March 2022

Oral Narrative Interventions Implemented by Teachers, Speech-Language Pathologists, and Parents

Megan S. Kirby

University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd

Part of the Elementary Education and Teaching Commons, and the Social and Behavioral Sciences Commons

Scholar Commons Citation


This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.
Oral Narrative Interventions Implemented by Teachers,

Speech-Language Pathologists, and Parents

by

Megan S. Kirby

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Behavioral and Community Sciences Department of Child and Family Studies College of Behavioral and Community Sciences University of South Florida

Major Professor: Trina D. Spencer, Ph.D., BCBA-D Matthew E. Foster, Ph.D. Howard Goldstein, Ph.D. Elizabeth B. Hadley, Ph.D.

Date of Approval:
March 28, 2022

Keywords: participatory research, early language and literacy, online parent training, single case research design

Copyright © 2022, Megan S. Kirby
DEDICATION

It is with great pride that I dedicate this dissertation to my family and friends who have been my sounding board and source of strength during my doctoral student journey. To my husband, Chris, thank you for keeping me focused, reminding me of my goals, and for moving to Florida with me. I LOVE you. Thank you to my extended family and friends who continue to be my cheerleaders and a very reinforcing audience. Your text messages and visits have helped me get through periods of self-doubt. And finally, I dedicate this work to the memory of my father, Bruce E. Sullivan, who taught me everything I know about networking, persistence, generosity, and caring for all humans. I miss you, Dad.
ACKNOWLEDGMENTS

Be wary of the researchers who publish alone. I must acknowledge that my doctoral journey only culminates with this dissertation because of the researchers and participants who helped me get here. My journey began in 2017, when Dr. Howard Goldstein told me over the phone that I would be the perfect student for an incoming professor at USF. He saw my potential and personality strengths early on and has continued to encourage me to do better and be better. I am grateful for the ongoing support.

I am also very lucky that my PhD advisor turned out to be a perfect match. I would not be successful as a researcher or author without the mentorship of Dr. Trina D. Spencer. Everything I know about academic writing is because of you, Trina. You spent copious amounts time teaching me the basics and I am successful for it. Your unwavering belief that I can and should “level up” with each learning opportunity has made me a better researcher. Thank you for making this arduous PhD journey so much fun.

I am sincerely grateful to my dissertation committee for their generosity and support in getting me to this point. Dr. Liz Hadley, thank you for carefully going through every page of the proposal and dissertation draft and providing comments and suggestions to better my work. Thank you to Dr. Matt Foster for spending your time off to teach me HLM processes and for encouraging me to keep learning. And thanks to Dr. Goldstein and Dr. Spencer for your guidance and setting a high bar for me to achieve. Finally, a huge thank you goes to Dr. John Ferron for
inspiring a now life-long love of single-case research design methodology and analyses. Your enthusiasm is contagious, and my dissertation studies are better for it.

I must also acknowledge the individuals who inspire me to be a better human. Thank you to Noé Erazo for being my friend and colleague, as well as last-minute SAS trouble-shooter. To my friends who also happen to be fellow women business-owners/behavior analysts/professors, Barb and Beth, thank you for being just a text message or plane ride away and listening to my research ideas (or complaints). And last, but not least, I must acknowledge the sacrifices my husband, Chris Kirby (Mr. Doctor), has made to bring my PhD dream to fruition. Thank you for keeping the lights on when I worked on this dissertation through the night. Thank you for making sure I always had a meal to eat and that I took breaks instead of breaking down. And thank you for helping me get to the finish line. I look forward to learning with you what life will be like when I am no longer a student.
# TABLE OF CONTENTS

List of Tables ...................................................................................................................................v

List of Figures ................................................................................................................................ vi

Abstract ......................................................................................................................................... vii

Chapter One: Introduction ...............................................................................................................1
  Narrative Language..............................................................................................................1
  Narrative Language Intervention .........................................................................................4
  References............................................................................................................................8

Chapter Two: Oral Narrative Instruction Improves Kindergarten Writing ...............................12
  Note to Reader ...................................................................................................................12
  Abstract..............................................................................................................................12
  Introduction........................................................................................................................13
  Early Writing Instruction ...................................................................................................13
  Narrative Language............................................................................................................15
  Purpose and Research Questions .......................................................................................17
  Method...............................................................................................................................17
    Participants and Setting..............................................................................................17
    Research Design..........................................................................................................18
    Practical Action Research Framework.........................................................................19
    Dependent Measures .................................................................................................20
      Language complexity.................................................................................21
      Narrative structure .....................................................................................21
      Composite scores .......................................................................................21
    Oral Language Instruction ......................................................................................22
    Procedures..............................................................................................................23
    Research Assistant Training and Inter-rater Reliability .........................................25
    Data analysis ..........................................................................................................26
  Results...........................................................................................................................26
    Group 1: Hank and Beth ........................................................................................27
    Group 2: Carter and Zach ......................................................................................28
    Group 3: Emi and Sam ...........................................................................................28
  Discussion..........................................................................................................................29
    Limitations ..................................................................................................................32
    Implications ...............................................................................................................33
  References..........................................................................................................................34
Chapter Three: Effects of a Multitiered Oral Narrative Intervention on First Grade Language and Literacy Skills

- Oral Language Intervention ................................................................. 46
- Narrative Intervention ........................................................................ 48
- Multitiered Systems of Language Support (MTSLS) .......................... 51
- MTSLS and Narrative Intervention ...................................................... 53

Method ........................................................................................................ 59

- Context .................................................................................................... 59
- Study Design .......................................................................................... 60
- Setting and Participants ........................................................................ 60
- Schools .................................................................................................. 60
- Interventionists ...................................................................................... 60
- Students ................................................................................................ 61
  - Screening ............................................................................................ 62
  - Demographics ................................................................................... 62

General Procedures .................................................................................. 63

- Multitiered oral narrative intervention .............................................. 63
- Training and coaching .......................................................................... 65

Measures .................................................................................................... 66

- Proximal measures .............................................................................. 66
  - Narrative retell .................................................................................. 67
  - Vocabulary inferencing .................................................................... 68
- Intermediate measures ....................................................................... 68
  - Expository oral retell ........................................................................ 68
  - Narrative writing ................................................................................ 70
- Distal measures ................................................................................... 70
  - School-level outcome measures ..................................................... 71
  - Transcription fidelity and reliability .............................................. 72
  - Scoring reliability ............................................................................ 72
    - Expository oral retell ..................................................................... 72
    - Narrative writing ........................................................................... 73

Data Preparation and Analysis ................................................................. 73

- Examination of ICCs .......................................................................... 73
- Missing data ......................................................................................... 74
- Two-level hierarchical linear model specification ............................ 74

Results ......................................................................................................... 77

- Preliminary Analysis .......................................................................... 77
  - Equivalence of groups ...................................................................... 77

Intervention Implementation Fidelity ................................................... 77

- Examiner Fidelity ................................................................................ 78
  - Expository oral retell ........................................................................ 78
  - Narrative writing ................................................................................ 78

Student Language and Literacy Outcomes ......................................... 79

- Narrative retell .................................................................................... 79
- Vocabulary inferencing ....................................................................... 79
Chapter Four: Online Asynchronous Oral Narrative Language Training for Caregivers of Children Aged 3 to 5 Years Old .................................................................107

Introduction ......................................................................................................................107
  Parent-Implemented Early Childhood Intervention .........................................................109
  Online Parent Training ..........................................................................................110
  Behaviors Taught to Caregivers ...........................................................................111
  Oral Storytelling ........................................................................................................113
  Oral Narrative Intervention ..................................................................................114

Method .............................................................................................................................116
  Participants and Setting ........................................................................................116
  Study Design ........................................................................................................118
  General Procedures ..............................................................................................119
  Independent Variables .........................................................................................121
    Tell Me More modules .............................................................................122
      Module 1: Making sense with stories ..........................................122
      Module 2: Starting conversations with stories .............................123
      Module 3: Strengthening children’s stories ................................123
  Family activities ..................................................................................................126
  Data Collection ....................................................................................................129
  Transcription ......................................................................................................130
  Outcome Measures ...............................................................................................130
    Effectiveness ........................................................................................................130
      Oral narrative language ...........................................................................131
      Conversational turns .................................................................................132
      Story elicitation .........................................................................................132
      Story strengthening ....................................................................................133
    Scoring reliability .........................................................................................134
  Feasibility and acceptability ..........................................................................134
    Participation logs ..........................................................................................134
    End-of-module feedback survey ..................................................135
    End-of-study questionnaire ........................................................................135
  Data Analysis ......................................................................................................136
    Hierarchical linear modeling (HLM). ...................................................137
  Results ......................................................................................................................138
    Conceptual Analysis and Visual Inspection ...................................................138
      Oral narrative language ...........................................................................139
      Story elicitation skills .................................................................................140
LIST OF TABLES

Table 1: Steps of oral narrative instruction .................................................................41

Table 2: Participant outcome means, standard deviations, and within-case effect size
estimates by phase ........................................................................................................42

Table 3: Student participant demographics ...............................................................100

Table 4: Participant sample demographics—Interventionist ......................................101

Table 5: Descriptive data for variables in 2-level model with students nested in schools ....102

Table 6: Posttest and follow-up hierarchical linear models and ES estimates for
narrative retell and vocabulary inferencing after controlling for pretest and
school R-TFI scores ...................................................................................................103

Table 7: Posttest and follow-up hierarchical linear models and ES estimates for
expository oral retell and narrative writing after controlling for pretest,
examiner fidelity, and school R-TFI scores ................................................................104

Table 8: Posttest and follow-up hierarchical linear models and ES estimates for
WRMT-III Listening Comprehension and Passage Comprehension after
controlling for pretest and school R-TFI scores ........................................................105

Table 9: Participant demographics .............................................................................118

Table 10: Tell Me More intervention modules and related ingredients .........................121

Table 11: Means (M), standard deviations (SD), and range by variable and participant .....170

Table 12: Individual and group EB estimates of Tell Me More effect on caregiver oral
narrative language ......................................................................................................144

Table 13: Individual and group EB estimates of Tell Me More effect on caregiver story
elicitation skills .............................................................................................................144

Table 14: Individual and group EB estimates of Tell Me More effect on caregiver story
strengthening skills ....................................................................................................145

Table 15: Caregiver responses to end-of-study questionnaire: overall, training modules,
and family activities .................................................................................................171
LIST OF FIGURES

Figure 1: Results of multiple baseline design across three groups..............................................43
Figure 2: Student writing samples from baseline and intervention phases. ...............................44
Figure 3: Post-intervention participant (Carter) writing sample ................................................45
Figure 4: Example: Screen capture from family activity 2C..........................................................129
Figure 5: Multiple baseline design across behaviors with replications across participants ......172
ABSTRACT

Children’s oral language skills in preschool through early elementary grades can predict reading, writing, and social outcomes ten years into the future. Oral narrative language, which includes storytelling, is a long-established cultural practice in communities around the world. Narratives are the monologic re/telling of a real of fictitious event and people have used them for centuries to entertain, make sense of current and past events, and provide instruction. Oral storytelling does not require physical materials and can be tailored to the cultural and linguistic values of the community in which they are used. Thus, the portability, utility, and ubiquitous nature of storytelling makes it a prime candidate for use as a scalable oral language intervention.

This dissertation presents three studies of oral narrative interventions implemented with children in preschool through first grade by teachers, speech-language therapists, and parents. Interventions were delivered by their intended end-users in a variety of natural settings and contexts. Using two multiple baseline single-case experimental designs and one randomized controlled trial group experimental design, study results indicated that the intervention is effective for improving kindergarten and first grade students’ narrative writing, as well as first graders’ narrative retell, expository retell, and vocabulary inferencing skills. Across all three studies, end users found the intervention to be both feasible and acceptable. These three studies contribute additional knowledge of oral narrative interventions used in real-world conditions by a child’s natural caregivers.
CHAPTER ONE:
INTRODUCTION

For decades, researchers and community partners have attempted to reduce disparities in children’s health and academic outcomes by strategically focusing on improving language nutrition in early childhood (Zauche et al., 2017). Language nutrition is a term describing rich and substantive language ingredients present in a child’s environment that provide nourishment as they develop “neurologically, socially, and linguistically” (Zauche et al., 2016, p. 494). More than socioeconomic status, race, gender, and parental level of education, language nutrition in early childhood is the best predictor of children’s later academic achievement (Dickinson & Porche, 2011; Hart & Risley, 1995; Hoff, 2013; Rowe, 2012).

To be scaled at a level necessary to impact language nutrition disparities in early childhood and beyond, interventionists must design programs that build off the cultural, linguistic, and social capital existing in the child’s natural environment (Bourdieu, 1986; Bourdieu & Passeron, 1990). To do so will require the identification and use of existing behavioral practices in homes and schools, as well as the establishment of research-practice partnerships (RPPs; Alonzo et al., 2022) with primary stakeholders for whom the intervention is ultimately intended. Narrative language is one potential set of behaviors primed for intervention, in that storytelling is an established practice in many communities, homes, and schools.

Narrative Language

Narratives involve the monologic telling of a real or fictitious event, such as sharing a personal experience or generating a story, for entertainment and/or instructional purposes.
Narrative language develops in early childhood and is used in storytelling, helping children learn linguistically and culturally-specific social communication forms and functions. Narrative form refers to the overall structure and sequence of a story, with slight cultural variations in what parts of a story make a complete narrative and the ordering of those parts called “story grammar” (Mandler, 1977; Stein & Glenn, 1979). For example, American English linear narratives (written or spoken) usually present story grammar in chronological order. These narratives include an initiating event to introduce character(s), story context, and problem. Then, the remainder of the plot centers on the character’s actions or attempts to resolve the problem. While inclusion of a time variable enhances story coherence, only the sequential order of events is essential to the construction of a complete narrative. However, for American Sign Language (ASL), temporal information is integral to interpretation of signed narrative schema and nonverbal behavior like shifts in eye gaze and facial expressions indicate a change in character/perspective (Okrent, 2006; Rathmann et al., 2007). Although their expressive narrative structures may differ (e.g., noun, verb, and referent markers), the content and form of signed and spoken narratives ensure that the sharing of a story meets a specific purpose (Chamberlain & Mayberry, 1999). Story structure serves the purpose of creating meaningful and coherent stories; narrative form directly impacts its quality and function.

There are several different functions of narrative language. Storytelling can be used as a sense-making device, mechanism for entertainment, information channel, and early language and literacy instructional tool (Baynham, 2000). Individuals can generate (or retell) personal (or fictional) stories to meet any of these purposes. Oral and written narrative retells can be methods of preserving family history, or used to convey culturally-bound knowledge, traditions/practices, and values. Frequent conversations about shared experiences with parents and caregivers helps
children learn the cultural norms of their community as well as their community’s storytelling practices (Fivush et al., 2006; McCabe, 1997). As Cajete (2017, p. 114) writes, “…listening and thinking about stories is the first foundation of Indigenous education.” In addition, parents of young children sometimes generate fictional narratives to teach social and behavioral expectations (Silva & Rogoff, 2020). For example, Inuit parents use storytelling as a preventative disciplinary tactic to warn children of environmental dangers or the consequences of not following a rule, which is correlated with Inuit mothers having lower levels of stress compared to non-Indigenous mothers of western European ancestry (Briggs, 1970; Doucleeff, 2019).

In addition to conveying information, storytelling can be beneficial to individuals’ mental and behavioral health. To support individuals’ ability to make sense of lived experiences, narrative exposure therapy (NET) uses narrative recall of past trauma and logical sequencing of autobiographical memories which result in moderate to large effects on the reduction of post-traumatic stress in refugee and displaced children (Gwozdiewycz & Mehl-Madrona, 2013; Kasmani, 2021). Beyond the therapeutic context, everyday interactions with caregivers (e.g., play routines, storybook reading, and sharing of oral narratives) help children develop autobiographical and personal narrative skills (Reese, 2018). The extent and quality of parents’ narrative language plays a key role in their children’s development of oral narrative competencies (Jordan et al., 2000; Marjanovic-Umek et al., 2012; Peterson et al., 1999; Snow & Dickinson, 1990).

Decades of research have revealed a strong link between narrative abilities in early childhood and later reading achievement for a variety of student populations (Catts et al., 1999; Chamberlain & Mayberry, 1999; Griffin et al., 2004; Lervåg et al., 2018; Snow et al., 2007; Storch & Whitehurst, 2002; Suggate et al., 2018). For example, a longitudinal study by Babayigit
et al. (2020) found linguistic comprehension and narrative skills by age five predict listening comprehension, vocabulary, and reading achievement at age 14. Several studies have shown oral narrative skills in kindergarten to significantly predict students’ narrative writing skills in first grade (Berninger & Graham, 2008; Pinto et al., 2016). Chamberlain and Mayberry (1999) reported strong correlations between Deaf students’ narrative and reading comprehension skills in elementary and high school. The role of oral narrative skills in academic and social development, as well as being a ubiquitous practice across cultures and linguistic communities, makes it an ideal repertoire to target in early childhood. Parents and educators must be equipped with the tools and knowledge to strategically teach narrative language to young children.

**Narrative Language Intervention**

Systematic oral narrative language instruction emphasizes the use of complex linguistic forms, the structure of language, advanced vocabulary, and comprehension of new content through academic discourse (Peterson et al., 2020). It is critical that adults provide quality language models and appropriate levels of support to young children developing storytelling because children develop narrative skills through observation and social interactions with primary caregivers (Grigoroglou & Papafragou, 2019). It is well known that adult-child interactions before, during, and after storybook reading can have a significant influence on children’s acquisition of new vocabulary (e.g., Hadley et al., 2020; Seven & Goldstein, 2019). However, given the complex nature of conversational behaviors and its related linguistic features, oral narrative language acquisition may be difficult for some children to acquire through observation and hearing alone. Children who exhibit oral language problems in preschool through first grade may require direct and intentional oral narrative intervention. While evidence of oral narrative intervention effects on students’ listening comprehension,
vocabulary, reading comprehension, and writing skills is strong, there remains a need to study the extent to which effects may (not) be noticeable when programs are implemented by natural end users in various contexts.

This dissertation includes three intervention studies that examined oral narrative languages interventions when implemented by primary caregivers (teachers, speech-language pathologists, parents) in the natural teaching environment (classroom, home, car, neighborhood). Innovative research methods were employed to capture intervention effects in real world conditions, to include a multitiered oral narrative intervention and an online asynchronous parent training program. Outcomes of interest across all studies included treatment effects, as well as social validity, intervention feasibility, and acceptability.

The first study examined the effects of a teacher-implemented oral narrative language intervention on kindergarten students’ narrative writing. A multiple baseline across groups design was used to assess the extent to which the oral language intervention had an effect after six small group lessons, and to examine any maintenance effects after three to four weeks. Significant improvements in the six participants’ narrative writing quality were observed and gains were maintained after the intervention ceased. The experiment is one of the first to provide causal evidence of a functional relation between oral language instruction and narrative writing in kindergarten. The study was published in *Reading & Writing Quarterly* (see Kirby et al., 2021).

The second study investigated the extent to which a multitiered oral narrative intervention implemented within a Multitiered System of Language Supports (MTSLS) framework by teams of teachers and speech-language pathologists (SLPs) at 10 elementary schools improved the language and literacy skills of 155 first graders identified at risk for or having language
disabilities. Students attending schools assigned to the waitlist control condition received classroom-based English language arts instruction as-usual, whereas schools randomly assigned to the treatment condition delivered an adaptation of the multitiered oral narrative intervention, *Story Champs* (Spencer & Petersen, 2016). Proximal (narrative retell, vocabulary inferencing), intermediate (expository retell, narrative writing), and distal measures of students’ language and literacy skills were collected at pretest (fall 2018), posttest (spring 2019), and follow-up (fall 2019). We used two-level hierarchical linear modeling (HLM) to examine the extent to which the intervention resulted in statistically significant improvements in student outcomes, accounting for student scores nested within schools and level-1 and level-2 covariates. We found moderate to large effects for several outcomes at posttest and follow-up. Results of the second study contribute to small but growing evidence of the effectiveness of a multitiered oral narrative language curricula implemented by elementary schools teams using a MTSLS framework (Petersen et al., 2022).

The third study explored the preliminary effectiveness, feasibility, and acceptability of an online caregiver training to support preschoolers’ oral narrative development. Prior parent-delivered intervention research to develop preschoolers’ oral narrative skills has focused on maternal elaborative reminiscing style (e.g., Kelley, 2018) or the use of researcher-contrived events to elicit independent or co-constructed dyad conversation (Cleveland & Reese, 2005; Fivush et al., 2006). To detect potential treatment effects of the *Tell Me More* (TMM) program, we used a multiple baseline across behaviors design with four replications across participants. The multiple baseline design allowed us to visually inspect caregiver responding in three targeted conversational skill sets that corresponded to the staggered introduction of three different TMM modules. Caregivers videotaped informal conversations with their child and the
recordings were transcribed verbatim to measure caregiver behavior change at the individual word level. Surveys collected parents’ responses to questions about intervention acceptability and logs were maintained to document feasibility. In addition to reporting the effects of the intervention on changing caregiver verbal behavior, we obtained feedback from caregivers that allowed for the identification of barriers and facilitators to intervention uptake and program adherence. Outcomes will inform future changes to the online module content and procedures used to measure the dependent variables.

The presented studies provide a wide-angle view of facilitators and barriers to the implementation of oral narrative interventions by intended end users in children’s natural environments. Community-based participatory research, while a challenging approach, can produce outcomes worthy of practitioner action. This research line has also contributed to the knowledge of caregivers’ role in children’s narrative language development. In addition, results lend additional information to a small but growing pool of literature discussing implementation and facilitators in the design of early childhood parenting programs focused on the delivery of language nutrition. Results from all three studies have the potential to contribute to a greater understanding of the effect of oral narrative language intervention on a child’s academic, social, and emotional development. Further, efficacy results are to be taken into consideration alongside measures of feasibility and acceptability, as all interventionists were active participants in research procedures like study design, intervention implementation, and data collection. Ultimately, such knowledge has the potential to improve the design of effective and adaptive language interventions, acknowledging that effective interventions are only implemented if they are also feasible, acceptable, and liked by their intended end users.
References


CHAPTER TWO:

ORAL NARRATIVE INSTRUCTION IMPROVES KINDERGARTEN WRITING

Note to Reader

This chapter is reproduced with the permission © 2021 The Author(s) with license by Taylor and Francis Group, LLC. Citation: Kirby, M. S., Spencer, T. D., & Chen, Y-J. I. (2021). Oral narrative instruction improves kindergarten writing. *Reading & Writing Quarterly, 37*(6), 574-591. [https://doi.org/10.1080/10573569.2021.1879696](https://doi.org/10.1080/10573569.2021.1879696)

Abstract

Writing is a critical literacy skill that emerges in kindergarten. The research literature has only addressed transcription skills of kindergarteners and has failed to address text generation. The purpose of this action-research study was to investigate the effect of oral language instruction that focused on narrative text structures on kindergarten students’ ability to generate written narrative text. We conducted a concurrent multiple baseline design across three groups of students with two participants in each group. Students received six instructional sessions that involved the teacher modeling a story and supporting the students while they retold and generated oral stories. Pictures and icons were used to represent story grammar elements, but were faded within session to facilitate independent storytelling. The oral language instruction had an immediate positive effect on the narrative quality of students’ writing. Individual and overall effects were significant and maintained three to four weeks later. Findings suggest an efficient causal relation between oral language instruction and writing quality.
Introduction

Writing is a critical 21st century literacy skill and plays a conspicuous role across the lifespan. As children age and move beyond expressing wants in holiday wish lists and signing their name in cards, writing becomes a method of influence and a measure of achievement. Students with proficient writing skills can effectively communicate to others what they know and how well they know it. Furthermore, student writing can serve as an indicator of post-secondary success. In an increasingly competitive application process, most colleges and universities require written narratives and exams, evaluating student preparedness for higher-level cognitive demands. In the workforce, 82% of employers list written communication as a highly requisite job skill and use resumes, emails, and applications to evaluate applicants’ writing skills (National Association of Colleges and Employers, 2019). Sadly, 23% of U.S. eighth grade students and less than 33% of high school seniors demonstrate proficient writing skills (National Center for Education Statistics, National Assessment of Educational Progress [NAEP], 2011). To prepare students with the literacy skills necessary for success in today’s society, writing instruction must begin early and be effective for all students.

