (pp. 2–26). New York: Macmillan.

Wang, J. (2018). 汉语儿童叙事能力发展研究 [Research on the development of Mandarin-speaking children's narrative competence]. Beijing: Tsinghua University Press/北京：清华大学出版社.

Wang, X., & Hu, Y. (2007). 小学生家庭阅读情况的调查研究 [A survey of elementary students' reading at home]. *Journal of Inner Mongolia Normal University—Education Issue/内蒙古师范大学学报(教育科学版), 20*(6), 84–87.

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of educational psychology, 93*(2), 243–250. Retrieved from http://celi.olemiss.edu/wp-content/uploads/sites/7/2014/01/wasik-and-bond-article.pdf

Wasik, B. H. (2004). Assessment and evaluation. In B. H. Wasik (Eds.) *Handbook of family literacy* (Vol. 2, pp.501–502). Mahwah, NJ: Lawrence Erlbaum Associates.

Wechsler, D. (1967). *Wechsler preschool and primary scales of intelligence*. New York, NY: Psychological Corporation.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265–279. doi:10.1037/0012-1649.37.2.265

Wells, G. (1985a). *Language development in the preschool years*. New York, NY: Cambridge University Press.

Wells, G. (1985b). Preschool literacy-related activities and success in school. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language, and learning: The nature and consequences of reading and writing* (pp. 229–255). New York, NY: Cambridge University Press.

Wells, G., & Gordon, W. (1985). *Language development in the pre-school years* (Vol. 2). New York, NY: Cambridge University Press.

Westby, C. E. (1982). Cognitive and linguistic aspects of children's narrative development. In L. Bradford (Ed.), *Communication disorders* (Vol. 7, pp. 1–16). New York, NY: Grune and Stratton.

Westby, C. E. (1984). Development of narrative language abilities. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-aged children* (pp. 103–107). Baltimore, MD: Williams and Wilkins.

Westby, C., Van Dongen, R., & Maggart, Z. (1989). Assessing narrative competence. S*eminars in Speech and Language*, *10*(1), 63–76. doi:10.1055/s-0028-1082490

Wheeler, P. (1983). Context-related age characteristics in mothers' speech: Joint book reading. *Journal of Child Language, 10*(1), 259–263. doi:10.1017/S0305000900005304

Whitehurst, G. J. (1992). *Stony Brook family reading survey*. Stony Brook, NY: Author.

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, *30*(5), 679–689.

doi:10.1037/0012-1649.30.5.679

Whitehurst, G. J., Epstein, J. N., Angel, A. L., Payne, A. C., Crone, D. A. & Fischel J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology*, *86*(4), 542–555. doi:10.1037/0022-0663.86.4.542

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24*(4), 552–559. doi:10.1037/0012-1649.24.4.552

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, *69*(3), 848–872. doi:10.1111/j.1467-8624.1998.tb06247.x

Widdowson, H. G. (1989). Knowledge of language and ability for use. *Applied linguistics, 10*(2), 128–137. doi:10.1093/applin/10.2.128

Williams, K. T. (2008). *Expressive vocabulary test-2nd edition*. Pearson.

Wu, X., Li, H., Anderson, R. C., & Li, W. (2002). 拼音在儿童分享阅读中的作用 [The role of pinyin in shared reading instruction]. *Science of Psychology/心理科学, 25*(5), 548–551. Retrieved from: http://www.dangzheng.cnki.net/kcms/detail/detail.aspx?filename=XLKX200205010&dbcode=CFJD&dbname=

Wu, Y. (2014). *互动式分享阅读对 4-6 岁幼儿阅读兴趣, 叙事能力的影响 [The effect of shared reading on the reading motivation and narrative competence of 4-6 years old children]* (Doctoral dissertation, Shanghai Normal University). Retrieved from: http://d.wanfangdata.com.cn/Thesis/Y2527013

Zeng, W.-X. & Li J. (2006). 儿童叙事能力发展的促进与干预研究(综述) [Intervention on children's narrative competence (a review)]. *Chinese mental health journal/中国心理卫生杂志*, *20*(9), 572–575. doi:10.3321/j.issn:1000-6729.2006.09.004

Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic Reading: A shared picture book reading intervention for preschoolers. In P. H. Anne Van Kleeck, A. Van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (177–200). Mahwah, NJ: Lawrence Erlbaum Associates.