Early Writing Instruction

Most students are formally introduced to writing when they enter kindergarten. Very quickly, writing takes a prominent role in classroom activities and its importance is emphasized in grade level standards. For example, by the end of their kindergarten year, students should be
able to use a combination of drawing, dictating, and writing to narrate “a single event or several loosely linked events, talk about the events in the order in which they occurred, and provide a reaction to what happened” (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p. 19). This means kindergarteners are expected to transition from single-letter formation to writing complete sentences that communicate a story in a single year.

From the theoretical perspective of a simple view of writing (Berninger, 1999; Berninger & Graham, 1998; Juel et al., 1986), the process of writing requires multiple skills: working memory, executive functioning, transcription (e.g., handwriting and spelling), and text generation (e.g., word choice, syntax, grammar, and text structure). All aspects of writing are needed to translate ideas into working memory and transfer information from working memory to paper or screen. Thought to develop in parallel (Berninger et al., 1996), transcription and text generation skills are particularly important to the writing of a cohesive and coherent story. As students acquire lower-level transcription skills, they also develop higher-order oral language skills necessary for text generation. The link between oral and written language is so well accepted that in a review of literature measuring writing in early childhood, Quinn and Bingham (2018) found ten studies that used oral retells as an alternative measure of young children’s text generation skills. For example, Hsieh, Ku, and Chen (2013) asked young children to “read” a wordless picture book and dictate their retell to an adult. The transcriptions were read back to the children to allow for revision of the written product measuring story retell. While research helps to establish the relationship between oral language and text generation, there is little evidence that this influences how early writing instruction is conducted (Campbell et al., 2018; Datchuk & Kubina, 2012; Korth et al., 2017; McMaster et al., 2017; Schrodlt et al., 2019).
Early writing instruction typically is based on a linear theory of writing (Flower & Hayes, 1981; Berninger et al., 1996). From this perspective, writing develops in a predictable and sequential pattern (Berninger et al., 1997; Berninger & Swanson, 1994) and text generation is thought to emerge once proficient transcription skills are established (Graham et al., 1997; Puranik & Al Otaiba, 2012; Puranik et al., 2020). In what precious little time is made available to writing instruction, primary grade (K-2) teachers tend to focus on teaching transcription skills (Cantin & Hubert, 2019; Makdissi et al., 2019; Teale et al., 2020). Although legibility and spelling may increase text production and reduce effort needed to judge student writing, transcription training alone does not lead to improvements in other areas of writing (Graham et al., 2018). Therefore, if students have weaker transcription skills, teachers may sacrifice higher-level text generation, planning and revision instruction to provide more handwriting and spelling practice. To prevent emergent writers from being left behind, young students may benefit from writing instruction that facilitates the development of text generation skills regardless of handwriting proficiency. The question that remains is whether text generation skills can be enhanced through oral language instruction.

**Narrative Language**

One type of oral language that may bridge oral language and literacy is narrative language (McCabe & Bliss, 2003). For example, longitudinal studies have found children’s narrative-based oral language and listening comprehension skills at age five directly contribute to reading comprehension skills assessed at 10 years of age (Babayigit et al, 2020; Hayiou-Thomas et al., 2010). Instruction in oral narrative language has been shown to significantly improve the reading comprehension skills of students in second through fourth grades (Clarke et al., 2010;
Narratives have been long considered a tool of metacognition that promotes connectivity between listening and reading comprehension (Short & Ryan, 1984). Narratives are a familiar and accessible form of communication for children, and both oral and written narratives share similar linguistic features (Pinto et al., 2015). Narrative language also includes the ability to organize story elements in a meaningful, causal, and temporal sequence, a key part of text generation. Unlike informational structures that require explicit training due to their variety, complexity and unfamiliarity, there is a generally accepted pattern for stories (Duke et al., 2011). Moreover, narratives are ubiquitous in the lives and environments of young students (McCabe, 2017; Westby et al., 2016), whereas informational structures do not become commonplace until later elementary years (Duke, 2000). Oral language interventions do not require students to have proficiency in informational genres (Traga Phillippakos, 2019).

However, most of the research on writing has overlooked narratives and examined text-based interventions (Hebert et al., 2016; Pyle et al., 2017; Williams et al., 2014).

Supported by schema theory (Anderson, 1984; Mandler, 1984), the development of oral narrative language should also contribute to the development of written narrative skills. In a 2015 study, Kim and colleagues found that kindergarteners’ oral language predicted their narrative writing skills in third grade. Based on the schema theory and correlational research, as students become proficient in oral narrative skills and aware of the shared story schema (Stein & Glenn, 1979), key narrative elements, such as story grammar (e.g., character, setting, problem, plan/action, consequence, ending), may transfer from oral to written language. If true, instruction that capitalizes on oral and written narrative structures could be used to teach text generation and hasten the development of writing among young students without requiring proficient transcription as a pre-requisite. In fact, there is emerging evidence that this may be true. Building
upon numerous oral narrative language interventions that improve students’ oral language skills, Spencer and Petersen (2018) investigated the oral to written language relationship. Indeed, they found that an oral narrative language intervention improved first graders’ writing quality.

**Purpose and Research Questions**

What is known about writing instruction and the development of writing skills has been gleaned from intervention research with students in first grade or higher (Finlayson & McCrudden, 2020; Graham et al., 2012). However, to date, there are no writing intervention studies with kindergarten students that explicitly evaluate text generation outcomes (Kent & Wanzek, 2016; Graham et al., 2017). To address the dearth of kindergarten writing instruction research and to further document an oral to written language causal relationship, we replicated Spencer and Petersen’s (2018) study with kindergarten students. We hypothesized that teacher-implemented small group oral narrative language instruction would lead to improvements in kindergarten students’ narrative text generation. Using an action research framework (Biancarosa & Snow, 2004; Creswell & Guetterman, 2019), we addressed the following questions:

1. What is the impact of low-dose, small group oral narrative language instruction on the quality of kindergarten students’ narrative writing?

2. To what extent does oral narrative language instruction lead to maintained improvements in kindergarteners’ narrative writing skills once instruction is withdrawn?

**Method**

**Participants and Setting**

This study began in January of the academic calendar year in a mixed kindergarten-first grade general education classroom located in a western state. The classroom was led by a first-year teacher, dual-certified in elementary and special education. Kindergarten students were
selected for inclusion in the study based on their ability to consistently produce legible writing samples. Students unable to consistently produce interpretable writing samples were excluded because otherwise, we would not be able to measure written text generation. Oral retells could have been used as a proxy measure. However, several studies have already established the effect of oral narrative intervention on oral retells (Adlof et al., 2014; Brown et al., 2014; Petersen et al., 2016; Spencer et al., 2013). We were interested in taking this work a step further to investigate the impact on writing. While it was unnecessary to include students without a minimal level of transcription skills to answer our research questions, we expect that the results of this study would be relevant for promoting their text generation skills regardless of students’ mastery of transcription. In total, six typically developing kindergarteners were selected to participate in the study from a total classroom sample of 22 students. All six kindergarten participants spoke English and came from middle-class backgrounds. Four of the six participants were identified as White, Non-Hispanic and two participants identified as Latino (Beth and Emi). Neither the teacher nor students had exposure to oral narrative instruction prior to the study.

Research Design

To investigate the causal relationship between the introduction of the oral narrative instruction and the quality of the kindergarteners’ writing samples, we conducted a concurrent multiple baseline design across three groups of students (Ferron & Scott, 2005). Group 1 included Hank and Beth, Group 2 included Carter and Zach, and Group 3 included Emi and Sam. Each group experienced baseline, intervention, and maintenance conditions (see Procedures below). The research team did not have access to the daily writing samples during the study. As a result of not being able to start the treatment condition according to baseline stability, we
determined the most conservative approach was to randomly assign groups (Kratochwill & Levin, 2014) a priori to one of three baseline lengths: five, seven, or nine school days.

**Practical Action Research Framework**

Since the research was initiated by the classroom teacher to improve her writing instruction, a practical action research approach was used in the design of the study. In an educational context, practical action research typically starts with a teacher or a group of educators identifying a specific problem within the school setting and implementing selected solutions (Creswell & Guetterman, 2019; Mills, 2018). For this study, the teacher initiated a partnership with the researchers for the purpose of using the results to inform her classroom writing instruction. She reported having no intentional or systematic method for teaching writing other than giving the students time to write and allowing them to ask the teacher how to spell words. Researchers provided consultation and support to the teacher as she identified research questions and an interest in using oral narrative language instruction to improve students’ story writing.

Another advantage of practical action research is that it enhances the likelihood of results generalizing to practical classroom settings (Nugent et al., 2012; Postholm, 2020). For the current study, the teacher completed all of the instruction and collected students’ writing samples to be scored later by researchers. She integrated study activities into the classroom’s existing literacy block, in which she regularly provided students the opportunity to write stories and conducted small group reading instruction. Students were organized into small groups for literacy center rotations according to their reading levels and each reading group was composed of four to five kindergarten and first grade students.
Researchers provided ongoing planning and consultation (e.g., data collection and schedule phases), but were not present in the classroom, except for the first intervention session. While the practical action framework has several advantages, it resulted in some loss of experimenter control over the procedures, particularly the direct and independent measurement of intervention fidelity and the fidelity with which she collected writing samples from her students.

**Dependent Measures**

Writing quality was defined as the extent to which written samples reflect canonical narrative structure and complex language. The *Narrative Language Measures (NLM) Flow Chart* (Petersen & Spencer, 2019) is a scoring rubric with a decision-tree format that allows for quick scoring of narrative structure, language complexity, and writing conventions (see Appendix A). It was developed to align with oral and written academic language expectations set forth by the Common Core State Standards (CCSS; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). In research, the *NLM Flow Chart* has been used to detect incremental changes in children’s oral (Spencer et al., 2013) and written (Spencer & Petersen, 2018) narrative language related to intervention. These studies documented adequate *NLM Flow Chart* scoring agreements (87-96%) and reliability correlations (.57-.69). In this study, it was not our purpose nor is it logical that an oral language instruction would improve students’ transcription skills in just a few weeks. The teacher also allowed students to ask for spelling support. Therefore, students’ writing samples were only scored for features related to text generation (i.e., Narrative Structure and Language Complexity scales) and not writing conventions such as punctuation, capitalization, and spelling.
**Language Complexity.** The Language Complexity section of the *NLM Flow Chart* includes common linguistic features expected of elementary students’ oral and written language. These include the following items: Relative Pronouns, Verb/Noun Modifiers, Vocabulary/Rhetoric, Temporal Ties, Causal Ties, and Dialogue. Except for Dialogue (which is scored on a 0-2 scale), Language Complexity items are scored on a 0-3 scale. Higher scores represent greater quantity and quality of those linguistic features in the language sample. A subscale score of the complexity of language included in the sample is calculated.

**Narrative Structure.** The Narrative Structure section of the *NLM Flow Chart* highlights specific story grammar elements similar to and modeled after the story schema outlined by Stein and Glenn (1979). This section measures the presence or absence of an appropriate story sequence, as well as the clarity and completeness of the following story grammar elements: Character, Setting, Problem, Plan/Attempt, Consequence, Ending, and Emotion. Although most structural elements are given a score ranging between 0 (not present) and 3 (complete and clear), samples that include more than one Problem, Plan/Attempt, and/or Consequence are awarded an additional point for each additional component. Furthermore, the *NLM Flow Chart* is sensitive to variations in the magnitude of episodic complexity, with weight given to samples that include a minimum of two complete and clear episodic elements (i.e., Problem, Plan/Attempt, Consequence, and/or Ending). Episodic sophistication differentiates basic or incomplete stories from episodes that are more complex, while remaining clear and complete.

**Composite Scores.** A composite writing quality score for a specific sample was obtained by summing the Language Complexity and Narrative Structure scores. Composite scores were graphed according to multiple baseline design conventions and are shown in Figure 1. A composite score of 50 is possible at the ceiling of *NLM Flow Chart*, with 33 points possible for
Narrative Structure and 17 points possible for Language Complexity. However, students in this study produced writing samples ranging from 0-26 points. Because the NLM Flow Chart is designed for students in grades K-5, scores in this range are developmentally appropriate for kindergarten students and align with CCSS K-1 writing expectations.

**Oral Language Instruction**

A multi-tiered academic language program called *Story Champs* (Spencer & Petersen, 2016) was used for the oral narrative instruction. Several evidence-based practices of the program include modeling, explicit instruction of story grammar, visual and verbal scaffolding, and multiple opportunities for students to retell the model story. One of the critical features of *Story Champs* is the use of 24 carefully constructed stories that feature themes familiar to young children (e.g., falling and getting hurt, being a picky eater, losing a game, disagreeing with a sibling or friend) and contain complex academic language. Although *Story Champs* can be used flexibly with students from preschool to third grade, the stories written according to kindergarten standards were used in this study. These stories are 108-110 words in length and include the following story grammar elements: Character, Setting, Problem, Feeling, Action, Consequence, Ending and End Feeling. Stories are written to be equivalent in terms of linguistic complexity too (i.e., same number of subordinate and relative clauses). A set of five illustrated picture cards (3 inches by 5 inches, printed in color on laminated cardstock) accompany each story. Story grammar icons (1 inch in diameter, printed in color on cardboard circles), representing the main story grammar elements, are used in the lessons as visual supports. Although the picture cards are specific to each different story, the story grammar icons are applicable to all stories.

Every lesson followed the same steps but featured a different *Story Champs* story (See Table 1). Each lesson lasted approximately 20 minutes. In the first step, the teacher displayed the
picture cards on the table in front of the students and read the model story. As she read each part, she laid a story grammar icon on the corresponding picture card. In the second step, the teacher pointed to the icons and named the story grammar parts. Students were asked to repeat the story parts (i.e., “Character, Setting, Problem, Feeling, Action, Ending, and End Feeling”). For step three, the teacher picked up the icons and distributed one or two icons to each student at random. Students individually contributed to a group retell based on their assigned story grammar part until the entire story was retold. At the end of the group retell, the teacher summarized the whole story to present a more cohesive model. In the fourth step, students retold the story, one at a time. For the student who retold the story first, the picture cards and story grammar icons were left on the table for reference. However, with each student retell to follow, the teacher systematically removed the pictures and then the icons. For example, the first student retold the story having access to both story grammar icons and picture cards, while the last student retold the same story in its entirety without a single visual support (picture card or icon). Finally, the teacher redistributed the icons to the students. Based on the one or two icons each student received, the students collectively generated a novel story as the teacher drew stick figures on the white board. The teacher used a two-step prompting procedure to support individual students to retell each part. This involved first asking an open-ended question such as “Who was the story about?” or “What did he do to fix his problem?” If that level of prompting was successful, the student continued to retell the story. If it was unsuccessful, the teacher followed the question with a model and asked the student to repeat it. For example, “Hannah was walking downstairs. Now you say that.” Or “Say it like this: Hannah asked her sister if they could play together.” The teacher used this prompting procedure only when necessary so that students retold the story as independently as possible.
*Procedures*

During a group’s baseline condition, students in that group did not receive the oral narrative instruction. Once the instruction phase began with each small group, the teacher delivered six oral narrative lessons (each lasting about 20 minutes) over a two-week period as part of students’ core instruction. Although not all students in the groups were research participants, they all received the instruction. Across baseline, treatment, and maintenance conditions, the teacher gave students opportunities to write stories. This practice occurred at the beginning of the literacy block, prior to small group rotations. Students were given writing paper and asked to produce a story based on a topic of interest or a general prompt (e.g., “Write about what you did last weekend.”). They were accustomed to the task. Not all students were given writing practice on the same day, but most students were asked to write a story two to three times per week. Students absent from the classroom (or school) during the writing period did not produce writing samples at other times of the day; therefore, missing samples are considered missing at random.

Through email and phone calls, researchers provided the teacher with schedules and directions regarding when to begin and end phases for each group. Additionally, researchers directed the teacher to temporarily pause students’ production of writing samples for a period of three to four weeks after each group’s instruction ended. Then, after the break, students produced additional writing samples. The purpose of this was to document the extent to which the students’ maintained their writing quality after three to four weeks without writing instruction or writing practice. As instruction started with Group 1, the second author demonstrated the first lesson as part of the teacher’s training and explained how to use the step-by-step procedural checklist of the lesson to guide her instructional sessions (similar to those in Table 1, but with
more specificity). The teacher completed the lessons thereafter independently. No other training
was provided. The teacher-maintained session logs to document the dates in which lessons
occurred and followed the procedural checklist in every session to ensure adherence to the
manualized procedures. Albeit self-report, the teacher completed all procedural checklist steps in
all of the sessions (i.e., 100% self-reported fidelity). Not unusual for methods assuming an action
research approach, the researchers were not involved in the daily implementation and did not
conduct additional fidelity observations. Alternatively, researchers collected the log and
completed checklists at the end of the study, along with the students’ writing samples. The
teacher ensured identifying information was removed and that each sample was dated properly
before giving them to the researchers to score. The teacher did not score the writing samples
because she was familiar with the students’ work (and handwriting). Writing samples were de-
identified and randomized during the measurement process to reduce the risk of bias in scoring
participant writing samples.

**Research Assistant Training and Inter-rater Reliability**

The second author trained two research assistants to score the samples for this study.
Following didactic training and practice, research assistants were required to achieve a minimum
criterion of 85% scoring accuracy on a set of three writing samples unrelated to this study (range:
87-100%). Before the research assistants began scoring, the second author masked all samples by
removing the dates and identification numbers. They were rearranged in a random sequence so
that scorers could not identify the participant or condition to which the sample belonged. Only
the second author had access to the identification keys. The primary scorer was given all of the
samples to score and the secondary scorer was given a randomly selected subset (42%) of the
samples to document scoring reliability. For the portion of samples scored by a second research
assistant, point-by-point scoring agreement was calculated by dividing the number of agreements by the total number of scored items (n=14), multiplied by 100. Mean scoring agreement was 96% (range: 87 to 100%).

**Data Analysis**

To determine the causal relation between the oral narrative instruction and writing quality, we analyzed graphical displays of data according to level, trend, variability, and immediacy (Kratochwill et al., 2010; Kratochwill et al., 2014). Such analyses were conducted comparing students’ responding in the treatment and maintenance conditions with baseline patterns of responding. We also calculated effect sizes to document the magnitude of the oral narrative instruction’s impact on writing quality. Within-case effect sizes, which reflect practical significance for each student individually, are useful when comparing findings from multiple single case design studies. For the within-case effect sizes, we subtracted baseline means from treatment and maintenance means and divided each mean difference by within-case variability

\[
\frac{\text{Mean}_{T or M} - \text{Mean}_B}{SD_{\text{within}}}
\]  

(Shadish et al., 2015). To allow for comparison of group results, we also calculated between-case effect size estimates using the R package, *scdhlm* (Pustejovsky et al., 2014). We used the code `effect_size_MB(outcome, treatment, id, time, phi, rho)` to calculate the between-case effect size (see Hedges et al., 2013 for a detailed estimation process). The statistical significance of effect sizes was derived using \(z\) statistics.

**Results**

A functional relation between oral narrative instruction and written narrative performance was examined through a concurrent multiple baseline design study across three groups of kindergarten students. Majority of the improvements in student writing were noted for Narrative Structure, with little to no increase (only 0-2 points) in Language Complexity. As a
result, it was not helpful to graph the Language Complexity scores separately. We relied exclusively on the single *NLM Flow Chart* composite scores for analysis. Each participant’s graphed composite scores are displayed in Figure 1. Several permanent product records of participant writing samples are provided in Figures 2 and 3. Across all three groups, visual analysis of individual observations revealed a functional relation between the oral narrative language instruction and written narrative skills, although the size of the effect varied across cases. Across participants, the corrected between-case effect size for the treatment condition was 2.85, with a standard error of 0.48. This effect was statistically significant; $z = 5.94$, $p < .001$, 95% CI [1.90, 3.79]. A more thorough description of group and participant-specific results are provided below. Within-case effect sizes are included in Table 2. Each participants’ treatment and maintenance effect size estimates are considered large and meaningful (Cohen, 1988; Durlak, 2009).

**Group 1: Hank and Beth**

A total of four samples were collected during baseline for Group 1. Hank’s written narrative performance was low and stable during baseline, with an average writing sample score of 3.33. Beth only produced one writing sample with a score of 3.0 during baseline, due to absences during writing periods scheduled by the teacher. Following the introduction of the intervention, a large and immediate increase in composite scores was noted for both participants. During the treatment phase, Hank’s mean composite score was 23.25 and Beth’s writing composite score average was 22.75. Neither participant’s scores during intervention overlapped with scores reported during baseline. Across all participants and corrected for bias, Beth had the largest estimated within-case treatment effect size, Hedges’ $g = 2.26$. The estimated size of treatment effect for Hank was $g = 2.0$. Following conclusion of the study, writing samples
collected from Hank revealed that he maintained high levels of writing performance, averaging a composite score of 25.5. Hank’s estimated maintenance effect size was $g = 2.23$.

**Group 2: Carter and Zach**

Carter and Zach provided a total of seven samples during the baseline phase, with both students’ writing samples scored at near-zero levels. Carter’s average writing sample composite score at baseline was 1.75 and Zach’s average was 1.67. Upon receiving the oral language instruction, both students showed an immediate improvement in writing performance, although the size of initial improvement varied by participant. Overall, Carter’s average composite score during intervention was 16.89 and Zach’s average score was 17.25. Variability in Zach’s written performance during the first three intervention days influenced the overall average, as two samples were assigned scores of 2 and 4. However, his remaining intervention composite scores ranged from 15 to 25. Despite the influence of these outliers, Zach’s treatment effect size estimate was $g = 1.41$. Carter’s estimated treatment effect was $g = 1.75$. At follow-up, Carter continued to show improvements in narrative writing, with a mean maintenance score of 15.67 and an estimated maintenance effect size of $g = 1.65$. See Figure 3 for an example of Carter’s writing at follow-up. Zach also demonstrated maintenance of treatment effects, with an average writing sample composite score of 24.50 post-intervention and maintenance effect size of $g = 2.10$.

**Group 3: Emi and Sam**

As shown in Figure 1, Group 3 had a baseline length of nine school days during which seven samples were collected from Emi and three from Sam. During baseline, Emi’s average composite score was 2.43 and Sam’s baseline level of written performance was 3.67. Both participants’ composite scores increased with the introduction of the oral language instruction.
and there was no overlap between phases. The size of immediate effect varied by participant, but both students showed a positive level change. Emi’s mean score during the treatment phase was 16.17, with an estimated treatment effect size of $g = 1.66$. Sam’s average writing sample score during the treatment phase was 14.67 and the within-case treatment effect was estimated as $g = 4.31$ significant but small. One sample collected from Emi after withdrawal of treatment revealed a writing sample score of 19, slightly higher than the average composite score during the treatment phase. Emi’s estimated bias-corrected maintenance effect size was $g = 2.06$.

**Discussion**

The aim of our study was to examine the extent to which oral language instruction had immediate and lasting effects on kindergarten narrative writing skills. Oral narrative-based instruction has been shown to lead to improvements in writing for first grade students (Spencer & Petersen, 2018). Until this study, the impact of oral language instruction on writing development had yet to be extended to kindergarteners. Furthermore, most kindergarten writing research focuses on teaching transcription, with little explicit instruction on other aspects of writing until students first develop handwriting fluency (Graham et al., 1997; Puranik & Al Otaiba, 2012). The design and results of our study offer an alternative instructional direction, suggesting that kindergarten students can enhance their skills necessary for text generation before their transcription skills become proficient. The results from this study further support the schema theory related to narrative structure (Anderson, 1984; Mandler, 1984) as well as provide evidence that oral language is a significant contributor to writing quality (Kim & Schatschneider, 2017).

Our main finding was that brief, low-dose oral narrative instruction led to immediate improvements in kindergarteners’ writing that continued to improve as students received more
instruction. Prior to the narrative focused oral language instruction, the kindergarteners’ writing samples had an average composite score of 2.64. Although the immediacy of effect varied by student, the average composite score during the treatment phase was 18.5. The growth was specifically linked to greater inclusion of key story grammar elements in student writing samples, with an example shown in Figure 2. Our findings indicate that, when oral language instruction capitalizes on critical text structures such as story grammar, it is powerful enough to improve kindergarteners’ writing.

Our second main finding is that oral narrative language instruction had lasting effects on kindergarteners’ narrative writing. Samples collected several weeks after instruction ended revealed that gains maintained above baseline, suggesting that once the schema was acquired, it was firm in students’ repertoires. Because there was no explicit instruction provided to students regarding how to write a story, the durable effects on text generation are remarkable. It should be noted that follow up samples were not available for two participants. Therefore, we suggest further research is needed to substantiate the long-term impacts of oral narrative instruction.

Because there was little to no improvement observed on the Language Complexity scale, we concluded that oral language instruction was too weak to impact sentence structure in written form. This finding is not surprising since the language complexity features such as temporal and causal subordination were not explicitly taught or prompted during instruction like story grammar was—only modeled by the teacher when she read the stories. In an oral narrative intervention study, Spencer and Slocum (2010) discovered that preschoolers rarely used complex language features in their oral narrative retells unless the use of words like because, when, and after were explicitly prompted (Spencer et al., 2013; Weddle et al., 2016). The lack of complex language improvements in our participants’ writing may be because they were only emerging in
their oral language repertoires or that transfer to writing requires explicit instruction of complex sentences. Because we did not also obtain oral language samples in the current study, we were unable to explore this further. Therefore, future research should examine the effects of intentionally teaching and prompting complex sentences and the extent to which these features are transferable from oral to written modalities.

Although the NLM Flow Chart documented no significant growth in Language Complexity, we view the ability to measure various aspects of writing essential for informing instruction and our knowledge of how writing develops. Scientists and practitioners alike continue to be hindered by a lack of reliable, user-friendly, and comprehensive tools available to measure multiple dimensions of young students’ writing (Puranik et al., 2020). Because the NLM Flow Chart was designed to be used for oral and written language samples (Spencer et al., 2013; Spencer & Petersen, 2018), the focus of this tool was on the features that are shared across modalities. In this study, we did not measure transcription skills such as spelling because the teacher allowed students to ask how to spell words and because it was not logical that oral language instruction would increase spelling in just a few sessions. However, for writing research and practice to advance in primary grades, further developments in the assessment of writing are needed (Coker & Ritchey, 2014; Ritchey & Coker, 2014). To be able to inform instruction and serve as research outcome measures, there is a significant need for writing assessment tools to include metrics for multiple dimensions of writing (e.g., text structures, sentence structures, and transcription skills). The NLM Flow Chart can measure text and sentence structures in written form. A mean scoring agreement of 96% suggests that it can be used reliably to document a student’s present level of narrative writing performance, as well to track their development over time. Nonetheless, this assessment tool requires further
development and research. In addition to examining the sensitivity of the Language Complexity subscale, future research should include the scoring of students’ spelling, punctuation, and capitalization alongside text and sentence structures.

Limitations

The Story Champs oral language instruction shows potential for enhancing early learners’ writing quality as it relates to narrative structure. However, these findings must be interpreted with limitations in mind. For example, we developed the study design *a priori*, to include randomization of baseline lengths assigned to different groups. On pre-selected data collection dates, if students were absent from the classroom during scheduled writing periods, no other opportunity was provided to the student to produce a sample. Although it was important to maintain consistency in the methods for collecting writing samples, the natural consequences of conducting research in schools and classrooms with minimal disruption to the teacher’s routines limited the number of writing samples that were available.

In addition, an action research approach was used in the design of the study. Therefore, procedures did not include programming for an independent observer to measure, track, and respond to the teacher’s ability to implement Story Champs lessons with fidelity. Rather, keeping with action research traditions (Biancarosa & Snow, 2004), we relied on up front training and the teacher’s self-monitoring during each session to ensure manualized lesson implementation. Although a lack of independent observer is considered a limitation in most single-case research designs (Kratochwill et al., 2013), the results of the study suggest that the teacher utilized the training and checklists as the students’ writing outcomes improved markedly following the introduction of the oral narrative instruction. Ruling out chance, the effects were also visible across all three groups of students.
**Implications**

The statistically significant improvements in kindergarteners’ narrative writing abilities and the large effect sizes provide evidence of the power of explicit oral language instruction on the development of text generation in written form. During the intervention, the students were exposed to multiple exemplars of the targeted narrative schema through the introduction of unfamiliar stories each day. Thus, the focus of instruction was not mastery of any specific story (i.e., rote memorization) but procedural knowledge, such as story pattern recognition (e.g., story sequence and story grammar). Unexpectedly, a few students spontaneously and independently drew representations of the story grammar icons on their papers during writing activities (see Figure 3). The icons, whether used during oral language instruction or self-drawn during independent writing activities, prompted students to remember to include key story grammar elements in their narratives. We found this to be exciting evidence of schema activation while writing, even though the students had never been told to use the story grammar organization for writing. This suggests that the teaching methods worked as hypothesized. Mastery of the narrative schema during the oral language instruction is likely to have mediated students’ generalization from oral to written output (Anderson, 1984; Mandler, 1984).

Preparing students to become effective communicators across their lifespan requires dedicated time to writing instruction in the classroom. However, in a survey of Canadian K-6 elementary school teachers, only 57% of respondents reported explicitly teaching writing in their classrooms and less than half of those teachers used a specific program of instruction to support their students’ acquisition of handwriting skills (Makdissi et al., 2019). When early writing instruction focuses exclusively on visible and salient handwriting skills, students with delayed decoding and transcription skills may be left out of more complex language instruction needed
for reading comprehension and written composition. There is small but important causal evidence connecting oral narrative instruction to reading comprehension improvements (Clarke et al., 2010; Lervág et al., 2018), but research linking oral language instruction to writing outcomes is almost non-existent (Spencer & Petersen, 2018; Traga Philippakos, 2019). Nonetheless, this study provides additional evidence of a functional relation between oral language skills and written narrative skills. Given that the structure of oral narrative language maps to both reading and writing, the shared knowledge theory (Shanahan, 2006) posits that there is an efficient path from oral narrative instruction to writing composition and reading comprehension outcomes (Lervág et al., 2018; Shanahan, 2006; 2016). Therefore, we recommend that researchers and teachers use multi-faceted oral language instruction to facilitate the development of reading and writing simultaneously (Troia & Olinghouse, 2013) and that oral language instruction should not be delayed until students’ transcription skills are proficient. It is possible to build narrative text structures, needed for story reading and writing, before students can read and write. This is particularly critical for students with delayed or disabled decoding or transcription skills because without oral language instruction, they will have few opportunities to generate text and identify the narrative structures in the written language. Importantly, oral language instruction must go beyond vocabulary and grammar to teach powerful text generation skills such as text structures, perspective taking, and inferencing (Kim & Schatschneider, 2017), all of which are vital to young students’ development of oral and written narrative competencies.

References


**Table 1**

*Steps of Oral Narrative Instruction*

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Teacher displayed picture cards and read the model story. She laid icons down as she read each part of the story.</td>
</tr>
<tr>
<td>2</td>
<td>Teacher pointed to the icons and named the story grammar parts. Students repeated the story parts.</td>
</tr>
<tr>
<td>3</td>
<td>Teacher distributed one or two icons to each student. Students individually retold the part(s) of the story based on their icons.</td>
</tr>
<tr>
<td>4</td>
<td>Students retold the model story, one at a time. As students took their turns, visual materials were gradually withdrawn, first pictures then icons.</td>
</tr>
<tr>
<td>5</td>
<td>Teacher redistributed the icons. Based on the icons each student received, they collectively generated a novel story as the teacher drew stick figures on the white board.</td>
</tr>
</tbody>
</table>
Table 2

Participant Outcome Means, Standard Deviations and Within-Case Effect Size Estimates by Phase

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Treatment</th>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>Hedges’ g ES</td>
</tr>
<tr>
<td>Hank</td>
<td>3.33 (1.53)</td>
<td>23.25 (2.95)</td>
<td>2.00</td>
</tr>
<tr>
<td>Beth</td>
<td>3.00 (0.0)</td>
<td>22.75 (2.22)</td>
<td>2.26</td>
</tr>
<tr>
<td>Carter</td>
<td>1.75 (0.96)</td>
<td>16.89 (4.79)</td>
<td>1.75</td>
</tr>
<tr>
<td>Zach</td>
<td>1.67 (0.58)</td>
<td>17.25 (9.79)</td>
<td>1.41</td>
</tr>
<tr>
<td>Emi</td>
<td>2.43 (0.79)</td>
<td>16.17 (5.77)</td>
<td>1.66</td>
</tr>
<tr>
<td>Sam</td>
<td>3.67 (2.9)</td>
<td>14.67 (7.0)</td>
<td>4.31</td>
</tr>
</tbody>
</table>

*Note.* ES = Effect size; Within-case effect size (Hedges’ g) corrected for bias
Figure 1. Results of multiple baseline design across three groups
Figure 2. Student writing samples from baseline and intervention phases

Transcription:

“it raining in the ocean. I am a killer whale I am hungry. I’m going to go. catch a shark to eat.”

Transcription:

“A dolphin lived in a swimming pool. But he wanted to live with other dolphins and was sad so he walked on his fins to the ocean and found some friends.”
Figure 3. Post-intervention participant (Carter) writing sample.

Note. Participant spontaneously and independently drew story icons on their paper.
Transcription: “one day Zac was at my house and he was playing with my [cnroc]. could I play with it in 5 minutes.” On back of page (not shown), participant wrote: “ok. I was very happy.”
CHAPTER THREE:
EFFECTS OF A MULTITIERED ORAL NARRATIVE INTERVENTION ON FIRST GRADE LANGUAGE AND LITERACY SKILLS

What a child learns in their first few years of school can play a pivotal role in later academic, social-emotional, and employment outcomes. More specifically, reading and writing proficiency is correlated with adult quality of life indicators such as employment and salary, physical health, housing stability, and long-term relationships (Kern & Friedman, 2008; Lesnick et al., 2010; Smart et al., 2017; Snow et al., 2021). One of the earliest and strongest predictors of later reading and writing abilities is oral language (Hjetland et al., 2019; National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN], 2005). Oral language is a broad construct that encompasses meaning-related skills like semantics, vocabulary, syntax, linguistic structure, listening comprehension, and storytelling (Storch & Whitehurst, 2002; Tunmer & Chapman, 2012). Students’ development of oral language in preschool can predict their word reading, reading comprehension, and listening comprehension in elementary (Kim, 2016; Kim et al., 2018; Quinn et al., 2015; Lepola et al., 2012) and high school (Babayigit et al., 2021; Cunningham & Stanovich, 1997; Spencer et al., 2017a; Stothard et al., 1998; Suggate et al., 2018). Proficiency in these skills is critical to students’ use and understanding of more sophisticated and advanced academic language as they successfully progress through school.
It is important to identify students at risk for language and literacy problems and intervene early to prevent language and literacy problems from persisting into adulthood and limiting one’s opportunities for continued education and vocational advancement (Lyon, 1998; Pennington & Bishop, 2009). Yet, the number of young children identified as having specific learning disabilities (SLD) affecting language and literacy achievement may be underreported, as many children may struggle with literacy throughout their early academic years until they meet eligibility requirements for specialized intervention, often when they fail their third grade reading exams (Benson et al., 2020; Horowitz et al., 2017; Hwang & Cappella, 2018). On average, schools identify only 6% of students with a SLD in first grade, but in fourth grade the rate of identification of students as having a SLD increases to over 40% (Horowitz et al., 2017). Unfortunately, those who do not receive intervention to ameliorate problems by the end of second grade are more likely than their peers to experience persistent reading difficulties and worsening academic problems into fourth grade and beyond (Ferrer et al., 2015; Nation et al., 2010; Spira et al., 2005). When examining the 2019 National Assessment of Educational Progress (NAEP) reading assessment results (NCES, 2019), the degree to which students are experiencing language and literacy problems becomes more apparent—approximately 2 out of 3 U.S. fourth graders are not proficient in grade-level reading and writing expectations. This state of student reading failure requires a critical examination of current methods for teaching foundational literacy skills. Rather than assume students may have unidentified SLD, it may be possible that not all language and literacy skills are given equal instructional time and attention in the primary grades.

A student’s first grade year can be pivotal for their development of critical language and literacy skills. As of 2020, only 19 states and the District of Columbia mandate kindergarten
Grade 1 is some students’ first introduction to formal schooling and curricula that introduces basic reading, writing, and math skills (Education Commission of the States, 2020). Several longitudinal studies have found first grade academic and behavioral outcomes to be positively correlated with educational attainment and employment in adulthood (Entwisle et al., 2002; 2003; 2005). Examining developmental changes that occur within a child’s first grade year, other skills known to affect lifelong outcomes also start to emerge. For example, students demonstrate the largest degree of gains in reading and math in K-1 than they do in later grades, even after controlling for differences in ability at school entry (Burchinal et al., 2020). First graders also begin to develop and strengthen complex oral language skills like inferencing and understanding of text structures that are shown to be positively and significantly associated with other academic skills, as well as problem solving and theory of mind (Dawes et al., 2018; Zelazo & Carlson, 2020). Suggate and colleagues (2018) found that children’s oral language abilities in first grade were predictive of their reading comprehension skills in sixth grade. Given the importance of first grade in setting the trajectory for future academic success, it is imperative that students are provided with instruction in all skills critical to lifelong outcomes, including oral language (Clarke et al., 2010).

**Oral Language Intervention**

Young students’ ability to understand and use oral language to communicate effectively is key to their social and academic success. Despite its critical role in the development of reading and writing (Gersten et al., 2017; Gischar & Vesay, 2018; Sparapani et al., 2018; Uccelli & Snow, 2008), few early elementary curricula go beyond the explicit teaching of vocabulary to intentional programming in the other aspects of oral language. Most teachers use programs that target only the relationship between print and sound (e.g., alphabet knowledge, decoding,
phonics), with few programs available to support comprehensive instruction in the relationship between print and meaning, or language comprehension (e.g., Bowyer-Crane et al., 2008; Fien et al., 2015; Silverman et al., 2020; Tarvainen et al., 2020; Wilcox et al., 2020). By omitting more comprehensive oral language instruction and only targeting vocabulary skills or word-level reading, literacy curricula are less likely to have significant, meaningful, and lasting effects on students’ overall academic achievement (Chambers et al., 2016; Chiang et al., 2017; Levesque et al., 2018; Metsala et al., 2021). Rather, language instruction that targets the entire range of literacy skills including oral language will have the greatest impact on young students’ listening comprehension, vocabulary, reading comprehension, writing, science, and math skills.

Oral language interventions that include explicit vocabulary and listening comprehension instruction contribute to a variety of foundational literacy skills in early childhood. For example, vocabulary skills of 4- and 5-year-old children are positively and significantly related to their morphological awareness and pragmatic skills one year later, with the latter two skills independently predicting first graders’ oral narrative retell (listening comprehension) abilities (Ralli et al., 2021). There are also several large-scale studies reporting evidence of a causal relationship between oral language instruction, language comprehension, and vocabulary. For example, in two randomized controlled trials (RCTs) of the *Let’s Know!* oral language program, 30-minute supplementary lessons delivered by classroom teachers four times per week led to statistically significant improvements in first grade comprehension monitoring and vocabulary (LARRC et al., 2017; 2019).

Intervention research has also found causal evidence of an effect of oral language instruction on reading comprehension outcomes. For example, Westerveld and colleagues (2020) followed a cohort of first grade students for five years and examined the effects of a
supplementary oral language intervention delivered in a response to intervention model. The authors reported that over 90% of the students who received the oral language intervention in first grade were proficient in reading comprehension skills at second grade, compared to only 56.4% of students who had not participated in the oral language program.

In addition to the improvement of reading comprehension, oral language interventions can also improve students’ argumentative and procedural writing skills. In two RCTs, first grade students who experienced interventions with embedded oral language instruction significantly outperformed students in comparison or control groups. In the first study, Traga Phillippakos (2019) taught first grade teachers to implement a nine-day writing intervention with embedded oral discourse and strategy instruction. Traga Phillippakos (2019) reported statistically significant improvements in the quality of procedural writing for students in the treatment group. Kim et al. (2021) found that a language and literacy intervention delivered during science class had positive and statistically significant effects on students’ domain-specific vocabulary, listening comprehension, reading comprehension, and argumentative writing skills.

There is growing evidence that embedding oral language instruction within domain-specific curricula has strong positive effects on a variety of student outcomes, like math skills. Oral language interventions that address text comprehension are also shown to have positive and significant effects on students’ development of math word problem solving in first (Fuchs et al., 2021) and second grade (Fuchs et al., 2018). For both second grade boys and girls, oral language skills significantly predicted math word problem solving in fourth grade, $\beta = .351$, $SE = .056$, $p < .001$ (Spencer et al., 2020a). Further, meta-analysts have found significant and positive relationships between math and language comprehension development, across age groups and populations of monolingual and second language learners (Arizmendi et al., 2021; Peng et al.,
2020). This body of research suggests that interventions targeting skills like vocabulary, inferencing, and text structure may be essential in developing a strong foundation in science and math content areas.

**Narrative Intervention**

Oral language interventions can vary in format, to include but not limited to dialogic reading, fictional narrative or expository retells, and group discussion about text read (e.g., summarizing, evaluating, questioning). Narrative intervention is one form of oral language instruction that teachers and speech-language pathologists (SLPs) use to promote a multitude of language and literacy skills. The key ingredient of narrative interventions is the use of oral storytelling or retelling activities to intentionally practice integrated word, sentence, and discourse level language targets (Spencer & Petersen, 2020). The aim of such interventions can be on developing students’ language comprehension (e.g., listening comprehension, vocabulary, inferencing), language production (i.e., written, spoken), or both.

There is a growing evidence base of the effects for oral narrative language instruction on first graders’ language skills. For example, Paris and Paris (2007) studied the effects of direct narrative strategy instruction using wordless picture books and reported significant student gains in taught narrative comprehension (e.g., story elements, inferencing) and untaught listening comprehension skills (e.g., predictions, story questions). Baumann and Bergeron (1993) taught story mapping to first grade students, finding the brief intervention to have positive and significant effects on their oral recall of key story grammar. In a quasi-experimental study of a brief 10-week oral narrative intervention, Wright and Dunsmuir (2019) also found statistically significant positive effects on first grade oral retells.
Slightly more evidence is available to evaluate narrative-focused language intervention effects on first grade text generation skills (Coker et al., 2018; Spencer & Petersen, 2018). When students learn background knowledge, language structure, and vocabulary through oral language interventions, they can produce higher-quality written narratives (Graham et al., 2021; Graham & Hebert, 2010). For example, after controlling for students’ handwriting and spelling abilities, Coker and colleagues (2018) found generative narrative writing tasks coupled with language composition instruction lead to statistically significant improvements on first graders’ reading achievement. Using a multiple baseline across participants design, Spencer and Petersen (2018) investigated the effects of a brief small group oral narrative intervention on first graders’ narrative writing and reported immediate and sustained improvements. The promise of narrative intervention for enhancing first grade text generation skills warrants further research. However, overall, these studies demonstrate the effectiveness of narrative-focused oral language interventions on first graders’ language and literacy skills.

**Multitiered Systems of Language Support (MTSLS)**

Given the importance of addressing language difficulties in first grade to prevent later reading and writing problems, oral language intervention used within a multitiered system of supports (MTSS) model may reach the largest number of students in the shortest amount of time. In the past two decades, states have adopted educational policies that encourage the use of MTSS to organize differentiated instruction for diverse students that aims to prevent the worsening of academic difficulties.

Many schools routinely use tiered instruction to teach elementary reading and writing skills (Schiller et al., 2020), but the oral language aspects of literacy are not systematically executed within MTSS. To emphasize the need for early and intentional oral language instruction
within an MTSS framework, Petersen et al. (2022) promoted and examined the effect of a Multitiered System of Language Supports (MTSLS) on critical language-related outcomes of young students. The MTSLS framework aligns with recommended national and state reading policies that mandate universal screening to identify at-risk students, frequent progress monitoring, and data-based problem solving that result from effective partnerships between families and schools (Abbott & Wills, 2012; Mahdavi & Tensfeldt, 2013). When school teams use the MTSLS model, all students receive evidence-based oral language instruction at Tier 1, traditionally delivered by general education classroom teachers. Based on regular benchmarking, students who do not make sufficient progress at Tier 1 receive supplementary oral language intervention at Tier 2 (small group). Students receiving intervention participate in ongoing progress monitoring. If necessary and according to school resources, students can be referred for individualized and intensive intervention at Tier 3.

An important component of MTSLS is the involvement of language experts, who are usually school-based SLPs. In Petersen et al. (2022), SLPs provided Tier 2 small group intervention to young students with oral language risk in addition to students with language disabilities and provided ongoing consultation to teachers regarding Tier 1 oral language instruction. SLP-delivered supplemental language interventions during the early primary grades help to prevent language difficulties from developing into reading and writing difficulties such that a SLD eligibility is warranted in later primary grades. Given that MTSLS is a framework that demands collaboration between teachers and SLPs, the oral language intervention approach should be suitable for whole class and small group delivery and focus on the most critical skills for literacy development.

**MTSLS and Narrative Intervention**
A number of studies have examined the implementation of oral language interventions in a tiered fashion (Zucker et al., 2013, 2021), but the use of narrative interventions is among the most common approaches to MTSLS (Nelson et al., 2022; Petersen et al., 2020; 2022; Peterson et al., 2020; Spencer et al., 2017b; Spencer et al., 2020b; Weddle et al., 2016). In a recent meta-analysis of interventions reporting effects on school-aged children’s narrative language by Pico and colleagues (2021), the researchers found eight interventions delivered in whole-group settings. However, more narrative interventions have been delivered in small group or individualized arrangements (e.g., Brown et al., 2014; Favot et al., 2021; Glisson et al., 2019) and only a few involved multitiered implementations (e.g., Garzarek et al., 2019; Gillam et al., 2014).

One narrative intervention program was specifically designed for MTSLS. Story Champs® (Spencer & Petersen, 2016) is a multitiered oral narrative language curriculum that uses carefully constructed stories and illustrations to teach complex academic language to students through storytelling activities. Story Champs lessons come in three formats for implementation in whole group, small group, or individual instructional settings and can be used for instruction within mainstream classrooms or intervention for students at risk for or having language disabilities. Story Champs is designed so that anyone trained in the basic procedures can implement lessons in any format. With minimal training, the program can be implemented by literacy specialists, reading coaches, classroom teachers, instructional assistants, clinicians (e.g., SLPs, behavior analysts), undergraduate students, parents, and researchers. To deliver lessons, users are provided with a set of child-relatable narratives that are leveled based on their inclusion of complex language (e.g., and number of structural elements). At a minimum, each story includes five structural elements (i.e., character, problem, feeling, action, ending), but the
first-grade stories also include settings and end feelings. Stories are purposefully written to include complex linguistic features like advanced vocabulary, modifiers, causal ties, and temporal ties.

A typical lesson for improving students’ oral narrative retell skills follows the basic principles of effective instruction, starting with the adult presenting a series of illustrations and reading aloud a story script while simultaneously gesturing to and/or using icons to highlight key parts of the story as they read. Procedures require the transfer of stimulus control away from visual and verbal prompts so that by the end of the lesson, students should be able to independently and accurately retell the story without needing illustrations or icons. See Spencer et al., 2015 for a more detailed description of Story Champs procedures for whole class implementation.

In numerous experimental and quasi-experimental studies, the efficacy of Story Champs has been examined in elementary (Petersen et al., 2020; 2022) and preschool settings (Spencer & Slocum, 2010; Spencer et al., 2013; 2014), at all tiers independently and combined (Spencer et al., 2018), with neurotypical children (Petersen et al., 2020; 2022; Spencer & Petersen, 2018) as well as children with disabilities (Hessling & Schuele, 2020; Petersen et al., 2014) or children at risk of academic difficulties (Spencer et al., 2015; 2020). In addition to the numerous English language outcomes observed such as listening comprehension, personal story generation, story writing, and reading comprehension, a dual language version of Story Champs, delivered at Tier 1 (large group), Tier 2 (small group), and Tier 3 (individual) improved both Spanish and English oral language outcomes of preschool dual language learners (DLLs). Moreover, Story Champs has been examined in at least three international studies, one in Mexico (Gutierrez Arvizu &
Not all the previous research investigating the effects of *Story Champs* have involved end users—that is, teachers and SLPs. However, there is value in conducting research with the individuals who would use the program in typical practice, especially when efficacy research is already available. As noted in a call-to-action for language intervention researchers by Curran and colleagues (2022), studies should be designed and conducted in a manner that allows for easy adaptation and efficient uptake by practitioners like SLPs based on their knowledge of client needs and/or context of delivery. Interventions having positive outcomes under strict lab-like conditions are less likely to be effective in sometimes chaotic and messy real-world settings. As a result, clinicians and teachers tend to trust outcomes when the primary ingredients for effectiveness are identified while allowing for adaptation of flexible ingredients that better fit with their own practice setting and experience (Goldberg et al., 2018; Institutes of Medicine, 2001; Kim, 2019). Thus, methodological flexibility and research partnerships with community members are essential for researchers to move away from sample-specific studies of effect toward the study of effective interventions applied in broader contexts with more diverse populations.