Zevenbergen, A., Whitehurst, G., & Zevenebergen, J. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Applied Development Psychology, 24*(1), 1–15. doi:10.1016/S0193-3973(03)00021-2

Zhao, L. (2003). Relations between children's early language education and their Chinese literacy development—A study on early family education [儿童早期语言教育与其后继语文能力发展关系的研究报告——一项早期家庭教育的追溯研究]. *Preschool Education Research /学前教育研究, 11*, 15–18. doi:10.13861/j.cnki.sece.2003.11.004

Zhou, J. & Zhang, L.(2009). 汉语儿童平均语句长度发展研究 [Research on the development of Chinese children's mean length of utterance] In J. Zhou & J. R. Chang (Eds.). 汉语儿童语言发展研究:国际儿童语料库研究方法的应用与发展 (pp. 40–58). Beijing: Educational Science Press/北京: 教育科学出版社.

Zhu, J. (2005). 使用分享阅读有效促进汉语幼儿语言发展的实验研究 [*Using shared reading approach to promote children's language development*] (Unpublished master's thesis). Beijing Normal University, Beijing.

Zhu, M. & Li, J. (2015). 儿童口语叙事中叙事评价的发展研究 [Development of evaluation devices in oral narratives of children]. *Chinese Journal of Special Education/中国特殊教育*, (08), 92–96. Retrieved from: http://www.cnki.com.cn/Article/CJFDTotal-ZDTJ201508016.htm

Appendices

Appendix A:

Family-Information Questionnaire

Instructions

This questionnaire includes questions about you and your family. Your answers will be treated confidentially. Please fill it out, put in the envelope provided and seal it, and return it to your child's kindergarten teacher, who will pass it on to me.

Your Child

Please answer the following questions about your child.

1. Child's name_____ **2. Gender** ☐a. Female ☐b. Male

3. Date of birth_____ **4. Child's school**_____

5. Kindergarten class level ☐a. Junior ☐b. Middle ☐c. Senior

Your Family

Please answer the following questions about your family.

6. Contact phone number: _____

7. Mailing address: _____

8. WeChat number: _____

9. Language spoken at home:

☐a. Mandarin ☐b. language of minorities ☐c. other

10. Annual income (please check one of the following):

☐a. Below 30,000 RMB ☐b. 30,000~80,000 RMB

☐c. 80,000~120,000 RMB ☐d. 120,000~300,000 RMB

☐e. 300,000~1,000,000 RMB ☐f. Above 1,000,000 RMB

11. Mother's name: _____

12. Mother's occupation: _____

13. Mother's highest level of education. (*Please check one of the following.*)

☐a. Completed junior high school
☐b. Completed secondary/high school
☐c. Completed technical college
☐d. Completed undergraduate university degree
☐e. Completed graduate/advanced university degree

14. Father's name: _____

15. Father's occupation: _____

16. Father's (or other adult in the home) highest level of education. (*Please check one of the following.*)

☐a. Completed junior high school
☐b. Completed secondary/high school
☐c. Completed technical college
☐d. Completed undergraduate university degree
☐e. Completed graduate/advanced university degree

17. What do you hope to gain from this program?

_____

_____

18. If your child has attended or currently attends other programs related to reading, writing, or language development please describe below:

Name of the program: _____

Purpose of the program: _____

Number of hours per week: _____

Home Literacy Environment

Please answer the following questions about your current activities at home.

19. How often do you or another family member read a picture book with your child?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

20. How often does your child ask to be read to?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

21. How often does your child look at books by himself or herself?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

22. How often do you go to the library (picture book libraries, public libraries, or library in your community/kindergarten/other institutions) with your child?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

23. How often do you go to a bookstore with your child?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

24. How many minutes did you or another family member read to your child yesterday?

☐a. 0 minutes ☐b. 1–10 minutes ☐c. 11–20 minutes
☐d. 20–30 minutes ☐e. More than 30 minutes

25. Approximately how many picture books do you have in your home for your child's use?

☐a. 0–2 ☐b. 3–8 ☐c. 9–15 ☐d. 16–25 ☐e. 26-40 ☐f. 41-60
☐g. More than 60

26. At what age did you or another family member begin to read to your child?

☐a. 0–6 months ☐b. 7–12 months ☐c. 13–18 months
☐d. 19 months – 2 years ☐e. After their second birthday

27. How many minutes per day do you spend reading adult books, magazines, etc.?

☐a. None ☐b. 1–15 minutes ☐c. 16–30 minutes ☐d. 31–60 minutes
☐e. more than an hour

28. Please rate your enjoyment in reading adult books, magazines, etc. (electronic or paper version) on a scale of 1 to 10, where 1 means "do not enjoy at all" and 10 means "enjoy immensely".

Do not enjoy at all Enjoy immensely

☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

29. How often do you listen to your child and encourage your child to talk?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

30. How often do you show your child Chinese characters (in books/magazines/ instruction books, on television/on wall posters, or other places in the home)?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

31. How often does your child pretend to read familiar books?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3-5 times a week ☐e. Almost daily

32. During parent-child reading, how often does your child show the following behaviours <u>without being prompted to</u>? **My child…**	a. Never	b. Rarely	c. Occasionally	d. Frequently	e. Always
32-1. …talks about the story					
32-2. …raises questions with or without giving her/his own answer					
32-3. …retells the story or part of it in the picture book					
32-4. …creates his/her own stories with the aid of the pictures in the book					

33. Please rate your enjoyment in reading with your child on a scale of 1 to 10, where 1 means "do not enjoy at all" and 10 means "enjoy immensely".

Do not enjoy at all → Enjoy immensely

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

34. Please rate your confidence in reading with your child on a scale of 1 to 10, where 1 means "not confident at all" and 10 means "extremely confident".

Not confident at all → Extremely confident

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

35. During parent-child reading, how often do you usually demonstrate the following behaviours with your child? I…	a. Never	b. Rarely	c. Occasionally	d. Frequently	e. Always
35-1. …talk about the story					
35-2. …ask my child questions					
35-3. …ask my child to retell the story in the picture book or part of it					
35-4. …ask my child to create his/her own stories with the aid of the pictures in the book					

36. To what extent do you agree with the following statements

36-1. When you read storybooks with a child, you should pause and talk about the story and pictures as you go along.

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

36-2. Children should learn to listen quietly and not interrupt when a story is being read aloud.

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

36-3. It is good to ask children increasingly challenging questions about the book as you read.

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

36-4. It is good for parents to control the reading process (e.g., turning pages, deciding on the pace of reading, holding the book, etc.).

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

Appendix B

Follow-Up Questionnaire

Directions

This questionnaire includes questions about your feelings about your participation in the reading program. Your answers will be treated confidentially. Please fill it out, put in the envelope provided and seal it, and return it to the kindergarten teacher, who will pass it on to me.

1. How often does your child ask to be read to?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

2. How often does your child look at books by himself or herself?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

3. How often does your child pretend to read familiar books?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

4. How often do you go to the library (picture book libraries, public libraries, or the library in your community/kindergarten/other institution) with your child?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3-5 times a week ☐e. Almost daily

5. How often do you go to the bookstore with your child?

☐a. Rarely ☐b. Once or twice a month ☐c. Once or twice a week
☐d. 3–5 times a week ☐e. Almost daily

6. During parent-child reading, how often does your child show the following behaviours <u>without being prompted to</u>? **My child…**	a. Never	b. Rarely	c. Occasionally	d. Frequently	e. Always
6-1 …talks about the story					
6-2 …raises questions with or without giving her/his own answer					
6-3 …retells the story or part of it in the picture book					
6-4 …creates his/her own stories with the aid of the pictures in the book					

7. Please rate your enjoyment in reading with your child on a scale of 1 to 10, where 1 means "do not enjoy at all" and 10 means "extremely enjoy intensely".

Do not enjoy at all → Enjoy intensely

☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

8. Please rate your confidence in reading with your child on a scale of 1 to 10, where 1 means "not confident at all" and 10 means "extremely confident".

Not confident at all → Extremely confident

☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

9. During parent-child reading, how often do you usually demonstrate the following behaviours with your child? I…	a. Never	b. Rarely	c. Occasionally	d. Frequently	e. Always
9-1. …talk about the story					
9-2. …ask my child questions					
9-3. …ask my child to retell the story in the picture book or part of it					
9-4. …ask my child to create his/her own stories with the aid of the pictures in the book					

10. To what extent do you agree with the following statements:

10-1. When you read storybooks with a child, you should pause and talk about the story and pictures as you go along.