Three *Story Champs* studies are particularly relevant to the current investigation, in that they featured a multitiered delivery of the intervention by end users. Spencer et al. (2018) studied the effectiveness of *Story Champs* on the oral language skills of preschoolers attending a Head Start program. The researchers were interested in the effectiveness of the multitiered intervention when implemented in Head Start classrooms by teachers and classroom assistants. In comparison to students who did not receive the *Story Champs* intervention, preschoolers in the classrooms
assigned to the treatment condition demonstrated greater narrative retell and listening comprehension skills and the group differences were statistically significant. Medium effect sizes were noted for all winter and spring measures (partial $\eta^2$ range: .06–.08) Furthermore, the staff who implemented the intervention reported high intervention feasibility after several months of implementation.

In a recent pragmatic trial with 686 kindergarteners, Petersen et al. (2022) randomly assigned 28 classrooms to treatment or control groups. Story Champs was delivered in an MTSLS framework and several academic language outcomes were measured, including personal story generation, narrative and expository retells, and writing. Teachers and school-based SLPs assigned to the treatment group implemented the multitiered oral narrative language intervention for one year. Kindergarteners who did not immediately respond to whole group Tier 1 instruction received supplemental small-group Tier 2 intervention from SLPs. At the end of the year, students in the treatment group showed statistically significant improvements in all outcomes compared to their peers who received traditional language arts instruction (ES = 0.19 to 0.49). Further investigation of effects at the Tier 2 level revealed that the intervention led to substantial and statistically significant improvements in kindergarteners’ expository retells, personal story generation, narrative retells when compared to matched at-risk peers in the control group and peers who performed at or above grade level expectations (smallest ES = 0.35).

In an additional study of an MTSLS delivery of Story Champs, Petersen and colleagues (2020) included classroom teachers (Tier 1) and school-based SLPs in the delivery of small group (Tier 2) instruction to students in the second grade. Using the MTSLS model, the SLPs delivered small group intervention to students in the treatment group if they did not demonstrate progress in response to whole-class instruction alone. In comparison to students who did not
receive the multitiered narrative intervention, students in the treatment condition demonstrated significantly larger gains in reading comprehension, writing, and narrative retell skills from pretest to posttest. Four students in the treatment group were identified as needing supplemental Tier 2 intervention, demonstrating the importance of including SLPs in the prevention of language and literacy problems.

Although the evidence base for multitiered oral narrative intervention is promising and calls for multitiered models for the promotion of oral language are increasing (Meaux et al., 2020; Sylvan, 2018), there is a gap in the literature at first grade. To the best of our knowledge, the investigations of oral narrative intervention delivered in a multitiered fashion have not included first-grade participants. Most of the research examining multitiered oral narrative interventions have been conducted with students in preschool (Spencer et al., 2015, 2018, 2020; Weddle et al., 2016), kindergarten (Brown et al., 2014; Coyne et al., 2007; Petersen et al., 2022), or grades 2 and 3 (Hessling & Schuele; Lee, 2020; Nelson et al., 2022; Petersen et al., 2020).

The goal of the present study is to examine the immediate and long-term effects of an oral narrative language intervention (i.e., Story Champs Curriculum) delivered in an MTSLS framework on first graders’ language and literacy skills. We hypothesized that implementation of the MTSLS model would help identify students in need of more intensive and individualized levels of language instruction and provide necessary infrastructure to organize the oral narrative language instruction and intervention. The following research questions were addressed in this study:

1. Compared to standard (business-as-usual) instruction, does multitiered narrative-based language instruction increase the vocabulary, listening comprehension, reading comprehension, and writing skills of first grade students with language disabilities or
language delays when implemented by speech and language pathologists and teachers?

2. To what extent do the multitiered oral narrative intervention effects, if any, maintain after 5-6 months in absence of the intervention and summer language instruction?

3. How much of the variance in students’ language and literacy outcomes across is attributed to students’ language skills at baseline, examiner fidelity, and school-level factors like adherence to MTSSL?

Method

The genesis of this study emerged from a decades-strong research-practice partnership (RPP; see Alonzo et al., in press; Coburn et al., 2013; Coburn & Penuel, 2016) between university researchers, university-based clinical instructors, and school district program administrators. One benefit of RPPs is that they allow for research to occur organically, with the community partner bringing problems or needs to the attention of researchers so they can assist in the creation and/or implementation of solutions. From the outset, all stakeholders in this RPP contributedmeaningfully to the design, implementation, and interpretation of results. Rather than evaluate intervention effects under tight researcher-controlled conditions, we examined its effects when implemented by its intended end-users. In the following section, we describe the context of intervention delivery and the involvement of all RPP stakeholders.

Context

In conjunction with district initiatives, each of the 10 participating schools implemented the key ingredients of MTSSL (i.e., universal screening and progress monitoring, and monthly team meetings and data-based decision making). In the current study, we limit the independent
variable to the multitiered oral narrative instruction and intervention and examine their effects on students’ language and literacy performance.

**Study Design**

We used a cluster randomized waitlist control group design to study the effects of the multitiered oral narrative intervention delivered to first graders at risk for or with language disorders. We used this design because it offers the strongest possible controls for threats to internal validity (Hemming et al., 2017). Furthermore, the design is suitable for examining an intervention with student data nested within schools. Since MTSS activities were implemented school-wide, we randomly assigned five schools to the treatment condition and five schools to the waitlist control condition. In schools assigned to the waitlist control condition, first grade general education teachers engaged in literacy instruction as-usual, meaning that they used the English Language Arts (ELA) curriculum already in-place. In schools assigned to the treatment group, a portion of their ELA block was replaced with Story Champs (Spencer & Petersen, 2016). First grade teachers implemented the Story Champs oral language instruction to all students in their classrooms (Tier 1) while a portion of students received Story Champs supplemental oral language intervention (Tier 2) from SLPs. For all 10 schools, we measured outcomes at three time points: pretest (fall 2018), posttest (spring 2019), and follow-up (fall 2019).

**Setting and Participants**

**Schools.** We conducted the study in 10 elementary schools from two districts in the upper-Midwest region of the United States. A primary investigator presented information about the study in meeting with all the district principals and central office administrators. Each principal had the opportunity to volunteer their school to be included in the study. All
administrators, leadership teams, first grade teachers, and SLPs from each school expressed interest in participating.

The two school districts were located within the metropolitan suburbs of a large Midwestern city. In the fall of 2018, District 1 ($k = 6$ schools) had a total enrollment of 2,558 primary grade students across a geographically diverse suburban area, with 39% of students that qualified for free or reduced-cost lunch. The racial composition of District 1 was 76% White, 13% African American, 3% Hispanic, and 2% Asian. In comparison, District 2 ($k = 4$ schools) had a total enrollment of 4,200 students from a mixed urban/suburban area. In District 2, 41% of students were eligible for free or reduced-cost lunch. The district reported student racial demographics as 77% White, 12.9% African American, and 3.5% Hispanic.

**Interventionists.** Prior to the start of the study, each elementary school building had one SLP dedicated to serving the students at that school. In addition, resulting from an ongoing partnership with a local university speech-language pathology graduate training program, every building had at least one practicum student who was assigned to work alongside the school-based SLP. The SLPs included the students in administration of the language assessments because they considered it to be a natural and valuable real-world experience. However, each SLP was ultimately responsible for collecting the student measures, as the students were not mandated to assist with study procedures as a part of their university practicum course. Therefore, we did not consider them members of the research team. In total, 26 teachers and SLPs participated in the implementation of MTSLS and completed demographic surveys so that we could describe them. As shown in Table 4, all teachers and SLPs were White and female. More than half of the teachers and SLPs held master’s degrees (65.4%) and 27% held bachelor’s degrees. First grade
teachers’ experience ranged from 8–32 years. On average, the SLPs had 11 years of experience (range: 1–23).

**Students.** Student eligibility to participate in the study was dependent on the results of universal screening and identification of language disabilities by the school teams.

**Screening.** Schools conducted universal screening of all first graders in the Fall of 2018 using the Narrative Language Measures (NLM) Listening subtest of the CUBED assessment (Petersen & Spencer, 2016). Based on the fall benchmark cut score of 14 reported in the CUBED manual, teachers and SLPs created a list of first graders with language-related disabilities (i.e., existing IEP and scored 14 or below) or considered at risk for language-related disabilities. All students, regardless of an existing special education classification, who scored between 0–14 were invited to participate. Teachers and SLPs contacted parents and guardians of the students regarding permission for them to participate and the researchers were available to answer questions, as needed. Although the schools included all first-grade students in their MTSLS activities, only students meeting the MTSLS screening criterion and receiving signed parental permission served as research participants. The local university institutional review board (IRB) approved recruitment, enrollment, and ongoing research activities related to the collection of student outcomes. Researchers complied with all applicable federal, state, and local standards related to the ethical treatment of human subjects.

**Demographics.** In Table 3, we report demographics for all student participants based on a survey parents and guardians completed at the time of consent. A total of 155 first grade students participated in the study which included 77 males and 78 females, with an average of 15 student participants per school. At pretest, the average student was 6 years; 5 months old (range: 5;8–7;9). In total, 13 students had an IEP and school teams classified 139 students at risk for
language and learning disabilities (classification data were missing for three students). All
students spoke fluent English and most of the students came from monolingual English-speaking
households. However, three students spoke both English and another language at home and two
students spoke only Spanish at home. Participants were predominantly non-Hispanic White
(74.2%). Fifteen students were multiracial, 14 were African American, three were Asian
American and two were Latino/a. Six parents declined to provide information about their child’s
race and ethnicity.

**General Procedures**

**Multitiered Oral Narrative Intervention.** We investigated a non-commercialized
version of *Story Champs* in this study (i.e., *Story Champs Curriculum*). More than 10 years of
experimental and quasi-experimental intervention studies informed the development of *Story
Champs Curriculum* (e.g., Petersen et al., 2022; Petersen et al., 2020; Spencer et al., 2020;
Spencer et al., 2017b; Spencer et al., 2018; Spencer & Slocum, 2010) with the intent to improve
the implementation fidelity of comprehensive academic language instruction. As a result, the
primary difference between *Story Champs* and *Story Champs Curriculum* is the explicitness of
the lessons and detailed scope and sequence. With more scripted directions for the user, *Story
Champs Curriculum* offers teachers and interventionists greater specificity for lesson delivery as
well as how to build and integrate complex language across the year. Curriculum materials
comprised a spiral-bound lesson book, a spiral-bound picture book containing photos, a set of
story illustrations, story grammar icons, vocabulary posters, peer-tutoring checklists, and a USB
drive storing reproducible materials such as writing graphic organizers and word journals
available to print as needed.
The *Story Champs Curriculum* is organized into five strands: 1) story retell, 2) story writing, 3) information retell, 4) information writing, and 5) flex for differentiation. The first two strands (strands 1 and 2) feature stories with themes considered familiar to most U.S. first graders (e.g., falling off a bicycle, sharing toys, etc.) and are approximately 135–150 words in length. The stories include the canonical story grammar (i.e., character, setting, problem, feeling, plan, action, consequence, ending, and end feeling) and complex sentences (e.g., causal and temporal subordination, modifiers) that increase in variety and number across the year. Each story contains two general academic vocabulary words targeted during explicit instruction and practice within the lesson. Context clues are embedded in every story to facilitate guided practice of inferring word meanings from context.

Informational passages in strands 3 and 4 align with first grade science and social studies topics such as habitats, layers of the earth, and ancient Egypt (Core Knowledge Foundation, www.coreknowledge.org). Containing increasingly complex sentences (e.g., elaborated noun phrases, relative clauses, and transition words) similar to the stories, each informational passage is 125–135 words long and consists of two paragraphs, each with a main idea and four key details. Two general academic words and corresponding context clues are embedded in each passage. In addition, each informational passage contains four domain-specific vocabulary words related to the social studies or science topic.

*Story Champs Curriculum* lessons are scripted and manualized based on extensive literature on explicit instruction teaching procedures (Archer & Hughes, 2011) and contextualized language intervention (Ukrainetz, 2006), but there is also built-in flexibility and multiple ways to differentiate. Lessons for strands 1 and 3 contain six scripted activities; lessons for strands 2 and 4 contain nine scripted activities; the flex lesson contains four choices of
activities. Activities include teacher-student dialogue for teaching story/passage structures, vocabulary, and complex sentences. Initially, scripts help educators learn the instructional formats (e.g., model, lead, try) and efficient prompting and error correction techniques (e.g., first ask a question, then model).

*Story Champs Curriculum* for first graders includes 45 lessons distributed across 15 weeks, depending on the frequency of use. Each of the five strands contains a set of 15 lessons. General education teachers delivered lessons for strands 1–4 in order, every day (e.g., Monday through Thursday). On the fifth day (e.g., Friday), teachers chose an activity from strand 5, which was a flex lesson for differentiation. Given the option, teachers delivered the selected flex activity to the entire class or broke students into small groups and assigned a different flex activity to each group. For students needing more targeted language practice, SLPs or instructional assistants supervised by the SLP delivered the Tier 2 intervention using *Story Champs Complete Curriculum* two to three times a week. However, for Tier 2 intervention, the SLPs chose which strands to focus on with their students and determined how long students received Tier 2 intervention. This afforded them flexibility and the ability to individualize based on ongoing student progress monitoring data. Each lesson, regardless of arrangement, took approximately 30 minutes to complete.

**Training and Coaching.** Researchers trained all school assessment teams to administer and score the Narrative Language Measures (NLM) Listening and Reading for screening and progress monitoring purposes as part of their schools’ MTSLS. In addition, they trained the SLPs and speech and language pathology graduate research assistants to administer all standardized assessments. The research assistants demonstrated > 90% assessment administration fidelity and > 85% scoring reliability prior to engaging in study activities.
After completing participant screening, selection, pretesting, and random assignment phases, the Principal Investigators (PIs) delivered a one-day training on the multitiered language curriculum to all teachers, interventionists, and SLPs assigned to the treatment group. Training content included: the importance of oral language for academic skills; essential components of MTSLS; lesson format, collaboration, and delivery; data-based decision making; leadership team meetings; and instructional routines. Training methods included used didactic instruction, use of video/live model demonstrations of procedures, and opportunities for hands-on practice. The researchers observed each trainee in the delivery of a lesson, but due to the large number of teachers, SLPs, and interventionists, they did not use a fidelity check out system. Nonetheless, the Story Champs Curriculum was designed to maximize fidelity of lesson delivery with its scripted format. Story Champs Curriculum implementers were given a log to track dose and delivery of lessons.

Measures

School-based teams administered a series of assessments to the first-grade students at pretest, posttest, and follow-up. The third author (a language and literacy consultant employed by the school districts) was responsible for collecting and providing the first two authors with students’ original writing samples, audio recorders containing all student language samples, and original student record forms after each time point. Measures reflected a range of outcomes, from distal (i.e., Woodcock Reading Mastery Tests Third Edition, WRMT™-III; Woodcock, 2011) to more direct assessments of intervention effects like student oral and written language samples. We describe the measures and their procedures in detail below.

Proximal Measures. We used the NLM Listening subtest of the CUBED assessment (Petersen & Spencer, 2016; See Appendix B) as a proximal measure of students’ listening.
comprehension. School teams followed the standardized administration procedures to measure students’ narrative retell and vocabulary inferencing skills using a digital administration platform. The NLM Listening measure is organized into four sections: Listening Retell, Story Questions, Vocabulary Questions, and Personal Story Generation (optional). For this study, only the Listening Retell and Vocabulary Questions sections were completed. At pretest, posttest, and follow-up, each student was administered two stories with two opportunities for story retells and to answer the vocabulary questions. Administration took approximately 2–3 minutes to complete per student.

**Narrative Retell.** The Retell section of the NLM Listening requires students to listen to a brief narrative read aloud and then retell that narrative. Following the standardized administration procedures, examiners brought an individual child to a quiet location and said, “I’m going to read you a story. Please listen carefully. When I’m done, you are going to tell me the same story. Are you ready?” Examiners read a fictional (but realistic) story to the student aloud and with normal pace and inflection, then said, “Thanks for listening. Now you tell me that story.” As students retold the story, examiners scored their retell in real time. On a scale of 0–2, according to their inclusion, completeness, and clarity, examiners awarded points for each of the key story grammar elements (i.e., character, setting, problem, feeling, plan/attempt, consequence, ending). Scores of 0 indicates no inclusion of a particular component and a score of 2 indicates clear and complete use of the element in the retell. In addition to the story grammar, examiners awarded one point for each use of words that mark complex sentences (i.e., because, when, so that, after), with a total of three points for each word. Finally, bonus points were awarded according to how many episodic story grammar elements (i.e., problem, attempt, consequence,
and ending) received two points. The examiner added all points earned for story grammar, language complexity, and episode to report a total narrative retell score.

**Vocabulary Inferencing.** The Vocabulary Questions section of the NLM Listening was used to assess students’ ability to infer word meanings when provided context clues. Following the student retell, examiners administered the Vocabulary Questions section and scored student responses in real time. To do this, they read aloud a sentence that included one of the general academic vocabulary words embedded in the story and its related context clue. For example, “Scott kept tumbling off his skateboard. He kept getting hurt. What does tumble mean?” (CUBED manual; Petersen & Spencer, 2016, p. 43). There was a total of three Vocabulary Questions per NLM Listening story, each scored on a scale of 0–3. If a student was able to provide a definition upon request, their response was given a score of 3 (correct, complete and clear) or 2 (correct, but unclear or incomplete). If a student was unable to provide a definition, the examiner asked the student to identify the meaning of the word given two choices, one of which was the correct response. The student’s selection of the correct definition was assigned 1 point. If the student did not respond correctly or at all, the question was scored as 0. The total possible score range for the Vocabulary Questions section was 0–9.

**Intermediate Measures.** Students’ expository oral retell and narrative writing language samples were considered intermediate measures of student literacy skills.

**Expository Oral Retell.** Most often associated with disciplinary content (e.g., biology, history), expository language is considered an advanced academic language register (Cummins, 2014; Nippold et al., 2005). Given that students were exposed to information retelling in the *Story Champs Curriculum*, but the CUBED does not include a proximal measure akin to the NLM Listening, we elicited expository oral language samples that could reflect an intermediate
measure of intervention effect on students’ language comprehension outcomes. We provided school teams with a standardized elicitation script and procedural checklist, to include instructions for audio recording their elicitation and the students’ retell. The general procedures for collecting expository oral retell language samples included: 1) the presentation of four sets of photographs depicting different science or social studies-based information, 2) asking students to select one set of pictures to talk about, 3) reading an informational passage corresponding to the student’s selection, and 4) asking the students to retell the same information. SLPs elicited two expository oral retell samples from students at pretest, posttest, and follow-up. The SLPs used an audio recorder to save the oral language samples and sent the devices to the university research team for transcription and scoring.

To score the expository oral retell samples, we used a newly developed academic language discourse analysis tool called the Expository Language Measures (ELM) Flow Chart (Language Dynamics Group, 2020; See Appendix C). The ELM Flow Chart contains three subsections: writing conventions (optional; for written samples), language complexity, and passage structure. The six-item language complexity subsection evaluates students’ use of relative pronouns, modifiers, less-common content-specific vocabulary, and transition words. Five of the six items have a possible score range of 0–3 and transitions assigned scores between 0–4. The passage structure section of the ELM Flow Chart uses a decision-tree format to assign quality ratings to six items representing complete and coherent informational retells: main idea (score range: 0–3), information units (0–20), definitions and examples (0–3), passage cohesion (0–3), concluding statement (0–2), and exposition type (0–2). For this study, we chose to omit the item “Exposition Type” as it did not relate to the study aims and only provides categorical
information. We summed the total scores for both language complexity and passage structure subsections to report a single expository oral retell composite score ranging from 0 to 47.

**Narrative Writing.** At each time point, students wrote a fictional narrative in response to a set of selected photos. Examiners followed a standardized elicitation script to collect one narrative writing sample from each student. This involved the presentation of four sets of three photographs depicting different storylines and asking students to select the set of photographs they wanted to write about. The examiner removed the photos the student did not select and directed the student to write a story about the pictures. Examiners did not assist students with transcription or text generation but provided neutral encouragement as needed. School teams collected students’ writing samples and submitted them to the research team for scoring.

We used the Narrative Language Measures (NLM) Flow Chart (Language Dynamics Group, 2020; See Appendix A) to score the students’ writing samples ($N = 309$). The NLM Flow Chart includes three subscales: narrative structure, language complexity, and writing conventions. Using the flow chart decision tree, each language complexity subscale item was given a score between 0–3, and narrative structure items varied with a minimum range of 0–2 (sequence, ending), 0–3 (character, setting, emotion), 0-4 (problem, plan/attempt, consequence), and 0–8 for episode complexity. For writing outcomes, we reported each student’s total NLM Flow Chart score (range of total possible score: 0–59).

**Distal Measures.** The Listening Comprehension (WRMT LC) and Passage Comprehension (WRMT PC) subtests of the Woodcock Reading Mastery Test III (WRMT-III; Woodcock, 2011) served as distal outcome measures. To administer the LC subtest, examiners read aloud informational and narrative passages of increasing difficulty, asking students to answer direct questions or make inferences about the information or story they just heard. The
PC subtest requires a student to independently read and complete a brief sentence or passage with a missing word. We used the raw scores reported at each time point as an indirect and distal measure of intervention effect on students’ listening and reading comprehension.

**School-level Outcome Measures.** To address the second research question about moderating variables, we collected demographic information from students, teachers, and SLPs involved in the study. To provide a measure of MTSLS implementation fidelity at the school level, we used an existing elementary version of the R-TFI (St. Martin et al., 2015), designed to assist school teams in evaluating their progress toward full implementation of MTSS for reading. We made minor revisions to the form to add language to the inventory of items. School leadership teams used the form to rate their MTSLS implementation on a 0–2 scale for 27 Tier 1 features, 14 Tier 2 features, and 11 Tier 3 features. School leadership teams in both treatment and control groups completed the revised R-TFI in November 2018 and April 2019. We used R-TFI scores at posttest to reflect school-level MTSLS implementation.

We also measured the degree to which the school-based SLPs and speech-pathology practicum students already present within the schools elicited the students’ language samples according to standardized procedures. For both expository oral retell and written narrative outcomes, the first author trained undergraduate research assistants to listen to all audio recordings ($N = 895$) of language sample collection and use a checklist to record the number of administration steps to which the examiner adhered. Thirty percent of the audio files were selected for secondary scoring (expository oral retell: $n = 46$; narrative writing: $n = 65$). We used Cohen’s kappa to evaluate the degree of agreement between the two raters, resulting in 0.997 for expository oral retell samples and 1.0 for written narrative samples.
Transcription Fidelity and Reliability. The first author trained three undergraduate students to assist with word-for-word transcription of expository oral retell language samples and interpretation and transcription of narrative writing samples. The first author resolved any disagreements in the interpretation of students’ speech or handwriting. The first author randomly selected 20% of the expository retell ($n = 88$) and narrative writing ($n = 92$) samples per time point per school for transcription fidelity and reliability scoring. A trained independent research assistant blind to study aims and group assignment used a checklist to record the degree to which transcription procedures were followed and samples were transcribed verbatim. Average percent of transcription fidelity across expository oral retell samples was 95.2% (range: 80–100%) and word-for-word transcription reliability was 97.5% (range: 80–100%) with an average rate of disagreement per audio sample of 0.02. For narrative writing, average transcription procedural fidelity was 98% (range: 82–100%) and word-for-word transcription reliability was 96.8% (range: 82–100%) with an average rate of disagreement per sample of 0.03.

Scoring Reliability. Scoring reliability was calculated for all language sample outcomes.

Expository Oral Retell. The PI trained the first author and an independent research assistant who was unaware of group assignment in expository oral retell sample scoring procedures using the ELM Flow Chart. The first author completed primary scoring while blind to study condition. After primary scoring of all samples was complete, the independent research assistant completed secondary scoring for approximately 20% of the samples ($N = 92$) to assess item-by-item scoring reliability. The primary and secondary scorer met once to reconcile any differences in scores to the extent possible. If consensus was not reached, the primary score was used for reporting outcomes. The average item-by-item agreement was 86.5% (range: 72.7%–
The average difference between two conflicting scores (item level) was 1.09 (range: 1–5).

**Narrative Writing.** The first author was trained to fidelity in use of the NLM Flow Chart by the PI and served as the primary scorer. An independent research assistant trained by the PI and blind to group assignment conducted secondary scoring for approximately 20% of each school’s samples at pretest, posttest, and follow-up ($N = 90$). We compared the two raters’ sample scores at the item level to calculate percent agreement, or item-by-item scoring reliability, which was 92.9% (range: 77.8%–100%). The average difference between two conflicting scores (item level) was 0.91 (range: 0.0–3.0).

**Data Preparation and Analysis**

The school-level study teams collected all student outcome data and shared it with the university researchers. All samples, to include audio recordings and written products, were saved to a restricted university-housed server. All student products were de-identified, replacing students’ names with alpha-numeric study identifiers. We used Microsoft Word for all transcriptions. We scanned and saved all student writing samples (.pdf) by student identification number, as well. Demographic and quantitative data (scored outcomes) were prepared using Microsoft Excel and analyzed using SAS (v. 9.4).

**Examination of ICCs.** To examine the degree of variance in outcomes associated with student data nested in schools, we calculated the intraclass correlation (ICC) for all pretest variables using an unconditional model and calculated ICCs for posttest outcomes conditional on pretest scores. ICCs varied widely at pretest: $<.0001$ (WRMT PC), .011 (narrative retell), .027 (expository retell), .037 (vocabulary inferencing), .126 (narrative writing), and .995 (WRMT LC). Conditional on pretest scores, posttest ICCs by variable were .091 (expository retell), .134
(WRMT PC), .426 (narrative retell), .475 (WRMT LC), .525 (narrative writing), and .525
(vocabulary inferencing). These results supported the use of multilevel modeling with student
outcomes nested within the school clustering variable.

**Missing Data.** In this study, we assumed all missing data to occur at random or due to
participant attrition. As shown in Table 5, missing data accounted for no more than 13% of any
one outcome. At pretest, 26 out of 930 records were missing (15 from the intervention group, 11
from the control group), a total of 2.8% of all pretest records. At posttest, less than 1% of the
total dependent measures were missing in total ($n = 6$ records). However, approximately 9% of
all follow-up data were missing at random or due to participant attrition (changed schools), with
8.6% of data missing for the treatment group and 9.5% missing for control group participants.
We did not use multiple imputation methods given the rate of attrition per group was less than
8% and not significant between groups.

**Two-level Hierarchical Linear Model Specification.** We used a 2-level hierarchical
linear model to describe the effects of the intervention and pretest scores on students’ language
and literacy outcomes at posttest or follow-up, after controlling for differences between students’
pretest scores, assignment of schools to groups, and variability in school R-TFI and examiner
fidelity outcomes. We used the guidelines provided by Ferron et al. (2008) for defining the
model specification.

The level-1 model below describes student posttest or follow-up outcomes as a function
of their pretest score within a specific school. For student outcomes scored using language
samples (i.e., expository retell and narrative writing), we also included examiner fidelity
associated with each elicited sample as a predictor of students’ scores. The level-1 model
specification for all student posttest and follow-up outcomes, except for expository retell and narrative writing, was:

\[ Y_{ij} = \beta_{0j} + \beta_{1j} \text{Pretest}_{ij} + r_{ij}, \]

where \( Y_{ij} \) is the raw posttest or follow-up score for the \( i \)th student in the \( j \)th school, \( \beta_{0j} \) is the intercept of the regression equation predicting \( Y_{ij} \) in the \( j \)th school, \( \beta_{1j} \) is the regression coefficient indicating the relationship strength of the Pretest score with the posttest/follow-up outcome for the \( i \)th student in the \( j \)th school, and \( r_{ij} \) is the assumed autocorrelated error.

We specified the level-1 model for expository oral retell and narrative writing as:

\[ Y_{ij} = \beta_{0j} + \beta_{1j} \text{Pretest}_{ij} + \beta_{2j} \text{Examiner Fidelity}_{ij} + r_{ij}, \]

where we held all variables and interpretations consistent with the previous model but added the regression coefficient \( \beta_{2j} \) as an index of the relationship between students’ scores and the fidelity to which examiners adhered to procedures when eliciting those scored samples for the \( i \)th student in the \( j \)th school.

We then developed the level-2 model based on assumptions made from previous implementation of the intervention in a MTSLS model (Petersen et al., 2022; Spencer et al., 2020, 2017, 2015; Weddle et al., 2016). We defined the level-2 model to consider the effects of intervention context on students’ posttest and follow-up outcomes, as we assumed the oral narrative intervention delivered by school-specific intervention teams would moderate student outcomes. Therefore, we specified the level-2 model using schools’ random assignment to either treatment or waitlist control group (Group) and the measure of MTSLS implementation (i.e., R-TFI) as predictors of some of the level-1 coefficients. For the proximal and distal outcomes, we defined the level-2 model as shown below.

\[ \beta_{0j} = \gamma_{00} + \gamma_{01} \text{Group}_j + \gamma_{02} \text{R-TFI}_j + u_{0j} \]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11} \text{Group}_j + u_{1j} \]

\[ \beta_{2j} = \gamma_{20} \]

We coded Group\( j \) as 0 if school \( j \) was randomly assigned to waitlist control conditions and 1 if school \( j \) was randomly assigned to implement the multitiered oral narrative intervention. \( R - TFI_j \) was the measure of school-level MTSLS implementation fidelity at school \( j \). In review of raw data, we assumed level-2 errors \( u_{0j} \) and \( u_{1j} \) normally distributed and allowed for random variability in the intercepts and students’ pretest scores, as well as within schools. We did not include the interaction term between group and pretest outcomes, as the schools assigned to the treatment condition did not implement the intervention until after they collected all pretest measures.

The resulting full model specification for the proximal and distal outcomes was:

\[ Y_{ij} = \gamma_{00} + \gamma_{01} \text{Group}_j + \gamma_{02}R-TFI_j + \gamma_{10}\text{Pretest}_{ij} + u_{0j} + u_{1j} + \eta_j \]

For intermediate expository oral retell and narrative writing outcomes, the full model specification was:

\[ Y_{ij} = \gamma_{00} + \gamma_{01} \text{Group}_j + \gamma_{02}R-TFI_j + \gamma_{10}\text{Pretest}_{ij} + \gamma_{20}\text{Examiner Fidelity}_{ij} + u_{1j}\text{Pretest}_{ij} + u_{0j} + \eta_j \]

With some outcome variance parameters approaching zero, we used the Satterthwaite approach (Littell et al., 1996) in our assumptions for both variance parameters and fixed effects. We tested the null hypotheses using \( F \)-tests with approximate degrees of freedom reporting the results bounded by a 95% confidence interval (CI). We adjusted all \( p \)-values to account for multiple comparisons (Benjamini & Hochberg, 2000) and used the results when making assumptions about statistical significance. In addition, to provide a further measure of the
magnitude of intervention effects using student data nested within schools, we used effect size estimation methods appropriate for multilevel models (Feingold, 2009).

Results

Preliminary Analysis

We examined the raw data and descriptive statistics reported in Table 5 and checked for normality assumptions prior to analyzing the effects of the multiltiered oral narrative intervention on student’s language and literacy outcomes. Across all three time points, we estimated the effects using a series of separate two-level hierarchical linear models that allowed us to account for students nested within schools. We summarized parameter estimates for the unconditional and full models in Tables 6–8 organized by the anticipated relationship between each measure and its outcome (i.e., proximal to distal).

Equivalence of Groups. Table 5 displays each group’s descriptive statistics by outcome and study phase (i.e., pretest, posttest, and follow-up). Groups were approximately equal in size across all phases. We conducted multilevel ANOVAs for all pretest data to check for group differences in students’ outcomes controlling for examiner fidelity to language sampling procedures (level-1), as well as school R-TFI at posttest (level-2). Justifying inclusion of pretest scores as level-1 covariates in the HLM, pretest differences between groups were statistically significant for narrative writing ($F = 5.31, p = .023$) and expository retell ($F = 4.47, p = .036$).

Intervention Implementation Fidelity

As the study focused on intervention delivery by teachers and SLPs in routine practice, we used R-TFI outcomes as a measure of adherence to MTSLS intervention implementation procedures. On average, schools in the treatment group reported meeting 81.9% of the MTSLS implementation indicators, compared to 65.8% reported by the schools in the control group.
Follow-up t-tests indicated that significant differences between group R-TFI scores at posttest favored the schools assigned to the treatment condition, $t(1, 144.2) = 6.86, p < .0001$. As shown in Tables 6–8, the school-level R-TFI score was a significant contributor to the hierarchical linear models for narrative retell, vocabulary inferencing, narrative writing, and WRMT outcomes.

**Examiner Fidelity**

**Expository Oral Retell.** On average as shown in Table 5, the SLP examiners at schools assigned to waitlist control conditions adhered to the language sample elicitation procedural checklist with 97% fidelity at pretest (range: 72–100%), 95% at posttest (range: 35–100%), and 96% at follow-up (range: 74–100%). For examiners assigned to schools implementing the intervention, fidelity to language sample elicitation procedures was lower on average, with 88% at pretest (range: 41–100%), 90% at posttest (range: 65–100%), and 96% at follow-up (range: 67–100%). Follow-up t-tests indicated that significant differences between group examiner fidelity outcomes favored the schools assigned to the waitlist control at pretest, $t(1, 99.24) = 5.81, p < .0001$ and posttest, $t(152.54) = 2.86, p = .005$. When used as a covariate in the two-level models, examiner fidelity was not a significant predictor of student expository retell scores at either posttest or follow-up, as shown in Table 7. However, we kept the variable in the model to control for differences in language sampling across time points.

**Narrative Writing.** Overall, examiners across schools followed the writing sample elicitation procedures approximately 95% of the time. Writing samples were elicited from students in schools randomly assigned to the waitlist control condition with 96% fidelity at pretest (range: 73%–100%), and 95% at posttest and follow-up (range: 67%–100%). At pretest (95%) and posttest (93%), examiners at schools assigned to the treatment condition had lower
levels of adherence to the narrative writing sampling procedures; the range of examiner fidelity during sampling of students’ written narratives was 55%–100% across both time points. However, at follow-up, examiner fidelity at the treatment schools was 98% (range: 0.64–1.0), higher than the waitlist control group. Follow-up t-tests indicated that significant differences between group examiner fidelity outcomes existed only at follow-up favoring the treatment group, $t(129.51)= 1.99, p = .0492$. However, in the HLM, examiner fidelity to narrative writing sampling procedures was not a significant predictor of student outcomes at any time point.

**Student Language and Literacy Outcomes**

**Narrative Retell.** Review of descriptive statistics in Table 5 reveals a difference in group narrative retell outcomes from pretest to posttest. At the start of their first-grade year, prior to implementation of the multitiered oral narrative intervention, there was not a significant difference between groups’ narrative writing pretest scores. By the end of first grade, only students who received the intervention made large gains in their ability to retell a complete and coherent story using complex language, scoring an average of 3.7 points higher on the NLM Listening Retell measure than their peers attending schools assigned to the waitlist control condition. As shown in Table 6, after controlling for pretest scores and school R-TFI, the multitiered oral narrative intervention resulted in large and statistically significant positive effects on students’ posttest narrative retell scores ($\beta = 5.68, p = .007; ES = 2.64$). When assessed again at the start of their second-grade year (follow-up), however, the difference between groups that existed six months earlier was smaller and no longer significant, favoring the students who did not receive the intervention, $\beta = 1.12, p = .28$ (ES = .22).

**Vocabulary Inferencing.** As shown in Table 6, after controlling for students’ pretest and school R-TFI scores, the intervention resulted in statistically significant, positive, and large
effects on students’ vocabulary inferencing skills ($\beta = 1.62$, $p = .01$; $ES = 2.08$). As measured using the NLM Listening Vocabulary Questions section, students in the treatment group had a mean score of 7.28 ($SD=2.03$), compared to the control group mean of 6.17 ($SD=1.70$). However, after controlling for students’ pretest scores and school R-TFI, students who received the treatment did not appear to maintain the vocabulary inferencing gains made from pretest to posttest when administered a follow-up measure at the start of Grade 2. Table 6 shows that there was a positive moderate, but non-significant effect on vocabulary inferencing at follow-up for students who received the intervention in Grade 1 ($ES = .38$).

**Expository Oral Retell.** Table 7 presents the multilevel modeling results of intervention effects on students’ expository oral retell outcomes at posttest and follow-up. Students who received the Story Champs Complete intervention scored 2.2 points higher at posttest and 3.13 points higher at follow-up on the expository oral retell measure than their peers at the waitlist control group schools. After controlling for students’ pretest scores (a significant model predictor), R-TFI, and examiner fidelity scores, there was a moderate but non-significant intervention effect at posttest, $\beta = 1.98$, $p = .12$ ($ES = .30$). However, when examiners collected follow-up expository oral retell samples from the students six months later, in the fall of 2019, students in the treatment group scored higher on the outcome on average than did the peers who did not receive the multitiered intervention in first grade ($ES = 1.22$). After adjusting for multiple comparisons (Benjamini & Hochberg, 2000), that difference remained statistically significant at follow-up: $\beta = 3.27$, $p = .03$.

**Narrative Writing.** Students who received the intervention had narrative written scores at least 4.5 points higher than their control group peers at posttest (out of a total possible 59 points). Using pretest scores as a level-1 covariate and controlling for level-2 variables (R-TFI,
examiner fidelity), students who received the intervention in first grade performed significantly better on narrative writing outcomes at the start of second grade when compared to peers who did not receive the multitiered oral narrative intervention ($\beta = 7.01, p = .001, ES = .93$). Shown in Table 7, the treatment effect maintained to some degree over the summer, with students’ narrative writing differences at follow-up remaining almost 1.5 points higher than the waitlist control group. The intervention was associated with statistically significant large writing improvements ($ES = 2.15$). However, after adjusting for multiple comparisons, the differences between groups were no longer statistically significant ($\beta = 4.77, p = .09$).

**WRMT Listening Comprehension.** At the end of the first-grade year, students who did not receive the intervention slightly outperformed their peers in the intervention group on the distal WRMT LC measure, as shown in Table 5. However, review of multilevel model results in Table 8 shows that posttest group differences were non-significant after controlling for pretest score differences and school R-TFI outcomes, $\beta = .12, p = .81$ ($ES = -.16$). By the start of the second grade, students who received the multitiered oral narrative intervention scored higher than peers who did not receive the intervention, but the positive effect was small ($ES = .63$) and not statistically significant after controlling for pretest and school R-TFI variables, $\beta = .99, p = .51$.

**WRMT Passage Comprehension.** At posttest, there was a negligible difference between group means on the distal outcome measure, WRMT PC, as shown in Table 5. While intervention had a positive but small effect, performance was not statistically different between groups at posttest, $\beta = .21, p = .94$ ($ES = .21$). Yet, distal effects were more noticeable at follow-up. WRMT PC measures at the start of Grade 2 indicated moderate to large positive intervention effects, $ES = .997$. However, as shown in Table 8, the difference between groups’ follow-up
scores was not statistically significant after adjusting for multiple comparisons (Benjamini-Hochberg, 200) and controlling for students’ pretest and school R-TFI scores, $\beta = 1.08, p = .11$. 

**Discussion**

The current study compared the effects of existing classroom English language arts instruction to an oral narrative intervention delivered using a MTSLS framework on the language and literacy skills of first graders at risk for or having language disabilities. As shown in Tables 6–8, the *Story Champs Curriculum* resulted in students making at least small to moderate improvements on all measured outcomes from pretest to posttest (ES range: .21–2.64), apart from the distal WMRT-III Listening Comprehension subtest (ES = -.16). For students assigned to schools in the treatment condition, posttest gains were also significantly larger than their peers in the waitlist control condition for several outcomes. As expected, only students who received the multitiered oral narrative intervention showed immediate and significant gains in narrative retell and vocabulary inferencing skills at the end of their first year (ESs = 2.64–2.08, respectively). Students in treatment schools scored an average of 3.7 points higher on the NLM Listening retell subtest and more than 1.03 points (33%) higher on the NLM Vocabulary Questions subtest. These results are similar in effect to prior studies of *Story Champs* (Spencer et al., 2014; Spencer et al., 2015; Spencer et al., 2017b).

The narrative writing and expository oral retell outcomes were intended to be intermediate measures of effect. After controlling for differences in schools’ fidelity to the MTSLS model, as well as individual differences in pretest and examiner fidelity scores, there was evidence of a large and significant treatment effect for first graders’ narrative writing (ES = .93, $p = .001$). These results strengthen the argument for using oral narrative intervention alongside existing handwriting (transcription) instruction. The significant effects also extend
previous findings of a causal relation (Kirby et al., 2021; Petersen et al., 2022; Petersen et al., 2020; Spencer & Petersen, 2018).

Whereas prior evidence of intervention effect on writing outcomes is emerging, only one prior study of *Story Champs* found a treatment effect for expository oral language (e.g., Petersen et al., 2022). Thus, we were surprised to find that the oral narrative intervention led to a positive, large, and statistically significant improvement in young students’ expository language six months after the intervention ended (ES = 1.22). In absence of additional instruction, students who received the multitiered oral narrative intervention in first grade were able to listen to informational passages read aloud and retell the same information when in second grade, cohesively summarizing main ideas and key supporting details using advanced academic language at a statistically higher level than peers who did not receive the *Story Champs Curriculum* in first grade. This result may indicate that students’ developmental trajectory of expository language skills may be slower than for the narrative academic language counterpart. Replication studies in K-1 grades should be able to confirm or rule out these hypotheses. Alternatively, it may be possible that the expository oral retell measure at the end of first grade was not sensitive enough to confidently detect the presence of an effect, like we suspect happened with the distal standardized measures.

Although many language intervention studies include distal norm-referenced standardized measures of language and literacy skills (e.g., WRMT-III), these measures are not known to be sensitive to the presence of oral narrative intervention effects (Favot et al., 2021; Pico et al., 2021). We included these measures to examine the strength to which the oral narrative intervention may improve students’ scores on norm-referenced standardized measures, often used to identify students with language disabilities. However, students attending schools
randomly assigned to the waitlist control condition outperformed student in the treatment schools, resulting in our questioning of the appropriateness of such measures. There are possible reasons for why we did not see improvements in distal standardized measures of effect. With treatment and instruction-as-usual control groups having 77 and 78 students respectively, the study was underpowered. Yet, the differential attrition rate was within the threshold of acceptability specified by the What Works Clearinghouse (WWC). Therefore, it is possible that the intervention dosage was not sufficient enough for students to demonstrate significant improvements on the standardized measures within one school year. Overall, we find the positive effects on narrative and expository retell, vocabulary, and writing for students who received the oral narrative intervention at posttest and follow-up to be meaningful and encouraging.

The current study contributes more evidence of the impact of using school-based teams of teachers and SLPs in the prevention of worsening language and literacy conditions in the U.S. educational system. In schools assigned to the waitlist control condition, teachers implemented ELA curricula as-usual, meaning students received instruction as delivered in a one-size-fits-all model. Classroom teachers did not have regular access to the knowledge and expertise of SLPs, and vice versa. Rather than intervene only when students failed to meet grade-level expectations, treatment group schools implemented the oral narrative intervention in a MTSLS framework, which demanded more frequent collaboration. The MTSLS model promoted teaming between teachers and SLPs around a critical academic skill that serves as a foundation for reading comprehension and writing. Even when variability existed between schools’ MTSLS implementation, schools assigned to the treatment group were more likely to implement key ingredients of the model which significantly contributed to students’ overall outcomes.
Evidence is building in support for language and literacy curricula designed intentionally for delivery in a MTSLS model (Petersen et al., 2022). Educational interventions are usually designed for specific student populations and rely on individualized targeted instruction to ameliorate academic problems. However, where there is evidence of intervention effectiveness with students experiencing or at risk of language and literacy problems, scalability is necessary to prevent more students from experiencing future academic failure. One way to address this problem is to design language and literacy curricula with both general instruction and targeted intervention in mind, from the start. Oral language programs like Story Champs Curriculum that include lesson plans aligned with instructional tiers (i.e., whole group, small group, individual) benefit students by providing sufficient differentiation for each student’s needs. When provided with curricula with multitiered lesson plans, teachers and/or specialists do not require additional time to adjust plans to suit the setting of delivery. Although there is an obvious benefit when planning time is reduced, more research is needed to evaluate the time and financial cost benefit of well-designed multitiered curriculum (Forman et al., 2009; Simmons et al., 2007; Walker, 2004). Nonetheless, impact on student outcomes is readily apparent now.

Limitations and Future Directions

There are several limitations to this study that impact the interpretation of findings. First, we are limited in our ability to make assumptions of intervention effect because the study was underpowered with a total of 155 student participants. Although attrition was assumed exogenous (i.e., attributed to student’s moving out of the school district; WWC, 2020) and the rate was low (n = 3 students per group) overall, a total of 75 students in the treatment group and 74 students in the waitlist control group limited the power available to conduct statistical analyses of effect. In addition, these results are restricted to a population of first graders at risk of
or identified with language disabilities. In other words, these results do not generalize to all first-grade students because we did not include students performing at or above benchmark on the universal screener as a comparison group as Petersen et al. (2022) did.

We also caution others from making generalized assumptions about the intervention’s ameliorative and preventative effects. The evaluation of such effects would require data on the number of students with IEPs in the treatment and waitlist control groups who were no longer found eligible for special education services. This information was not available to us at that time. In addition, to examine longitudinal preventative effects, we would have needed to follow students through the end of their second-grade year, spring 2020. Unfortunately, due to the global COVID-19 pandemic, all students experienced a disruption in their schooling. Therefore, any outcome data collected in spring 2020 would not have been comparable to other time points. In the future, it would be useful to evaluate the degree to which effective interventions ameliorate existing language problems in addition to acting as a preventative strategy, so long as instructional contexts are comparable.

The development of research-informed language and literacy intervention usually starts with testing under optimal conditions, but we consider the RPP context to be a strength in that it allowed us to assess intervention effects in real-world conditions. Involvement of school staff and specialists in key aspects of the study (e.g., data collection, intervention implementation) enhanced the social validity of this research and provided data on implementation quality (Domitrovich et al., 2008; Lendrum & Humphrey, 2012). For example, by examining schools’ R-TFI scores and examiner fidelity to assessment procedures, we identified discrepancies between what RPP stakeholders planned for and what was actually implemented. This information was valuable when interpreting study outcomes. It also led to certain limitations in
our ability to make assumptions regarding effectiveness as the school-based teams did not perfectly implement the intervention to the degree expected of researcher-implemented programs.

Although fitting with practices of an effectiveness study within a RPP context (Breitenstein et al., 2010), treatment effects may have been larger if researchers solely implemented the study. By not directly collecting student language samples ourselves, we found that training and the provision of a procedural checklist alone was not enough to result in consistent adherence to language sample elicitation protocols. However, this is a problem faced by many community-based participatory researchers—when researchers prioritize end users, they must make tradeoffs that weaken the degree of experimental control (Foster, 2014; Nelson et al., 2009). In other words, by giving up tight control over the contingencies governing the behavior of the teachers and SLPs, we increased the potential for variability in outcomes. Future research is necessary to identify methods for supporting school staff responsible for key study activities rather than avoid intervention research studying effects when applied by school teams in real world contexts.

While significant differences in examiner language sampling procedural fidelity did not significantly influence students’ expository oral retell and narrative writing outcomes, future research using student outcomes derived from language samples may need to consider the influence of examiner behavior. The measurement and reporting of community partners’ fidelity to assessment procedures may be just as important as the more common method of measuring and reporting treatment fidelity. To date, only one large-scale study of an oral language intervention has reported a measure of school-based study teams’ fidelity of data collection procedures. In the Petersen et al. (2022) study of a kindergarten oral narrative intervention
implemented within a MTSLS framework, the authors randomly sampled 30% of all audio-recorded expository language samples and reported a range of 76%–100% examiner fidelity to data collection procedures (overall mean: 94.8%). For this study, we analyzed every audio recording and were surprised to find great variability among examiners’ adherence to language sampling procedures. Across groups, examiners’ fidelity to expository oral retell elicitation scripts was 35%–100%. Researchers interested in the effects of interventions implemented by intended end users should continue to include community partners as data collectors, but as we move away from colonial research practices that traditionally benefited university-based researchers (Pritchett et al., 2021; van der Westhuizen, 2013) and toward community-based participatory research models using RPPs, we will need to develop technologies and methods for ensuring data collected are valid measures of a student’s language abilities in real-world conditions. We will need to rely on innovations emerging from implementation science to foster greater fidelity in practice settings with the ultimate goal of reducing the educational research-practice gap.

**Conclusion**

If we are to address the lack of reading and writing proficiency of U.S. students in fourth grade, we will need to intensify our focus on the prevention of language and literacy problems and work collaboratively. That means we need to capitalize on the strengths of MTSLS, with teachers and SLPs working alongside one another to prevent persistent language problems from curtailing students’ overall academic success. Altogether, evidence of the effects of a multitiered oral narrative intervention in this study show its potential for preventing language and learning difficulties in first grade within a MTSLS context. Considering prior research in preschool (Spencer et al., 2015; 2017; 2020), kindergarten (Petersen et al., 2022), and second and third
graders (Petersen et al., 2020), the current study fills the gap at first grade, adding to a solid body of research supporting multitiered oral narrative intervention.