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

10-2. Children should learn to listen quietly and not interrupt when a story is being read aloud.

☐a. Strongly disagree ☐b. Disagree ☐ c. Neutral ☐d. Agree ☐e. Strongly agree

10-3. It is good to ask children increasingly challenging questions about the book as you read.

☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

10-4. It is good for parents to control the reading process (e.g., turning pages, setting the pace of reading, holding the book, etc.).
☐a. Strongly disagree ☐b. Disagree ☐c. Neutral ☐d. Agree ☐e. Strongly agree

11. Do you think there have been any benefits to you from participating in this program? Please rate the benefits of the following aspects on a scale of 1 to 5, where 1 means "not beneficial at all" and 5 means "extremely beneficial".

11-1. Benefits to your parent-child reading techniques:
☐a. Not beneficial at all ☐b. A little bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

11-2. Benefits to the frequency of your parent-child reading:
☐a. Not beneficial at all ☐b. A little bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

11-3. Benefits to your interest in parent-child reading:
☐a. Not beneficial at all ☐b. A little bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

11-4. Did you benefit in other ways?

_____

_____

12. Do you think there have been any benefits to your child from participating in this program? Please rate the benefits to the following aspects on a scale of 1 to 5, where 1 means "not beneficial at all" and 5 means "extremely beneficial".

12-1. Benefits to your child's overall language development:
☐a. Not beneficial at all ☐b. A bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

12-2. Benefits to your child's reading behaviour:
☐a. Not beneficial at all ☐b. A bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

12-3. Benefits to your child's frequency of reading by himself/herself:
☐a. Not beneficial at all ☐b. A bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

12-4. Benefits to your child's interest in reading:
☐a. Not beneficial at all ☐b. A bit beneficial ☐c. Moderately beneficial
☐d. Very beneficial ☐e. Extremely beneficial

12-5. Did your child benefit in other ways?

_____

_____

13. Has the way you think about parent-child reading changed because of the program?

_____

_____

14. How do you feel about the program? Please rate the helpfulness of the following elements on a scale of 1 to 5, where 1 means "not helpful at all" and 5 means "extremely helpful".

14-1. Helpfulness of the picture books provided by the program:
☐a. Not helpful at all ☐b. A bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-2. Helpfulness of the list of suggested questions for picture books
☐a. Not helpful at all ☐b. A l bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-3. Helpfulness of the handout (summary of workshop content)
☐a. Not helpful at all ☐b. A lbit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-4. Helpfulness of facilitator explanations during the workshops
☐a. Not helpful at all ☐b. A bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-5. Helpfulness of the video modelling during the workshops
☐a. Not helpful at all ☐b. A bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-6. Helpfulness of the practice activities you did during the workshop
☐a. Not helpful at all ☐b. A bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-7. Helpfulness of the tips for reading enjoyment
☐a. Not helpful at all ☐b. A bit helpful ☐c. Moderately helpful
☐d. Very helpful ☐e. Extremely helpful

14-8. Were other aspects of the program helpful?

15. What did you like the most about the program?

16. Would you change anything about the program?

17. Were there things that might have prevented you from being able to attend the program? If so, what kind of barriers?

18. Are there things that make it difficult at times in practise the DR approach with your child at home?

Appendix C

List of Selected Books

Title	Author	Publisher
《大脚丫游巴黎》 English Title: *Belinda in Paris*	埃米·扬 English name: Amy Young	启发童书馆 Translation: Inspiration Children Book House
《大脚丫跳芭蕾》 English Title: *Belinda, the Ballerina*	埃米·扬 English name: Amy Young	启发童书馆 Translation: Inspiration Children Book House
《你好，安东医生》 Translated title: *Hello, Dr. Anthony*	西村敏雄 English name: Toshio Nishimura	连环画出版社 Translation: Sinocomic Publishing House
《你好，安东医生--出诊记》 Translated title: *Hello, Dr. Anthony -- House Calls*	西村敏雄 English name: Toshio Nishimura	连环画出版社 Translation: Sinocomic Publishing House
《你好，理发店》 Translated title: *Hello, Barber*	乾荣里子 著 西村敏雄 绘 Author: Eriko Inui Illustrator: Toshio Nishimura	北京科学技术出版社 Translation: Beijing Science and Technology Press
《放屁的公主》 French title: *La princesse qui pète*	莫德·罗杰斯 Maud Roegiers	河南文艺出版社 Translation: He'nan Literature and Art Press