References


Kim, Y.-S. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension or oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141*, 101–120. https://doi.org/10.1016/j.jecp.2015.08.003


95


Sylvan, L. (2018). Tiers to communication success: How can SLPs join in the MTSS framework many schools are adopting to catch students’ special education needs earlier and provide levels of intervention? *ASHA Leader, 23*(8), 44–53. https://doi.org/10.1044/leader.FTR1.23082018.44


Table 3  

**Student Participant Sample Demographics**

<table>
<thead>
<tr>
<th></th>
<th>Treatment ($n=78$)</th>
<th>Control ($n=77$)</th>
<th>Total ($N=155$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
<td>41</td>
<td>78</td>
</tr>
<tr>
<td>Male</td>
<td>41</td>
<td>36</td>
<td>77</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian, Asian American</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Black</td>
<td>8</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Latino/a</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>White</td>
<td>62</td>
<td>53</td>
<td>115</td>
</tr>
<tr>
<td>Multiracial</td>
<td>4</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Not answered</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>75</td>
<td>72</td>
<td>147</td>
</tr>
<tr>
<td>Not answered</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>Home Language</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English only</td>
<td>73</td>
<td>72</td>
<td>145</td>
</tr>
<tr>
<td>Spanish only</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>English and 1+ other language</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Not answered</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>Avg. students per school</strong></td>
<td>15 (range: 12-18)</td>
<td>15 (range: 8-29)</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Information about race and ethnicity missing for 6 participants.
<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>% of total (N=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td><strong>Race, ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, Non-Hispanic</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td><strong>Professional license</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech-language pathologist</td>
<td>8</td>
<td>27.6</td>
</tr>
<tr>
<td>General education, elementary</td>
<td>13</td>
<td>44.8</td>
</tr>
<tr>
<td>Special education, K-12</td>
<td>1</td>
<td>3.4</td>
</tr>
<tr>
<td>Reading specialist</td>
<td>2</td>
<td>6.9</td>
</tr>
<tr>
<td>None</td>
<td>2</td>
<td>6.9</td>
</tr>
<tr>
<td><strong>Highest level of education completed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school diploma</td>
<td>1</td>
<td>3.8</td>
</tr>
<tr>
<td>Associates degree</td>
<td>1</td>
<td>3.8</td>
</tr>
<tr>
<td>4-year undergraduate degree</td>
<td>7</td>
<td>26.9</td>
</tr>
<tr>
<td>Master's degree</td>
<td>17</td>
<td>65.4</td>
</tr>
<tr>
<td><strong>Years of experience in education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 (new to field)</td>
<td>1</td>
<td>3.8</td>
</tr>
<tr>
<td>1–5 years</td>
<td>4</td>
<td>15.4</td>
</tr>
<tr>
<td>6–10 years</td>
<td>6</td>
<td>23.1</td>
</tr>
<tr>
<td>11–15 years</td>
<td>3</td>
<td>11.5</td>
</tr>
<tr>
<td>16+ years</td>
<td>12</td>
<td>46.2</td>
</tr>
</tbody>
</table>

*Note.* Participants may have more than 1 license (total of 29 licenses reported)
Table 5

Descriptive Data for Variables in 2-Level Model with Students Nested in Schools

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th></th>
<th></th>
<th>Posttest</th>
<th></th>
<th></th>
<th></th>
<th>Follow-up</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>% missing</td>
<td>m</td>
<td>sd</td>
<td>skew</td>
<td>n</td>
<td>% missing</td>
<td>m</td>
<td>sd</td>
<td>skew</td>
</tr>
<tr>
<td><strong>Story Champs Complete Intervention (N = 78)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrative retell</td>
<td>78</td>
<td>0</td>
<td>6.40</td>
<td>4.37</td>
<td>0.27</td>
<td>78</td>
<td>0.00</td>
<td>17.83</td>
<td>4.23</td>
<td>-0.49</td>
</tr>
<tr>
<td>Vocabulary inferencing</td>
<td>78</td>
<td>0</td>
<td>5.62</td>
<td>2.19</td>
<td>-0.48</td>
<td>78</td>
<td>0.00</td>
<td>7.28</td>
<td>2.03</td>
<td>-1.56</td>
</tr>
<tr>
<td>Expository oral retell</td>
<td>72</td>
<td>8</td>
<td>8.96</td>
<td>5.58</td>
<td>0.65</td>
<td>78</td>
<td>0.00</td>
<td>12.17</td>
<td>5.95</td>
<td>0.13</td>
</tr>
<tr>
<td>Narrative writing</td>
<td>78</td>
<td>0</td>
<td>8.29</td>
<td>6.03</td>
<td>0.58</td>
<td>78</td>
<td>0.00</td>
<td>18.62</td>
<td>6.60</td>
<td>-0.93</td>
</tr>
<tr>
<td>WRMT Listening Comprehension</td>
<td>75</td>
<td>4</td>
<td>6.64</td>
<td>1.63</td>
<td>0.47</td>
<td>77</td>
<td>0.01</td>
<td>6.90</td>
<td>1.99</td>
<td>0.11</td>
</tr>
<tr>
<td>WRMT Passage Comprehension</td>
<td>72</td>
<td>8</td>
<td>6.72</td>
<td>3.64</td>
<td>0.90</td>
<td>77</td>
<td>0.01</td>
<td>10.32</td>
<td>3.46</td>
<td>-0.67</td>
</tr>
<tr>
<td>Examiner fidelity (expository retell)</td>
<td>72</td>
<td>8</td>
<td>0.88</td>
<td>0.12</td>
<td>-1.37</td>
<td>78</td>
<td>0.00</td>
<td>0.90</td>
<td>0.10</td>
<td>-0.64</td>
</tr>
<tr>
<td>Examiner fidelity (narrative writing)</td>
<td>71</td>
<td>9</td>
<td>0.95</td>
<td>0.07</td>
<td>3.16</td>
<td>78</td>
<td>0.00</td>
<td>0.93</td>
<td>0.08</td>
<td>-2.55</td>
</tr>
<tr>
<td>R-TFI (level-2)</td>
<td>78</td>
<td>0.00</td>
<td>0.82</td>
<td>0.13</td>
<td>-0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Waitlist control (N = 77)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrative retell</td>
<td>77</td>
<td>1</td>
<td>6.94</td>
<td>4.31</td>
<td>-0.14</td>
<td>75</td>
<td>0.03</td>
<td>14.13</td>
<td>4.45</td>
<td>-0.68</td>
</tr>
<tr>
<td>Vocabulary inferencing</td>
<td>77</td>
<td>1</td>
<td>5.96</td>
<td>1.72</td>
<td>-0.24</td>
<td>75</td>
<td>0.03</td>
<td>6.17</td>
<td>1.70</td>
<td>-0.41</td>
</tr>
<tr>
<td>Expository oral retell</td>
<td>69</td>
<td>10</td>
<td>6.87</td>
<td>6.12</td>
<td>0.70</td>
<td>77</td>
<td>0.00</td>
<td>9.97</td>
<td>6.13</td>
<td>0.47</td>
</tr>
<tr>
<td>Narrative writing</td>
<td>76</td>
<td>1</td>
<td>11.21</td>
<td>6.91</td>
<td>-0.05</td>
<td>77</td>
<td>0.00</td>
<td>14.09</td>
<td>7.22</td>
<td>-0.14</td>
</tr>
<tr>
<td>WRMT Listening Comprehension</td>
<td>76</td>
<td>1</td>
<td>6.96</td>
<td>1.49</td>
<td>0.74</td>
<td>77</td>
<td>0.00</td>
<td>6.96</td>
<td>1.56</td>
<td>0.60</td>
</tr>
<tr>
<td>WRMT Passage Comprehension</td>
<td>76</td>
<td>1</td>
<td>7.41</td>
<td>4.78</td>
<td>0.41</td>
<td>77</td>
<td>0.00</td>
<td>10.69</td>
<td>3.97</td>
<td>-0.44</td>
</tr>
<tr>
<td>Examiner fidelity (expository retell)</td>
<td>69</td>
<td>10</td>
<td>0.97</td>
<td>0.06</td>
<td>-2.43</td>
<td>77</td>
<td>0.00</td>
<td>0.95</td>
<td>0.09</td>
<td>-4.28</td>
</tr>
<tr>
<td>Examiner fidelity (narrative writing)</td>
<td>68</td>
<td>12</td>
<td>0.96</td>
<td>0.06</td>
<td>-1.58</td>
<td>77</td>
<td>0.00</td>
<td>0.95</td>
<td>0.06</td>
<td>-1.19</td>
</tr>
<tr>
<td>R-TFI (level-2)</td>
<td>77</td>
<td>0.00</td>
<td>0.66</td>
<td>0.16</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Attrition: 3 students in treatment and 3 students in waitlist control groups moved and are not reported at posttest; % missing = percentage of total student samples missing (absent from school or missing at random), \( m = \) mean, \( sd = \) standard deviation, skew = skewness
Table 6

*Posttest and Follow-Up Hierarchical Linear Models and ES Estimates for Narrative Retell and Vocabulary Inferencing After Controlling for Pretest and School R-TFI Scores*

<table>
<thead>
<tr>
<th></th>
<th>Unconditional model</th>
<th>Posttest</th>
<th>Full model</th>
<th>ES</th>
<th>Unconditional model</th>
<th>Follow-up</th>
<th>Full model</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrator retell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>16.94**</td>
<td>1.11</td>
<td>13.54 to 18.54</td>
<td></td>
<td>26.79**</td>
<td>3.54</td>
<td>18.76 to 34.64</td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>0.08</td>
<td>0.07</td>
<td>-0.05 to 0.22</td>
<td></td>
<td>0.12</td>
<td>0.08</td>
<td>-0.04 to -0.28</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>5.68*</td>
<td>1.39</td>
<td>2.57 to 8.80</td>
<td></td>
<td>11.29*</td>
<td>4.18</td>
<td>1.9 to 20.67</td>
<td>2.64*</td>
</tr>
<tr>
<td>R-TFI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 (σ²)</td>
<td>12.70</td>
<td>1.50</td>
<td>8.02 to 17.38</td>
<td></td>
<td>11.68**</td>
<td>1.50</td>
<td>8.02 to 17.38</td>
<td></td>
</tr>
<tr>
<td>Intercept (τw)</td>
<td>11.25</td>
<td>5.74</td>
<td>1.56*</td>
<td></td>
<td>3.00</td>
<td>1.81</td>
<td>0.25 to 5.76</td>
<td></td>
</tr>
<tr>
<td>Error covariance</td>
<td>0.958</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary inferencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>6.63**</td>
<td>0.34</td>
<td>5.87 to 7.40</td>
<td></td>
<td>9.33**</td>
<td>1.18</td>
<td>6.86 to 11.81</td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>0.08</td>
<td>0.07</td>
<td>-0.08 to 0.20</td>
<td></td>
<td>0.10</td>
<td>0.12</td>
<td>-0.13 to -0.34</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>1.62*</td>
<td>0.46</td>
<td>0.61 to 2.62</td>
<td></td>
<td>1.58</td>
<td>0.79</td>
<td>0.37 to 2.79</td>
<td></td>
</tr>
<tr>
<td>R-TFI</td>
<td>3.01*</td>
<td>1.35</td>
<td>1.11 to 5.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 (σ²)</td>
<td>3.02</td>
<td>0.36</td>
<td>2.86 to 3.19</td>
<td></td>
<td>1.22</td>
<td>0.55</td>
<td>0.63 to 1.81</td>
<td></td>
</tr>
<tr>
<td>Intercept (τw)</td>
<td>0.93</td>
<td>0.54</td>
<td>0.17*</td>
<td></td>
<td>1.79</td>
<td>1.10</td>
<td>0.63 to 2.95</td>
<td></td>
</tr>
<tr>
<td>Error covariance</td>
<td>0.76</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7

Posttest and Follow-Up Hierarchical Linear Models and ES Estimates for Expository Oral Retell and Narrative Writing After Controlling for Pretest, Examiner Fidelity, and School R-TFI Scores

<table>
<thead>
<tr>
<th></th>
<th>Unconditional model</th>
<th>Posttest</th>
<th>Full model</th>
<th>ES</th>
<th>Unconditional model</th>
<th>Posttest</th>
<th>Full model</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>SE</td>
<td>95% CI</td>
<td>β</td>
<td>SE</td>
<td>95% CI</td>
<td>ES</td>
<td>β</td>
</tr>
<tr>
<td><strong>Expository oral retell</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>11.10**</td>
<td>0.54</td>
<td>9.88 to 12.32</td>
<td>14.32</td>
<td>7.22</td>
<td>-2.5 to 28.89</td>
<td>0.30</td>
<td>12.47**</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.41**</td>
<td>0.08</td>
<td>-2.5 to 56.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner fidelity, pre-</td>
<td>1.34</td>
<td>4.54</td>
<td>-8.05 to 10.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner fidelity, post-</td>
<td>-4.87</td>
<td>4.96</td>
<td>-14.68 to 4.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>1.98</td>
<td>1.12</td>
<td>-2.3 to 4.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-TFI</td>
<td>3.24</td>
<td>3.44</td>
<td>3.56 to 10.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 (σ²)</td>
<td>37.07</td>
<td>4.32</td>
<td>Z value = 8.57**</td>
<td>12.73</td>
<td>14.29</td>
<td></td>
<td></td>
<td>27.64</td>
</tr>
<tr>
<td>Intercept (to)</td>
<td>0.46</td>
<td>1.24</td>
<td>Z value = -3.7, p = .35</td>
<td>0.68</td>
<td>0</td>
<td></td>
<td></td>
<td>2.99</td>
</tr>
<tr>
<td>Error covariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Narrative writing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>16.52**</td>
<td>1.14</td>
<td>13.94 to 19.10</td>
<td>9.16</td>
<td>12.26</td>
<td>-15.08 to 33.40</td>
<td></td>
<td>18.39**</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.38**</td>
<td>0.09</td>
<td>21.0 to 55.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner fidelity, pre-</td>
<td>2.11</td>
<td>9.27</td>
<td>-16.23 to 20.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner fidelity, post-</td>
<td>7.54</td>
<td>7.32</td>
<td>-6.93 to 22.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>7.01*</td>
<td>1.28</td>
<td>4.49 to 9.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-TFI</td>
<td>2.91</td>
<td>4.03</td>
<td>-5.07 to 10.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 (σ²)</td>
<td>47.40</td>
<td>5.60</td>
<td>Z value = 8.47**</td>
<td>41.91</td>
<td>0.41</td>
<td></td>
<td></td>
<td>38.13</td>
</tr>
<tr>
<td>Intercept (to)</td>
<td>9.66</td>
<td>6.67</td>
<td>Z value = 1.45, p = .074</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>12.65</td>
</tr>
<tr>
<td>Error covariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8

Posttest and Follow-up Hierarchical Linear Models and ES Estimates for WRMT-III Listening Comprehension and Passage Comprehension After Controlling for Pretest and School R-TFI Scores

<table>
<thead>
<tr>
<th></th>
<th>Unconditional model</th>
<th>Posttest</th>
<th>Full model</th>
<th>ES</th>
<th>Unconditional model</th>
<th>Follow-up</th>
<th>Full model</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WRMT Listening Comprehension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>7.01**</td>
<td>0.29</td>
<td>6.37 to 7.65</td>
<td></td>
<td>3.49</td>
<td>1.23</td>
<td>.80 to 6.0</td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>-0.44**</td>
<td>0.09</td>
<td>-.72 to -.21</td>
<td></td>
<td>-0.12</td>
<td>0.48</td>
<td>-1.18 to 0.95</td>
<td>-.16</td>
</tr>
<tr>
<td>Group</td>
<td>-0.66</td>
<td>1.38</td>
<td>-2.68 to 1.36</td>
<td></td>
<td>-0.66</td>
<td>3.68 to 2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-TFI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 ($\sigma^2$)</td>
<td>2.64</td>
<td>0.31</td>
<td></td>
<td></td>
<td>0.99</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\alpha$)</td>
<td>0.61</td>
<td>0.37</td>
<td></td>
<td></td>
<td>1.64</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error covariance</td>
<td>0.63</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WRMT Passage Comprehension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>10.46**</td>
<td>0.42</td>
<td>9.52 to 11.40</td>
<td>7.91**</td>
<td>1.45</td>
<td>4.84 to 10.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>0.55**</td>
<td>0.05</td>
<td>.50 to .67</td>
<td></td>
<td>0.21</td>
<td>0.57</td>
<td>-1.04 to 1.45</td>
<td>0.21</td>
</tr>
<tr>
<td>Group</td>
<td>0.83</td>
<td>0.74</td>
<td></td>
<td></td>
<td>1.81</td>
<td>1.71</td>
<td>-1.9 to 5.51</td>
<td></td>
</tr>
<tr>
<td>R-TFI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-1 ($\sigma^2$)</td>
<td>13.01</td>
<td>1.53</td>
<td></td>
<td></td>
<td>7.03</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\alpha$)</td>
<td>0.53</td>
<td>0.74</td>
<td></td>
<td></td>
<td>1.12</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error covariance</td>
<td>0.69</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: β = Parameter estimate; SE = Standard error; 95% CI = 95% confidence interval; ES = Effect size estimated using Hedges' method for multilevel models. Group results for treatment group coded 1; $p$-values adjusted using Benjamini-Hochberg (2000) method: **$p < .001$, *$p < .05$; Full-model estimated using ML, degrees of freedom estimated using Satterthwaite method.
CHAPTER 4:
ONLINE ASYNCHRONOUS ORAL NARRATIVE LANGUAGE TRAINING FOR CAREGIVERS OF CHILDREN AGED 3 TO 5 YEARS OLD

Introduction

Oral language skills in early childhood are significant predictors of life-long academic, social-emotional, and behavioral outcomes (Carmiol et al., 2020; Claessens et al., 2009; Duncan et al., 2007; Jones et al., 2015). The degree to which children have strong oral language models in early childhood can be a better predictor of children’s later academic achievement than socioeconomic status, race, gender, and parental level of education (Dickinson & Porche, 2011; Hart & Risley, 1995; Hoff, 2013; Rowe, 2008). During their preschool years, it is important for children to be exposed to and participate in language rich environments because oral language abilities in preschool are closely related to later reading and writing skills (Griffin et al, 2004; Mol & Bus, 2011). For example, children’s vocabulary acquisition between the ages of 2 and 4 years can predict their academic achievement in elementary school (Rowe, 2012; Walker et al., 1994) and plot structure in preschoolers’ play narratives predict elementary students’ narrative writing skills (Griffin et al., 2004). Disparities in adults’ health and academic outcomes can also be tied back to language nutrition provided to children in their early years (Head Zauche et al., 2017). Language nutrition is a term describing rich and substantive language ingredients present in a child’s environment that nourish their neurological, social, and linguistic development (Head Zauche et al., 2016). If language nutrition differences are not addressed immediately, oral
language delays can become fixed reading and writing problems by the second grade (Ferrer et al., 2015; Spira et al., 2015). The provision of language nutrition in early childhood is critical to preschoolers’ development of language and literacy skills necessary for academic success in kindergarten and beyond (Sheridan et al., 2011; Steiner, 2014).

One way to reduce differences in children’s language nutrition is to provide early childhood education. Preschool attendance is associated with lower levels of grade retention, need for special education services, literacy scores at age 19, and overall higher levels of advanced education and wages (Garces et al., 2002; Schweinhart et al., 1993). However, fewer than half of all U.S. children aged three to five years attend preschool (The Annie E. Casey Foundation, 2020; NCES, 2020). As a result, students enter school without the benefit of the oral language preparation early childhood education offers. Not all communities have publicly funded pre-kindergarten programs for children to attend (Friese et al., 2017). For others, transportation barriers and conflicts with caregiver work schedules prevent young children from attending preschool (Johnson, 2018; Jones et al., 2013; Kazdin et al., 1997). Because disparities in language input and learning can set the stage for a child’s later academic and social success, it is important to adapt early education programs to reach all children within a community.

To remediate the preschool access gap, the U.S. government has increased funding allocated to communities that train volunteers or early childhood specialists as home visitors (Duggan et al., 2018; United States Government Accountability Office, 2019). However, these interventions are typically expensive, require high levels of integrity to be effective, and are difficult to sustain over long periods of time (Biel et al., 2020; Barton & Fettig, 2013; Gubbels et al., 2019; Roberts et al., 2019). Additionally, the intervention dosage can vary from family to family because home visitors typically deliver services in-person based on their own schedule of
availability (i.e., during hours of operation). As a result, 60% of families enrolled in a home visiting program reported receiving only half of the number of sessions recommended by program developers (Duggan et al., 2018). Given the importance of early childhood language development, interventions that rely on parents or caregivers as interventionists may be a more practical and effective solution to a more universal provision of language-building early education (Walker et al., 2020).

**Parent-Implemented Early Childhood Intervention**

As facilitators of enriched language environments, parents and caregivers play a key role in their children’s development. Extensive research has shown a strong relationship between a child’s home language environment and their developmental trajectory (Chang & Monaghan, 2018; Hart & Risley, 1995; Hurtado et al., 2008). Toddlers’ cognitive and language development is affected by the frequency and quality of social interactions with caretakers regardless of a family’s socio-economic status (Weisleder & Fernald, 2013). Communicative interactions between parents and their infants during storybook reading at six months (O’Farrelly et al., 2018) or play at ten months (Mermelshtine & Barnes, 2016) can predict social-emotional, cognitive, and oral language skills at one year and beyond. In their preschool years, several studies have shown that the diversity of words and quantity of complex language in parental speech can significantly improve children’s vocabulary (e.g., Hadley et al., 2017a, 2017b; Suggate et al., 2018; Zimmerman et al., 2009). Since parents are likely to be “first responders” in their children’s oral language development—the first to respond to, reinforce, and/or notice problems, they are ideally positioned to act as their children’s first teachers.

Early childhood researchers and practitioners have spent decades considering the best way to design and implement effective parenting programs to prevent developmental problems
that may arise in a child’s preschool years. By providing parents and caregivers with knowledge about their children’s development and coaching to support new practices, parent training interventions can counteract the effects of problem behavior or developmental delays (e.g., Agazzi et al., 2019; Bonin et al., 2011; Breider et al., 2019; Rakap & Rakap, 2014; Scott et al., 2001). Meta-analyses of parent training programs designed to supporting early childhood language development report significant and moderate to large effects on children’s vocabulary and other language skills (Anderson et al., 2021; Jeong et al., 2021). For example, Jeong et al. (2021) reviewed programs used around the world to train parents of children aged birth-3 years and found positive and moderate intervention effects on improving children’s language development (ES SMD = .25, \( p < .0001 \)). Anderson and colleagues (2021) found the largest effects in studies that assessed caregivers’ linguistic complexity and quantity in the context of their homes (i.e., naturalistic observation), versus measuring caregiver linguistic input during free play or structured tasks. However, in a review of language intervention research conducted with low-income families between 1980 and 2016, Greenwood and colleagues (2020) discovered that researchers led majority of parent training programs in small groups outside of the home. However, traditional models of parent training that utilize in-person individual or group delivery formats are associated with high rates of attrition due to financial, transportation, childcare, and scheduling barriers (Johnson, 2018; Kazdin et al., 1997; O’Brien et al., 2012). Thus, traditional methods of in-person, out-of-home program implementation may not be scalable to meet the needs of all members within a community.

**Online Parent Training**

As an alternative to in-person delivery, there is emerging evidence that online and traditional in-person methods may be equally as effective in the prevention or treatment of
children’s mental and behavioral health problems (Breitenstein et al., 2014; Breider et al., 2019; Spencer et al., 2020). Further, digital parent training can increase caregivers’ engagement and participation in learning experiences, as well as significantly improve their knowledge and skills (Corralejo & Domenech Rodríguez, 2018; Doty et al., 2016; Spencer et al., 2020; Suarez et al., 2016). However, few effective parent training programs targeting the improvement of children’s language involve online implementation (Bellon-Harn et al., 2020; Greenwood et al., 2020; Jeong et al., 2021). For example, of the 46 studies reporting child language outcomes, Jeong and colleagues (2021) found evidence of only one program using technology as a delivery mechanism.

Emerging evidence suggests that online parenting interventions may help ameliorate language nutrition differences in early childhood due to barriers in accessing early childhood education provided through preschool or home visiting programs. Baralt et al. (2020) used a phone application designed to teach Hispanic parents the importance of speaking Spanish while delivering language nutrition to their infants. The researchers reported high degrees of feasibility and acceptability and preliminary evidence of effectiveness for increasing child-directed speech. By using an online delivery mechanism, language interventions can reach a broader audience of parents and caregivers who otherwise cannot attend trainings in person. In the development of additional technology-based parent training interventions, researchers should also consider the behaviors to be taught to caregivers.

**Behaviors Taught to Caregivers**

Storybook reading is one of the most common activities recommended by researchers and used by parents to support children’s early language development. There is substantial evidence of the effectiveness of shared reading interventions as implemented by a variety of caregivers, to
include fathers (e.g., Seven & Goldstein, 2020) and extended family members (Coetzee et al., 2021). Meta-analyses examining parent-implemented storybook interventions found statistically significant and positive intervention effects on children’s vocabulary outcomes and language comprehension (Dowdall et al., 2019; Heidlage et al., 2020; Montag et al., 2015; Sénéchal & LeFevre, 2002). Many shared storybook reading interventions involve parent-child dialogue before, during, and after reading, and provide opportunities for children to connect story content to personal experiences. In addition, many storybooks include more advanced or content-area specific vocabulary (e.g., science) that parents would not otherwise use in everyday conversation. Therefore, they are associated with medium to large effects on children’s vocabulary outcomes (Bus et al., 1995; Mol et al., 2008).

The time requirement and availability of a book to read presents a barrier for effective storybook reading intervention uptake and sustainability at a larger scale. Families who live in rural areas may not have access to a local library and families experiencing poverty are less likely to purchase storybooks (Neuman & Celano, 2001; Strasser et al., 2017). Even when books are available, several studies have shown lower rates of shared book reading by parents employed full-time or who have more caregiving demands (Kalil et al., 2016; Karrass et al., 2003; Kuo et al., 2004). A 2019 KIDS COUNT report found that 45% of all surveyed parents in the U.S. read a storybook to their preschoolers less than 4 times per week, with lower rates reported for children identified as American Indian, Latino, or Asian and Pacific Islander (The Annie E. Casey Foundation, 2020). These results are also based on parental self-report of the number of books read or hours spent in reading books with children, an outcome that may be influenced by social desirability bias (Chung & Monroe, 2003; Hammer et al., 2003; Mol & Bus, 2011; Shen & Del Tufo, 2022). Therefore, if storybook reading is not the best contextual fit for
all families, researchers and practitioners may need to consider alternatives for helping parents support young children’s early language development (Hindman et al., 2016). One feasible option is oral storytelling.

**Oral Storytelling**

Storytelling involves a broad range of oral language skills that include vocabulary knowledge, listening comprehension, inferencing, recall, story structure, and morphological and syntactic skills. Narrative language—the sophisticated form of language used when telling stories—plays a pivotal role in kindergarten readiness, as well as the development of reading and writing in later elementary years (Griffin et al., 2004; Lervåg et al., 2018). Decontextualized language used by caregivers during storybook reading or informal conversations during daily routines helps young children develop vocabulary and knowledge of linguistic structure, skills important to reading and writing (Fiorentino & Howe, 2004; Zimmerman et al., 2009). Therefore, use of narratives is frequently recommended to parents and caregivers as a tool to strengthen children’s oral language (Dicataldo et al., 2020).

Practiced for centuries, oral storytelling involves the construction and sharing of an account about real or fictitious events, sometimes conveying cultural practices or rules, morals, and beliefs relevant to the family. Because it does not require physical materials, the portability of storytelling is likely the reason the practice has maintained across civilizations. Storytelling may also persist because of its utility as both a conveyor of information and language teaching tool. Starting around the age of 36 to 42 months, children begin to demonstrate the ability to recall and share previous experiences (Reese, 2018). It stands to reason that parental elaboration during shared storytelling predicts children’s later narrative language, memory, vocabulary, and other literacy skills (Fivush et al., 2006; Reese, 1995). Young children’s vocabulary knowledge
and use of complex grammar in narratives are also related to the lexical diversity and linguistic complexity in caregiver storytelling (Levya et al., 2020).

In addition to portability and utility benefits, many different cultures and languages practice storytelling using a similar formula for key story content (e.g., setting, character, problem, action, ending) and story structure (Spencer & Petersen, 2020; Stein & Glenn, 1978). For example, native speakers of English, Arabic, Hebrew, and Spanish typically use a similar narrative macrostructure, including the order of story elements (Abdallah et al., 2020; Berman & Slobin, 1994; Doan & Wang, 2010; Teepe et al., 2017). Not required to follow a specific text, parents can easily adapt story content to be more culturally relevant or for a more appropriate linguistic fit. A parent may choose to retell a fable using more contextually relevant vocabulary or more common character names. In addition, the linguistic flexibility of oral storytelling allows parents to easily adapt narrative structure to meet the needs of their audience. For example, while stories generally include the same key information (e.g., story parts), a Hearing parent would need to re-arrange the narrative sequence of a spoken story to fit the time-topic-context structure of ASL for their Deaf child. Thus, in the context of early childhood language intervention, narrative intervention is a flexible, adaptable, and familiar practice that may be easily adopted by parents and caregivers of young children.

**Oral Narrative Intervention**

The use of oral storytelling interventions in early childhood research has increased over the past decade as evidence of its effectiveness continues to expand (Pesco & Gagne, 2007; Spencer & Petersen, 2018). A narrative intervention is one that uses oral storytelling, including retelling and generating stories, as the primary teaching procedure (Spencer & Petersen, 2020). In research, oral narrative interventions have improved the oral language and vocabulary skills of
preschool dual-language learners (Petersen et al., 2016; Spencer et al., 2019; Spencer et al., 2020), students attending Head Start (McGregor, 2000; Spencer et al., 2015), and young children at risk for or having language disorders (Favot et al., 2020; Glisson et al., 2019; Hessling & Schuele, 2020).

For families experiencing poverty, oral narrative interventions that target extra-textual conversational behavior (e.g., extending talk beyond the storybook text) support the development of children’s oral narrative language than do interventions encouraging reading alone (Reese, 2018). However, research on the effects of oral narrative interventions implemented by parents and caregivers is sparse. Bailey and colleagues (2020) found that teaching mothers to shape their preschoolers’ oral retells positively predicted children’s narrative structure in first grade. In another study, Garcia et al. (2021) investigated the use of family engagement activities (FEAs) to supplement a classroom-based oral narrative curricula for preschool dual language learners. Caregivers’ brief use of the family-implemented home activities led to statistically significant gains in children’s Spanish vocabulary. However, neither study relied on caregivers to generate stories used to teach narrative skills to children. In addition, while evidence of a causal relationship exists between teacher- and clinician-delivered oral narrative intervention and children’s listening comprehension, vocabulary, reading comprehension, and writing outcomes, we are unaware of research to date that explicitly investigates the use of online programs designed to teach caregivers how to elicit, improve, and encourage preschoolers’ personal narratives.

The purpose of this study was to examine the preliminary effectiveness, feasibility, and acceptability of an online asynchronous parent training program focused on teaching caregivers skills to support children’s oral narrative language development. We designed the intervention,
Tell Me More, based on the hypothesis that caregivers can use oral storytelling to foster children’s oral language skills while not disrupting family routines and without necessitating books or literacy. Given the lack of established online asynchronous interventions for enhancing caregivers’ oral language and conversational behaviors, it was imperative that we approached the study inductively/dynamically by prioritizing an investigation of the intervention’s acceptability and feasibility. Thus, we set out to meet the study objectives by asking the following research questions:

1. What is the preliminary effect of the online asynchronous training package (i.e., online asynchronous training modules and practice activities) on three sets of targeted caregiver behaviors: oral narrative quality, the ability to elicit children’s stories, and skills used to strengthen children’s stories?

2. To what extent do caregivers of preschoolers find the intervention feasible (i.e., ease of access, time to complete; intervention dosage) and acceptable (e.g., relevant and appropriate magnitude of content)?

Method

Participants and Setting

The study was approved by the university institutional review board (IRB) and conducted between September 2021 and November 2021. Caregivers were eligible to participate if they were English speaking adults of preschool-aged children (ages 3–5 years) with at least one personal device (e.g., computer, smartphone, tablet, etc.) capable of recording video and audio, and reliable access to working internet. A working internet connection was defined as an internet connection speed that was fast and reliable enough to watch a YouTube or Facebook video in SD format without buffering. We recruited the participants by posting information in Facebook
parent groups and making public posts on Instagram and Twitter that could be re-shared by anyone online. If interested in learning more about the study, caregivers filled out an online form and gave permission to be contacted by the researchers. The first author contacted eight interested parties to determine eligibility to participate. If they were interested and eligible, the first author held meetings via phone or videoconferencing (e.g., Zoom) to provide further information about the study (e.g., time commitment, activities, data collection and storage procedures) and obtain consent and assent verbally and in writing. Informed consent and assent to participate was obtained through an online portal called REDCap (Research Electronic Data Capture; Harris et al., 2009; 2019). REDCap is a secure, web-based software platform hosted at the University of South Florida that was used to support data capture, management, and analyses for this study.

Of the six caregivers enrolled, five caregivers completed the study. The five caregivers represented four different families, with Family 2 including a triad (mother, father, and child). Remaining dyads were composed of mothers and their children. We collected participant demographics and background characteristics via questionnaire immediately after obtaining signed consent forms. The data collected included caregiver and child names, dates of birth, race and ethnicity of caregiver and child, caregiver’s highest level of education, and family and household characteristics at time of enrollment (e.g., number of persons living in the home and their age groups, zip code of home address). Demographic information is detailed in Table 9. Please note, pseudonyms chosen by the caregivers and their children are used instead of participants’ real names.

After consenting to participation in the study, we assigned the caregivers an alphanumeric study identification number unique to their family to remove any identifiable
information from data repositories. The study identification number included the study code (TMM), study phase (1A), and order of enrollment by date (01, 02, 03…). Consent forms, surveys, questionnaires, and quantitative data unique to each participant was saved with file names unique to their assigned study identification number on a secured university database, as well as in REDCap.

Table 9. Participant Demographics

<table>
<thead>
<tr>
<th>Family</th>
<th>Participant</th>
<th>Role</th>
<th>Gender</th>
<th>Age (years)</th>
<th>Highest degree</th>
<th>Adults in home</th>
<th>Children in home</th>
<th>Area of residence</th>
<th>U.S. region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Talia</td>
<td>Mother</td>
<td>F</td>
<td>38</td>
<td>Masters</td>
<td>2</td>
<td>3</td>
<td>Rural</td>
<td>NE</td>
</tr>
<tr>
<td></td>
<td>Lisa</td>
<td>Child</td>
<td>F</td>
<td>4.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Holly</td>
<td>Mother</td>
<td>F</td>
<td>35</td>
<td>PhD</td>
<td>2</td>
<td>2</td>
<td>Suburban</td>
<td>NE</td>
</tr>
<tr>
<td></td>
<td>Alexander</td>
<td>Father</td>
<td>M</td>
<td>36</td>
<td>JD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oliver</td>
<td>Child</td>
<td>M</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Freya</td>
<td>Mother</td>
<td>F</td>
<td>35</td>
<td>PhD</td>
<td>2</td>
<td>2</td>
<td>Rural</td>
<td>MW</td>
</tr>
<tr>
<td></td>
<td>Rune</td>
<td>Child</td>
<td>M</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Cameron</td>
<td>Mother</td>
<td>F</td>
<td>44</td>
<td>PhD</td>
<td>2</td>
<td>1</td>
<td>Rural</td>
<td>SE</td>
</tr>
<tr>
<td></td>
<td>Christopher</td>
<td>Child</td>
<td>M</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. F = female, M = male; NE = Northeast, MW = Midwest, SE = Southeast

Study Design

We were interested in exploring the preliminary effects of the Tell Me More intervention on parent/caregiver behavior during conversations with their preschool-aged children, particularly the quality of their oral narrative language, rate of story elicitation skills, and rate of skills used to strengthen their child’s stories. Thus, we used a concurrent multiple baseline design across three behaviors with four replications across participants to examine the effects of the Tell Me More intervention on the three caregiver skills (oral narrative language, story elicitation, story strengthening). The design is ideally suited to highlight the effects of a multi-component intervention on three different caregiver conversational skills over an 8-week period (Favot et al., 2020; Tarvainen et al., 2020). In addition, single-case experimental designs are appropriate for
initial intervention feasibility tests, allowing for the systematic identification of implementation barriers and facilitators at the individual case level (Ferron & Scott, 2005).

An overview of the multiple baseline design study activities and related timeline is available in Appendix D. We defined weeks as beginning on Sunday and ending on Saturday. The initial baseline phase was two weeks, having a maximum of six recorded conversations (heretofore referred to as “observations”). We introduced a new training module every two weeks, thus staggering the subsequent baseline lengths accordingly. The third and longest baseline (six weeks) had a maximum of 18 observations. We did not randomize the order of training modules because modules 2 and 3 built on skills learned in the previous module(s).

**General Procedures**

After caregivers completed all study enrollment tasks (e.g., consent forms, demographic questionnaires), the first author met with each caregiver individually via video-conferencing or phone. Prior to the onset of data collection, the first author provided the following instructions to caregivers on how to record themselves in conversation with their child during a typical daily routine: “*On three separate days every week, please record a conversation between you and your child talking about your day or a past event. You can choose any topic or experience to talk about and you can film at any time of day or location. Start filming before you start the conversation and stop the recording as soon as it is over. If your child does not want to participate, don’t worry. We do not want to force them to talk, so if you think you’ve made every attempt possible to get them to engage, feel free to stop the recording. A child’s non-participation is just as important for us to see as the wonderful conversations you’re going to have.*” A two-week baseline phase started on the Sunday following completion of all meetings.
with the caregivers (September 12, 2021). Caregivers submitted videos during baseline in absence of any other contact with the first author or research team.

The intervention phase began at week 3, at which time participants received an individualized email from the REDCap portal that contained the following: basic instructions for accessing and starting the training, a statement reporting the duration of time the module would be available, a general reminder to continue to submit three recorded conversations every week, a link to an end-of-module feedback survey, and hyperlinks to access the first training module and first set of three family activities. Distribution of the email using the REDCap survey feature allowed us to monitor parent access to the training (when they clicked on the hyperlinks) and avoid sending more than one email per week by simultaneously including the end-of-module feedback survey.

Each even week thereafter (weeks 4, 6, and 8), caregivers received an email with hyperlinks to a new set of family activities, a gentle reminder to continue to submit videos, and instructions to complete the end-of-module survey if they had not already done so. On each Sunday in the remaining odd weeks (weeks 5 and 7), the caregivers received an email similar to the email used in week 1, with the exception of general instructions that were no longer included.

The intervention pace was one training module per every two-week period. To hold intervention dose consistent across participants, we terminated access to a prior module when a new training module was introduced by manually changing its associated hyperlink. The intervention ended on the Saturday of week 8. Using the same instructions as in baseline and treatment conditions, we asked the caregivers to provide additional conversational recordings during weeks 9 through 11 to examine the degree to which they maintained skills in absence of the intervention.
Independent Variables

The primary purpose of the study was to test the initial effectiveness, acceptability, and feasibility of the new online caregiver training, *Tell Me More*. The asynchronous intervention was designed to teach caregivers about the benefits of shared oral storytelling, what makes a good story (e.g., story order and key parts), and strategies to support their child’s oral narrative language development. The independent variable had two components: online asynchronous training modules and practice activities based on the “introduce, illustrate, and practice” model (Snodgrass et al., 2017) that we characterize in Table 10, below, as “Learn, Watch, Do.”

Table 10. *Tell Me More* Intervention Modules and Related Ingredients

<table>
<thead>
<tr>
<th>Module</th>
<th>Learn</th>
<th>Watch</th>
<th>Do</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Making Sense with Stories</strong></td>
<td>(1) Introduction and overview</td>
<td>(8) Video of an adult and child telling personal stories</td>
<td>(4-6) Interactive activities to practice new material</td>
</tr>
<tr>
<td></td>
<td>(2) The importance of storytelling</td>
<td></td>
<td>(8) Check off all the elements as you listen to their stories</td>
</tr>
<tr>
<td></td>
<td>(3) How stories are used to make sense of experiences</td>
<td>(9) Parent child storytelling interaction</td>
<td>(11) Interactive activity: Steps of parent-child storytelling</td>
</tr>
<tr>
<td></td>
<td>(7) Developmentally appropriate story elements</td>
<td>(10) Video of parent and child telling reciprocal stories</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9) Parent child storytelling interaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12) Wrap up</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Starting Conversations with Stories</strong></td>
<td>(1) Quick review of Module 1 and overview of Module 2</td>
<td>(7) Video of parent using the “T” strategy</td>
<td>(4-5) Interactive activities to practice new material</td>
</tr>
<tr>
<td></td>
<td>(2) Less responsive children and what it could mean</td>
<td>(9) Video of parent using “T” &amp; “O” strategies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Strategies for starting conversations (TOSS)</td>
<td>(11) Video of parent using “T”, “O” &amp; “SS” strategies</td>
<td>(12-14) Interactive activities to practice the three strategies</td>
</tr>
<tr>
<td></td>
<td>(6) Tell me strategy (T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8) Open-ended questions (O)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10) Story share (SS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(15) Wrap up</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Strengthening Children’s Stories</strong></td>
<td>(1) Quick review of Module 2 and Intro to Module 3</td>
<td>(7) Video of parent recasting</td>
<td>(4-5) Interactive activities to practice new material</td>
</tr>
<tr>
<td></td>
<td>(2) Responsive parenting and why it is important</td>
<td>(9) Video of parent improving</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Skills for strengthening conversations (RIVER)</td>
<td>(11) Video of parent validating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6) Recast</td>
<td>(13) Video of parent elaborating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8) Improve</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10) Validate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12) Elaborate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14) Relate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(20) Wrap Up</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tell Me More Modules. Tell Me More module content includes textual, visual, and auditory stimuli. Behaviorist theory guided the identification of intervention ingredients based on the notion that language is an operant behavior improved through systematic instruction. In addition, the design framework of Tell Me More was informed by adult learning theory (Knowles, 1975) that emphasizes use of situational-based and practice-oriented instructional technologies. Thus, training modules followed the “Learn, Watch, Do” format and allowed caregivers to acquire knowledge through self-guided learning opportunities available through Tell Me More’s asynchronous online delivery. Both theories supported our purposeful use of a pedagogical agent as the narrator and placement of interactive activities interspersed within the overall informational content (Morris & Bellon-Harn, 2021). Table 10 includes a general overview of the Tell Me More content.

Tell Me More’s primary intervention ingredients include three asynchronous training modules accessed online at no cost to caregivers of preschool-aged children. We arranged the order of training so that each subsequent module builds upon each previously taught skill set. An asynchronous format was used to allow caregivers a judgement-free opportunity to stop and start the training on several different occasions, given the nature of the home environment or competing and unpredictable events. Each of the three training modules, described in detail in the following section, lasted no longer than 10 minutes in duration.

Module 1: Making Sense with Stories. In addition to the history of storytelling and its importance in many cultures, the ingredients of a good story were described in the first Tell Me More module. We designed the module to teach caregivers the necessary parts and sequence that define a complete and cohesive story: setting, character, problem, feeling, action/attempt, consequence, ending, and end feeling.
Module 2: Starting Conversations with Stories. In the second training module, we briefly reviewed key information from Module 1 prior to teaching caregivers a series of methods, using the acronym TOSS, to engage children in a storytelling conversation. The acronym TOSS refers to the following strategies: presenting a “Tell me about” directive (T) to elicit a story from their child. The “Tell me about” directive is an important skill for caregivers to learn because it establishes the child’s role as a leader in the construction of a story and elicits a story from the child’s perspective. When they follow their child’s interests in remembering and telling a story about a past event, caregivers support the development of social emotional skills and secure attachment (Cleveland & Morris, 2014; Cleveland & Reese, 2005). However, if children are unresponsive to the directive to tell a story, Module 2 covered the use of open-ended questions (O). The use of open-ended questions can increase the likelihood that children will remain engaged in and extend their story in conversation with a caregiver (Khan et al., 2014).

If the previous two strategies do not result in a story from the child, the caregiver was given information about the importance of sharing their own personal story. By telling a personal narrative, caregivers can reduce the response effort required by children not yet proficient with storytelling (Gentile, 1972). In summary, the purpose of the TOSS strategies was to provide caregivers with skills that encourage children to tell their own stories after caregivers have provided appropriate models (as learned in Module 1). Story elicitation was key to the caregivers’ use of skills presented in Module 3.

Module 3: Strengthening Children’s Stories. The content in the third and last module included a set of responsive behaviors intended to strengthen (i.e., keep it going and make it better) a child’s oral narrative. In general, the module introduced the importance of learning to wait, listen, and then respond to a child’s utterances in a manner that increases the likelihood of
subsequent conversational turns and future storytelling interactions for child language improvement (Dunst et al., 2012; Roberts et al., 2019). More specifically, caregivers learned five specific responding techniques that occur in close temporal proximity to a child’s utterance: repeat, improve, validate, encourage, and relate the child’s content with experiences of their own (RIVER; See Appendix E).

The first strategy, repeat, allows a caregiver to confirm or acknowledge a child’s production by repeating what the child just said with point-to-point correspondence. Research has shown repetition to be beneficial for child language development, as the behavior prevents caregivers from interrupting a child’s story to correct their speech (Camarata & Nelson, 2006; Leonard, 2011). Non-corrective repetitions do not need to be intrusive to be effective for improving children’s overall language development; rather, effective repetitions should maintain the basic meaning of the child’s emission (Cleave et al., 2015). However, when children make errors or have inaccuracies in their remembrance of a shared event, caregivers were trained to use “improving” statements. Improving statements generally entail the caregiver increasing the linguistic complexity or accuracy of the utterance (Fivush et al., 2006; Reese, 1995). These contributions are a form of error correction after the child has completed their turn, in which a caregiver corrects grammatical mistakes or inaccurate information. For example, if the child said, “He go dare,” the caregiver would wait until the child finished talking to respond, “Yes. He went over there.” To correct the accuracy of a child’s contribution without interrupting the flow of the conversation, like if the child misremembered a character name or used the wrong pronoun, the caregiver could use the improving strategy to present the correct information as unobtrusively as possible. The caregiver could ask questions like, “Do you mean Miss Sally?” or reply, “Oh, you mean Miss Sally took you to the bathroom.” In keeping with the goal of a fluid
and enjoyable conversation, the module also discourages caregivers from requiring the child to repeat back their improved utterance. In addition, caregivers may elaborate on their child’s contribution to a story. Elaboration is defined as any caregiver verbal response that extends the child’s storytelling by elaborating upon their utterance with new information, advanced vocabulary, or complex language. Caregivers could also demonstrate an elaboration if they answered a child’s question (“But was he there, too?”) with new information, “He was at the party, but only briefly. Grandma and Papa were also invited, although they could not come because they were in Florida.” In such example, the caregiver answered the question but also added more information and complex language (i.e., advanced vocabulary: briefly, invited; causal ties: although, because). Elaborations motivate or prompt the child to add on more information in their next utterance while also preventing the caregiver from quizzing their child (“Who else was there?”).

Similar to the use of repetition, caregivers can also validate their child’s input during storytelling interactions. When caregivers validate the child’s contribution by making affirmative comments, they are likely to reinforce the child’s current behavior and increase the number of conversational turns to follow, as well as increase the likelihood children will continue to share stories. For example, in response to a child’s shared story about getting hurt when falling outside, the parent could reply, “Yeah, you were sad when you hurt your knee.” In addition, children will usually focus on aspects of a story that they find most interesting. By delivering validating statements, the parent can support the child’s view and be sensitive to features of the remembered event that the child is most interested in discussing. In contrast, caregivers who force children to remember specific parts of a story and correct the child frequently will terminate the storytelling interaction prematurely.
Rather than terminate the story interaction, in Module 3 caregivers learned to make relating statements as a method of further children’s engagement in storytelling. Relating in conversation is similar to text-to-life connections made by a parent during shared storybook reading (Grolig, 2020), where caregivers demonstrate a connection between the child’s experience and similar experiences of their own—present day or when they were a child. This may happen through the caregiver’s sharing a portion of their own story, but with the intention to make it relevant to the child’s experience. For example, when talking about a severe weather event experienced the night before, the caregiver may relate by saying, “I was so scared by that lightning last night! But, I felt safe because you were there.” In doing so, the caregiver models social-emotional vocabulary and attempts to get the child to engage in the construction of a narrative centered on their shared experience. When relating the child’s experience to one of their own (fictional or factitious), caregivers can also provide additional information related to problem solving or dealing with challenges. In summary, Module 3 capitalizes on content taught in the preceding modules. After learning the importance of oral narrative language, what makes a good story, and how to increase storytelling opportunities, caregivers learn to improve in their child’s use of story structure, content, complex language, and extend mean length of story productions naturally and authentically. Given increased opportunities to tell stories through use of TOSS skills, RIVER skills allow caregivers to take an active role in the shaping of their child’s oral language skills.