Appendix D

GLM Tables in the Preliminary Analyses

Table D1

The Relationship between Pretest Expressive Vocabulary and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 506.019		$R^2 = 0.344$	
Gender*	1.505	3.269	0.460	0.647	0.2118
Age	1.053	0.199	5.303	< 0.001	28.1185
Income	-0.439	1.689	-0.260	0.796	0.0677
Medu	0.369	2.035	0.181	0.857	0.0329
Pedu	0.159	1.507	0.105	0.916	0.0111
Model with HLE factors		AIC = 519.328		$R^2 = 0.167$	
HLERes	11.0	6.46	1.696	0.095	2.876
ReadAct	-10.0	10.80	-0.927	0.358	0.859
OthLiterAct	14.7	7.87	1.865	0.067	3.478
EmoMoti	18.4	11.92	1.543	0.128	2.380

Note. *=Male compared with female, Medu=Maternal education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D2

The Relationship between Pretest Story Grammar and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 339.433		$R^2 = 0.281$	
Gender*	-1.172	0.8897	-1.318	0.193	1.736
Age	0.225	0.0541	4.158	< 0.001	17.291
Income	0.494	0.4597	1.076	0.287	1.157
Medu	0.327	0.5537	0.591	0.557	0.350
Pedu	-0.467	0.4102	-1.139	0.260	1.296
Model with HLE factors		AIC = 352.8660		$R^2 = 0.0853$	
HLERes	3.1482	1.760	1.7884	0.079	3.20
ReadAct	-0.0706	2.941	-0.0240	0.981	5.77e-4
OthLiterAct	2.1453	2.145	1.0001	0.321	1.00
EmoMoti	-3.2743	3.248	-1.0081	0.318	1.02

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D3

The Relationship between Pretest TNW and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 57.243		$R^2 = 0.143$	
Gender*	-0.23371	0.09813	-2.382	0.021	5.6717
Age	0.00707	0.00596	1.186	0.240	1.4064
Income	-0.03119	0.05070	-0.615	0.541	0.3784
Medu	-0.01687	0.06107	-0.276	0.783	0.0763
Pedu	-0.01800	0.04524	-0.398	0.692	0.1583
Model with HLE factors		AIC = 61.8030		$R^2 = 0.0505$	
HLERes	0.156	0.1812	0.859	0.394	0.737
ReadAct	0.186	0.3026	0.613	0.542	0.376
OthLiterAct	0.122	0.2207	0.552	0.583	0.305
EmoMoti	-0.423	0.3343	-1.267	0.210	1.604

Note. TNW=Total number of words, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D4

The Relationship between Pretest Vocd and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 448.6690		$R^2 = 0.0765$	
Gender*	2.5327	2.227	1.137	0.260	1.2937
Age	0.0814	0.134	0.607	0.546	0.3682
Income	0.1809	1.144	0.158	0.875	0.0250
Medu	-1.1879	1.438	-0.826	0.412	0.6824
Pedu	-0.7433	1.042	-0.713	0.479	0.5087
Model with HLE factors		AIC = 442.102		$R^2 = 0.141$	
HLERes	7.16	3.768	1.901	0.062	3.612
ReadAct	3.56	6.236	0.571	0.571	0.326
OthLiterAct	5.56	4.546	1.222	0.227	1.494
EmoMoti	-13.38	6.934	-1.929	0.059	3.721

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D5

The Relationship between Pretest Verb Density and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = -250.3070		R^2 = 0.1570	
Gender*	0.01410	0.00888	1.588	0.118	2.522
Age	7.23e-4	5.39e-4	1.341	0.185	1.797
Income	0.00312	0.00459	0.681	0.499	0.463
Medu	-0.01507	0.00552	-2.727	0.008	7.439
Pedu	0.00576	0.00409	1.408	0.165	1.981
Model with HLE factors		AIC = -242.6410		R^2 = 0.0196	
HLERes	-0.00971	0.01679	-0.578	0.565	0.3344
ReadAct	0.01452	0.02805	0.517	0.607	0.2678
OthLiterAct	-0.01804	0.02046	-0.881	0.382	0.7769
EmoMoti	-0.00400	0.03098	-0.129	0.898	0.0167

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D6

The Relationship between Pretest Adverb Density and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = -236.8540		R^2 = 0.0297	
Gender*	0.00365	0.00986	0.3696	0.713	0.13664
Age	-2.84e-5	5.99e-4	-0.0474	0.962	0.00225
Income	-0.00271	0.00510	-0.5313	0.597	0.28229
Medu	0.00106	0.00614	0.1725	0.864	0.02977
Pedu	-0.00362	0.00455	-0.7961	0.429	0.63376
Model with HLE factors		AIC = -240.6110		R^2 = 0.0560	
HLERes	3.93e-4	0.01706	0.0231	0.982	5.31e-4
ReadAct	-0.02097	0.02850	-0.7357	0.465	0.5413
OthLiterAct	-0.02234	0.02079	-1.0749	0.287	1.1554
EmoMoti	-0.00817	0.03148	-0.2597	0.796	0.0674

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D7

The Relationship between Pretest TNT and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = -369.89200		R^2 = 0.12754	
Gender*	0.00780	0.00349	2.237	0.029	5.003
Age	1.69e-4	2.12e-4	0.798	0.428	0.637
Income	0.00127	0.00180	0.703	0.485	0.494
Medu	0.00207	0.00217	0.955	0.344	0.911
Pedu	-5.40e-4	0.00161	-0.336	0.738	0.113
Model with HLE factors		AIC = -366.2860		R^2 = 0.0477	
HLERes	-2.60e-4	0.00639	-0.0406	0.968	0.00165
ReadAct	-0.00934	0.01068	-0.8747	0.385	0.76514
OthLiterAct	-9.21e-5	0.00779	-0.0118	0.991	1.40e-4
EmoMoti	0.01930	0.01179	1.6368	0.107	2.67914

Note. TNT=Total number of T-units, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D8

The Relationship between Pretest MLU5 and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 293.236		R^2 = 0.253	
Gender*	-1.4505	0.6202	-2.339	0.023	5.4698
Age	0.1162	0.0377	3.082	0.003	9.4995
Income	-0.0885	0.3204	-0.276	0.783	0.0763
Medu	-0.1199	0.3859	-0.311	0.757	0.0965
Pedu	-0.2309	0.2859	-0.808	0.423	0.6521
Model with HLE factors		AIC = 303.4470		R^2 = 0.0959	
HLERes	1.46	1.197	1.224	0.226	1.499
ReadAct	1.27	1.999	0.633	0.529	0.401
OthLiterAct	2.07	1.458	1.421	0.161	2.019
EmoMoti	-1.69	2.208	-0.764	0.448	0.583

Note. MLU5=Mean length of the five longest utterances, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D9

The Relationship between Pretest CONJ and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 207.0900		R^2 = 0.0767	
Gender*	-0.18585	0.3164	-0.5874	0.559	0.345
Age	-0.03098	0.0192	-1.6113	0.113	2.596
Income	-0.26136	0.1635	-1.5988	0.115	2.556
Medu	0.00511	0.1969	0.0259	0.979	6.73e-4
Pedu	0.07701	0.1459	0.5280	0.600	0.276
Model with HLE factors		AIC = 207.5990		R^2 = 0.0398	
HLERes	-0.0445	0.566	-0.0786	0.938	0.00618
ReadAct	-0.1726	0.945	-0.1826	0.856	0.03335
OthLiterAct	-0.9808	0.690	-1.4224	0.160	2.02322
EmoMoti	0.1762	1.044	0.1688	0.867	0.02849

Note. CONJ=The number of conjunctions, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D10

The Relationship between Pretest DfCONJ and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 186.7630		R^2 = 0.0865	
Gender*	-0.4036	0.2699	-1.4950	0.140	2.23504
Age	-9.03e-4	0.0164	-0.0550	0.956	0.00303
Income	-0.2699	0.1395	-1.9352	0.058	3.74505
Medu	0.2442	0.1680	1.4539	0.151	2.11376
Pedu	-0.0771	0.1244	-0.6194	0.538	0.38366
Model with HLE factors		AIC = 187.0500		R^2 = 0.0532	
HLERes	0.253	0.482	0.524	0.602	0.275
ReadAct	0.556	0.805	0.690	0.493	0.476
OthLiterAct	0.624	0.587	1.062	0.293	1.128
EmoMoti	-1.200	0.889	-1.349	0.182	1.821

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D11

The Relationship between Pretest Character Speech and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 143.837		$R^2 = 0.108$	
Gender*	-0.3537	0.1930	-1.832	0.072	3.3570
Age	0.0161	0.0117	1.373	0.175	1.8840
Income	0.0170	0.0997	0.170	0.866	0.0289
Medu	-0.1296	0.1201	-1.079	0.285	1.1636
Pedu	0.0889	0.0890	0.999	0.322	0.9984
Model with HLE factors		AIC = 148.54700		$R^2 = 0.00917$	
HLERes	0.2005	0.3567	0.5620	0.576	0.31585
ReadAct	-0.3326	0.5960	-0.5580	0.579	0.31137
OthLiterAct	-0.0272	0.4347	-0.0625	0.950	0.00390
EmoMoti	0.3112	0.6583	0.4727	0.638	0.22348

Note. *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D12

The Relationship between Pretest RIS and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 211.8300		$R^2 = 0.0529$	
Gender*	-0.3737	0.3283	-1.138	0.260	1.2955
Age	0.0265	0.0200	1.329	0.189	1.7654
Income	-0.0825	0.1696	-0.486	0.629	0.2364
Medu	0.0448	0.2043	0.219	0.827	0.0480
Pedu	0.0253	0.1514	0.167	0.868	0.0280
Model with HLE factors		AIC = 212.1580		$R^2 = 0.0179$	
HLERes	0.491	0.586	0.837	0.406	0.7001
ReadAct	-0.212	0.980	-0.216	0.830	0.0468
OthLiterAct	0.382	0.715	0.534	0.595	0.2853
EmoMoti	-0.447	1.082	-0.413	0.681	0.1704

Note. RIS=References to internal states, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D13

The Relationship between Pretest SGQ and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 390.889		$R^2 = 0.180$	
Gender*	1.019	1.3300	0.766	0.447	0.5867
Age	0.270	0.0808	3.342	0.001	11.1693
Income	1.199	0.6872	1.746	0.086	3.0469
Medu	-0.745	0.8277	-0.900	0.372	0.8108
Pedu	0.176	0.6131	0.287	0.775	0.0825
Model with HLE factors		AIC = 397.4840		$R^2 = 0.0618$	
HLERes	2.29	2.495	0.918	0.362	0.843
ReadAct	-7.13	4.167	-1.712	0.092	2.931
OthLiterAct	2.39	3.040	0.785	0.436	0.616
EmoMoti	5.23	4.603	1.137	0.260	1.292

Note. SGQ=Story grammar questions, *=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment

Table D14

The Relationship between Pretest PRQ and Control Variables

Variables	*β*	*SE*	*z*	*p*	Likelihood Ratio
Model with demographic factors		AIC = 70.5550		$R^2 = 0.0752$	
Gender*	-0.05384	0.10889	-0.494	0.623	0.2445
Age	-0.00557	0.00662	-0.841	0.404	0.7077
Income	0.01092	0.05626	0.194	0.847	0.0377
Medu	-0.03699	0.06776	-0.546	0.587	0.2980
Pedu	0.08317	0.05020	1.657	0.103	2.7456
Model with HLE factors		AIC = 72.2030		$R^2 = 0.0209$	
HLERes	-0.0848	0.1965	-0.432	0.668	0.1863
ReadAct	0.1161	0.3282	0.354	0.725	0.1250
OthLiterAct	-0.0705	0.2394	-0.294	0.770	0.0866
EmoMoti	0.3098	0.3625	0.854	0.396	0.7300

Note. PRQ=Problem resolution questions*=Male compared with female, Medu=Maternal' education, Pedu=Paternal education, HLE=Home literacy environment, HLERes=HLE resources, ReadAct=Reading activities, OthLiterAct=Other literacy activities, EmoMoti=Emotional and motivational environment