**Family Activities.** Every week during the treatment phase, caregivers received a set of three hyperlinks to access the optional family activities, which expired after two weeks. During those two weeks, however, caregivers could access and complete the activities with their child as often as they wanted. The six sets of family activities were designed and executed through
Boom! Learning (www.boomlearning.com), which is a digital platform for extending practice of instructional content outside of classrooms. While most Boom! decks are designed by teachers for elementary-aged students, its flexible digital design capabilities and hyperlinking to the lessons make Boom! applicable to any type of learner. For the Tell Me More family activities, we adapted the instructional design that has been successful in a school-based, non-digital, oral narrative language program (i.e., Story Champs; Spencer, 2021; Spencer & Petersen, 2016) for a digital, parent-child version of oral storytelling activities. We crafted the activities to include developmentally appropriate vocabulary, story content and length, and follow-up questions to foster child storytelling with increasing independence and complexity. Each of the six sets of family activities contained three different lesson types (A, B, and C) centered on the same narrative. Informal pilot testing of the family activities by the first author’s family and friends resulted in an estimated average completion time of 8 minutes per lesson, although the duration depended on the amount of caregiver support needed for children to complete activities like story retell.

The first lesson type (Activity A) in each set involved the caregiver navigating through a series of illustrations and reading aloud a story script. A set of five icons accompanied the illustrations and script to emphasize the key elements of a structured narrative, such as character, problem, feeling, action, and ending. After reading the story, the caregiver asked the child a set of prepared questions about the story (e.g., “Who was the story about?” “What was the character’s problem?” “How did the character feel about their problem?”). If the child was unable to answer a question, the caregiver had the option to select a HELP button that would take them to another screen with the answer and instructions for the caregiver to read aloud the
answer, point to the illustration provided, and help the child repeat the answer to the question
given the caregiver’s model.

The next activity and lesson type (Activity B) focused on guiding the caregiver in helping
their child retell the story with fading verbal and visual prompts. Activity B used the same story
presented in Activity A and required the caregiver to re-read the story with the child and then ask
the child to retell the story using the illustrations and icons on the screen. Like Activity A, a
HELP button was available in the top left of each page to be accessed as the caregiver needed.
The HELP screen provided caregivers with a script to model the correct response, gesture to the
illustration, and encourage the child to repeat the modeled response.

After telling and retelling the same story several times in Activities A and B, the last
lesson type focused on social-emotional vocabulary presented in the story and the child
generating their own story. In Activity C, caregivers retold the story with all illustrations and
icons provided, but focused conversation on the main character’s emotional reaction to the
problem. Scripts were given to the caregiver to read aloud to prompt the child to make inferences
about the character’s feelings using the targeted social-emotional vocabulary word, as shown in
Figure 4. If the child was unable to define the word, a HELP button on the screen was available
to assist the caregiver to support the child as needed. After asking about the social-emotional
vocabulary word, the caregiver asked the child to make predictions about what the character
would do in the future when they felt similarly. Finally, children were asked to tell a story about
when they experienced something like that. As they generated their personal story, children used
a graphic organizer in the Boom! lesson to move the story icons into place, much like simple
checklist. We purposefully wrote personally-relevant stories for the activities so that caregivers
could tailor content of the child’s story in Activity C to be more culturally or contextually
relevant to their family experiences, without changing the sequence and structure of the narrative.

Figure 4. Example: Screen Capture from Family Activity 2C

Data Collection

Caregivers recorded conversations with their children and sent them to the first author immediately thereafter. Although the intent was to capture conversations on video, some caregivers submitted audio recordings instead. The parents reported it was necessary to limit their children’s distractibility. Given the dependent measures are all based on what the children and caregivers said, audio recordings were considered an acceptable alternative to videos. After caregivers submitted the recordings, we referenced the file properties to confirm the recording date. This was done to rule out the risk of two or more observations occurring on the same date. If a caregiver had not submitted a recorded conversation by Wednesday evening of each week,
the first author used response prompts or “nudges” (Tagliabue & Simon, 2018; Tagliabue, 2022) to gently remind them to record and submit the recordings by the end of the week.

**Transcription**

The first author trained three undergraduate research assistants who were unaware of the study aims and conditions in transcription procedures prior to the start of the study. The study team transcribed the caregiver-child conversations verbatim from the provided recordings into MS Word. To protect confidentiality and mask study condition when scoring, the first author removed or replaced any identifying information contained within the transcription (e.g., names, locations) with secondary codes. Each transcribed language sample included the participant’s study identification number, the corresponding audio/video file name, and original date, duration, and location of the recording. In addition, the documents included the initials of the person who transcribed the sample and date of transcription.

Following training but prior to working on study recordings, reliability and fidelity measures were documented. The research assistants’ average word-for-word transcription reliability was 97.8% (range: 96.7%–99.1%). Fidelity to transcription conventions was 87.9% as measured by a fidelity checklist (range: 84.2%–92.3%). In addition, during the study, the first author randomly selected 20% of each participant’s samples to assess transcription fidelity and reliability measurement \( n = 22 \). Transcription fidelity was 93.3% on average (range: 81%–100%). Word-for-word transcription reliability was 95.0% on average (range: 85.1%–99.7%).

**Outcome Measures**

**Effectiveness.** Following completion of transcription fidelity and reliability assessments, the first author served as the primary scorer and used the transcribed language samples to code and quantify the three targeted verbal behavior outcomes at the caregiver level: oral narrative
language quality, rate of story elicitation skills, and rate of skills used to strengthen children’s stories. We limited examination of treatment effects to caregiver behavior based on the assumption that the Tell Me More intervention must affect adult language outcomes before we would detect child language effects. We provide definitions for each outcome related to treatment effects and their measurement procedures below. A sample scoring record form used in the study is provided in Appendix F.

**Oral Narrative Language.** We defined caregiver oral narrative language as the quality of caregiver storytelling based on the presence of story parts and use of complex language. The first author examined each caregiver’s conversational contribution and used a practitioner-friendly rubric, the Narrative Language Measures (NLM) Flow Chart (Petersen & Spencer, 2020), to quantify the caregiver’s overall narrative quality per language sample. While the NLM Flow Chart has only been used to measure children’s oral or written narrative language (e.g., Kirby et al., 2021; Spencer et al., 2014; Spencer & Petersen, 2018), we determined the NLM Flow Chart to be the best available tool for measuring narrative structure and language complexity in adult oral language samples. As provided in Appendix A, the NLM Flow Chart contains three sections, Writing Conventions, Narrative Structure (NS), and Linguistic Complexity (LC). For this study, we only used the NS and LC sections of the NLM Flow Chart to reflect the quality of discourse-level completeness and sentence-level complexity, respectively.

The NLM Flow Chart (Petersen & Spencer, 2020) is organized in a decision-tree format that allowed for quick scoring of the NS and LC subsections. There were 9 items within the NS section of the NLM Flow chart representing narrative completion and coherence. For each item, we used the flow-chart decision rules to measure the presence and quality of each of the following key story elements: character, setting, problem, sequence, plan/attempt, consequence,
ending, emotion, and episode complexity. Based on the degree to which each was present in the caregiver’s language sample, we used the NS rubric to assign a score to each category. NS scores of 0 indicated the absence of a particular story element, whereas scores of 3 (or higher) represented the caregivers’ inclusion of a clear and complete story part. The range of scores possible for the NS section was 0 to 33. The LC section of the NLM Flow Chart represented 6 indices of sentence-level features that appeared within a language sample. Whereas conventional use of the LC subsection assigns a rating of 0–3 similar to the NS section, we ruled out the chance of a ceiling effect by reporting count, a continuous variable, for each category. The first author visually inspected each transcribed language sample and counted each occurrence of the following six indices: relative pronouns, verb/noun modifiers, advanced vocabulary (e.g., similes, metaphors), temporal ties, causal ties, and character dialogue.

The first author also counted the frequency of social-emotional vocabulary in each sample. This additional measure was taken to see if the caregivers generalized social-emotional vocabulary purposefully embedded in the family activity narratives to use in conversations with their child. The count of social-emotional vocabulary in the sample was added to the NLM Flow Chart score. The NLM Flow Chart and social-emotional totals were summed to report a single outcome representative of caregiver’s oral narrative language quality.

**Conversational Turns.** Anticipating variability in conversational length and turns within and across participants, we reported conversational turns as a continuous outcome and used the variable to calculate story elicitation and story strengthening skills as rates per conversational turn. We defined conversational turns as beginning with the first speaker, either a caregiver or child, directed at the other partner. With the exception of Family 2, conversational exchanges between caregivers and other adults or siblings not participating in the study did not count as
conversational turns in the caregiver-child conversation. For Family 2 (Holly and Alexander), we viewed the parents as one caregiving unit in the conversational relationship with their child rather than unique conversational partners. The parents specifically requested this arrangement during the initial meeting with the first author because they believed having conversations about their days was an important part of their family routine in the afternoons and evenings.

**Story Elicitation.** Appendix G includes operationalized definitions of the individual *Tell Me About* (T), *Open-ended Questioning* (O), and *Personal Story Share* (SS) components taught in Module 2. Story elicitation was a composite of three specific strategies and represented by the acronym TOSS. The first author used a researcher-developed rubric (see Appendix G) to measure the extent to which caregivers used any or all strategies within the transcribed language samples. The first author visually inspected each transcription and counted the number of TOSS sub-skills observed within the sample. For example, during examination of a transcription, the first author counted every appearance of the phrase “Tell me…” and recorded the total count in the “T” column of the scoring record form (see Appendix F). The same process of sample examination was used to report a total count for open-ended questions and personal story shares. The totals were summed across the three strategies to report a cumulative TOSS score, a continuous variable, per transcription. For the purposes of visual analysis and statistical analyses, we divided the total count by the total number of conversational turns per sample to report the rate of TOSS skills used per observation.

**Story Strengthening.** The first author also examined each transcribed language sample to measure the extent to which caregivers used any or all RIVER strategies during the conversation, using a researcher-developed rubric shown in Appendix E. The five strategy subtotals were summed and recorded on the scoring record form as a continuous variable (see Appendix F). The
story strengthening outcome was reported as a rate of skills used per conversational turn per sample; the total count of all skills observed across the sample was divided by the total number of conversational turns per sample.

**Scoring Reliability.** We randomly selected 20% of all samples to assess scoring reliability \( n = 20 \). An independent research assistant, unaware of the study condition, conducted secondary scoring for the purpose of item-by-item reliability analyses. The primary and secondary scorer met once to reconcile any differences in scores to the extent possible. If consensus was not reached, the primary score was used for reporting outcomes. Across selected samples, the average item-by-item agreement was 88.9\% (range: 83.7\%–96.5\%). The average difference in conflicting scores per item per sample was 0.46 (range: 0.08–0.80).

**Feasibility and Acceptability.** Feasibility measures included per-participant tracking of module and family activity completion, as well as attrition. This is based on the notion that if the Tell Me More intervention is feasible, caregivers will be able to complete all intervention components and remain in the study. Feasibility data were collected through participation logs. To measure intervention acceptability, we obtained caregiver opinions regarding Tell Me More during and after the intervention using three end-of-module feedback surveys and an end-of-study questionnaire. In addition, all spontaneous communication received from caregivers during the study (i.e., unsolicited feedback shared with the first author via email or text message) was de-identified and added to the participant’s record in REDCap. We used these qualitative data to supplement quantitative measures of intervention feasibility and acceptability.

**Participation Logs.** We used background data in REDCap documenting caregivers’ opening of emails distributing training module and family activity hyperlinks and end-of-module survey completion as evidence of module and family activity completion. We also collected
information on voluntary participant withdrawal (i.e., attrition) to include rate and reason for
drop-out. We defined attrition as occurring when multiple (3 or more) communicative attempts
using the participant’s stated preferred method(s) of contact are unsuccessful, resulting in the
absence of data collection during the study timeline (e.g., no videos submitted, surveys not
completed, etc.). Prior research demonstrates an attrition rate as high as 70% for asynchronous
parent training programs (Doty et al., 2016; O’Brien et al., 2012), but we expected to see a lower
rate of attrition in this study because Tell Me More was designed so that caregivers could access
it at any time of day within a two-week window.

End-of-Module Feedback Survey. Caregivers completed a six-item feedback survey at
the end of each training module (weeks 4, 6, 8). The surveys were designed and distributed using
REDCap alongside the training and family activity links so that caregivers would immediately
provide feedback after module completion. Caregivers could complete only one survey per
module (i.e., no duplicate entries). For each module-specific survey, caregivers reported whether
the online training: 1) was easy to access, to include a working hyperlink and ability to start the
training without issue, 2) was easy to complete in one sitting, 3) had good examples and videos,
and 4) was easy to relate to (relevant to them and their child). The responses were reported as
binary (yes/no) outcomes. Two open-ended items allowed the caregivers to provide
individualized feedback on problems or issues experienced while completing the training module
and recommendations for improving the modules. However, caregivers only completed surveys
after completing the module. Therefore, in the event caregivers encountered barriers to accessing
a training module, we instructed them to report the problem immediately (in-situ) to the first
author using their preferred method of communication (e.g., email or text message). The first
author documented all reported issues and worked with the development team to immediately
make changes to allow successful training navigation. We resolved any issues related to intervention access and completion in less than two hours.

**End-of-Study Questionnaire.** We also obtained an overall measure of *Tell Me More* acceptability using a 19-item end-of-study questionnaire. We sent all questionnaires via email using the survey function in the REDCap project database, where results were automatically stored in the password-protected account. The end-of-study questionnaire included 17 Likert items (scale: 1–5) asking caregivers to reflect on the knowledge and skills taught, as well as rate the degree to which they thought *Tell Me More* was effective in changing their behavior. The questionnaire also included two open-ended items that allowed caregivers to provide individualized feedback about how researchers could improve module and family activity ingredients.

**Data Analysis**

Our primary method of analyzing the study outcomes was through visual analysis of single case data. Since this is the first study to date examining the effects of *Tell Me More* on caregivers’ behavior change in conversation with preschoolers, we made two different assumptions about treatment effects. For the oral narrative quality and story elicitation baseline phases, we assumed no trend and an overall low level of responding, but we expected to see high degrees of variability due to the lack of constraints on the setting, duration, and topic of conversation selected by the caregivers. However, we expected to see a positive change in level and trend from baseline to treatment, although we hypothesized that high degrees of variability would persist. For the third and longest baseline phase, we predicted a positive trend as caregivers and their children became more comfortable having, and to some extent, recording their conversations over time. We assumed the upward trend would continue into the treatment
phase. However, like the assumptions made for oral narrative quality and story elicitation skill sets, we expected a change in level of responding after introducing the last training module, in addition to persistent variability across observations in both phases.

**Hierarchical Linear Modeling (HLM).** Following inspection of raw data and visual analyses of participant multiple baseline graphs, we supplemented the results using statistical analyses to further examine the degree to which any change in level, trend, or variability occurred within and across participants for each dependent variable. We used hierarchical linear modeling (HLM) design and analysis recommendations for single-case experimental design research (Rindskopf & Ferron, 2014; Van den Noorgate & Onghena, 2007). Given predicted variability in effects within and across participants, we used a two-level HLM with repeated observations nested within participants (i.e., cases). Consistent with visual inspection of raw data and graphical displays, as well as our predictions about participants’ response to intervention, two different 2-level HLMs were specified so that parameter estimates provided an average treatment effect, as well as estimates of variability within and between cases. We defined the level-1 models, specified below, to predict the interrupted time series effects at the individual caregiver level using restricted maximum likelihood (empirical bayes) estimation. Given the small sample size ($n = 4$), we specified each level-2 model to allow for examination of overall between-participant (group) treatment effects using the Kenward-Roger method of estimating degrees of freedom and assumed positive autocorrelation among the observations across phases and participants autocorrelation to obtain the best possible fit (Ferron et al., 2009). Finally, we centered the time variable for oral narrative quality, story elicitation, and story strengthening outcomes so that the intercept represented the estimated shift or change in responding from baseline to treatment phase.
For oral narrative language and story elicitation skills, we defined the model based on the assumption of no trend in baseline and an immediate change in level and trend with the introduction of the treatment phase. For oral narrative language \((Y_{ONL})\) and story elicitation \((Y_{SE})\) outcomes, the full models were specified as:

\[
Y_{ONL} = \beta_{00} + \beta_{10} \text{Phase}_{t} + \beta_{20} \text{Phase}_{t} \text{Time}_{t} + r_{0t} + r_{1t} \text{Phase} + r_{2t} \text{Phase}_{t} \text{Time}_{t} + e_{t}
\]

\[
Y_{SE} = \beta_{00} + \beta_{10} \text{Phase}_{t} + \beta_{20} \text{Phase}_{t} \text{Time}_{t} + r_{0i} + r_{1i} \text{Phase} + r_{2i} \text{Phase}_{t} \text{Time}_{t} + e_{ti}
\]

For story strengthening skills, we assumed that caregivers had some degree of skills already within their repertoire, like repeating or validating. In addition, we expected an increase in the display of individual story strengthening skills as families became comfortable recording themselves on video. Therefore, we defined the HLM based on the assumptions of a positive trend in baseline continuing in the treatment phase with a noticeable increase in level change following the introduction of the third training module. The full model for story strengthening skills \((Y_{SS})\) was specified as:

\[
Y_{SS} = \beta_{00} + \beta_{10} \text{Phase}_{t} + \beta_{20} \text{Time}_{t} + r_{0i} + r_{1i} \text{Phase} + r_{2i} \text{Time}_{t} + e_{ti}
\]

**Results**

*Conceptual Analysis and Visual Inspection*

We used participant graphs shown in Figure 5 for visual analysis of treatment effects. Each column contains a panel of three graphs representing observed outcomes for a particular participant, starting with Talia on the far left and Cameron’s graphs on the far right. Each panel row represents the three different outcomes measured during the study. Scores representing the quality of narrative language used by caregivers per observation appear in the top row, with the y-axis scale ranging from 0 to 104. The second (middle) row of graphs have baseline lengths of four weeks and report caregivers’ rate of story elicitation skill use per conversational turn, with
the y-axis ranging from 0 to 1.25. The third baseline (bottom row) graphs depict caregivers’ rate
of story strengthening skills per conversational turn, with the y-axis scaled from 0 to 3.0. For all
graphs, the x-axis reflects dates across the study in two-week increments. The first observation
and last observations for all participants occurred on 11 September 2021 and 25 November 2021
respectively. We provide an overall summary of the preliminary experimental effects of the
intervention below, followed by detailed analysis of within-participant effects. Individual
multiple baseline design graphs for each participant appear in Appendix I.

**Oral Narrative Language.** All participants responded positively to treatment, with
average scores in treatment higher than baseline means (see Table 11), but visual inspection of
the graphs reveal variability in the amount and quality of oral narrative language used by
caregivers in conversation with their children prior to, during, and after the intervention. We
predicted that adults would use some oral narrative skills during the baseline phase, but at low
levels without a trend. However, within- and between-participants, baseline data were not
consistently stable, and trends varied across participants. For example, as shown in Figure 5 (top
row), Talia had relatively low and stable oral narrative skills in baseline whereas Holly and
Alexander had a downward trend in scores over time in the baseline phase. Only graphs for Talia
and Holly and Alexander aligned with our a priori assumptions of positive level change and
trend in caregivers’ oral narrative language quality during the treatment phase. After introduction
of the second training module, variability in responding persisted although most participants
demonstrated a general improvement in their average oral narrative quality score per
observation. However, we observed little to no experimental effects for Freya and Cameron
because of the overlap between narrative scores in baseline and treatment phases.
**Story Elicitation Skills.** Most caregivers also used few story elicitation skills prior to training introduced during weeks 5 and 6. As shown in Table 11, baseline observations across participants varied around a mean of .34 ($SD = .14$). After the introduction of training in story elicitation skills, caregivers’ rates increased to an overall mean of .55 ($SD = .34$). Nonetheless, clear causal effects of training on rates of story elicitation are visually apparent for only two of the four participants.

**Story Strengthening Skills.** Table 11 reports the means and standard deviations of caregivers’ skills used to strengthen children’s oral narrative language. Across participant graphs in Figure 5, data reflect an upward trend in caregivers’ use of story strengthening skills over time, to include baseline observations. The baseline trends extend into the treatment phase for most participants resulting in a high degree of data overlap between adjacent phases, weakening any argument for an overall experimental effect of the training on caregivers’ story strengthening skills. The results of within-participant visual analyses and descriptive statistics are detailed below.

**Individual Participant Results**

**Family 1: Talia (mother) and Lisa (daughter).** Talia reported difficulty attempting to engage their child, Lisa, in talks about their day or shared experiences during the baseline phase. Visual analysis of graphs in Figure 5 show low and stable patterns in oral narrative language quality across these five baseline observations. After the introduction of the first training module, there was a shift in both the trend and level of oral narrative language quality. The first treatment observation overlapped with baseline, but an upward trend of oral narrative quality in all subsequent observations. Talia also had low and relatively stable story elicitation rates during baseline, but unlike oral narrative skills, visual analysis of the second baseline graph showed an
immediate and large shift in level and no overlap across phases. Due to an upward trend that begins in baseline and extends into the treatment phase, there were no discernable effects of the intervention on Talia’s use of skills to strengthen Lisa’s storytelling.

**Family 2: Holly, Alexander (mother, father) and Oliver (son).** Family 2 was comprised of two parents, Holly and Alexander. Four of the 26 samples are dyadic conversations between Holly and her son, whereas 22 of the 26 samples are triadic conversations that include Alexander. There was no significant difference in the duration or number of conversational turns between dyadic and triadic conversations. Visual analysis of Figure 5 show that prior to the intervention, the quality of oral narrative language used by both caregivers was variable with scores ranging from 18 to 63. Whereas oral narrative quality had a decreasing trend in baseline, an upward trend occurred after introduction of the first training. However, Holly and Alexander’s conversations varied in the quality and quantity of narrative skills used during treatment observations and 82% of scores overlapped with baseline. Similarly, the intervention had no visually discernable effects on improving their rate of story elicitation skills, with low and relatively stable trends in baseline carried over into the treatment phase. Holly and Alexander’s rate of story strengthening skills varied around a baseline average of 1.7 skills per conversational turn, with a declining trend in rate of use across the baseline phase. On the second treatment observation, Holly and Alexander’s rate of story strengthening skills sharply increased and remained at levels greater than baseline during both treatment and follow up. Only two treatment observations overlapped with the adjacent phase, suggesting the intervention may have resulted in desired improvements in story strengthening skills.

**Family 3: Freya (mother) and Rune (son).** Freya’s multiple baseline design graph appears in Figure 5, second column from right. The data indicate a pattern of improving skills
over time with upward trends extending from baselines into treatment phases. In baseline, Freya had narrative quality scores ranging from 14 to 50 on the first and last baseline observation, respectively. The trend leveled off in the treatment phase with Freya’s oral narrative quality scores varying around a mean of 44.6, but a large amount of treatment observations overlapped with those in baseline. Like other participants, visual inspection of Freya’s story strengthening skills revealed a pattern of increasing rates over time, but the average rate increased from a baseline mean of 1.18 to 1.66 in treatment, as reported in Table 11. Due to a large amount of overlap between adjacent phases, we could not rule out the null hypothesis of no treatment effect for any of the three outcomes.

**Family 4: Cameron (mother) and Christopher (son).** Visual inspection of the graphs indicate trends in baseline for all three outcomes, and immediate level change in the treatment phase for both oral narrative quality and story elicitation skills. Addressing the first outcome, the top right graph in Figure 5 shows an upward trend in Cameron’s quality of oral narrative language prior to training, but a large degree of variability and a slight downward trend as time progressed throughout the treatment phase. Yet, at follow-up, Cameron’s oral narrative quality score was the highest in the study at 104. Similarly, the largest rate of story elicitation skills in the study was Cameron’s rate of 2.10 which occurred in the treatment phase. However, treatment observations before the outlier were relatively stable above baseline level. Observations later in the treatment phase returned to baseline levels (see Table 11), but trend upward closer to the end of the study. Examination of the story strengthening graph shows a gradual increase in rates with the trend extending into the treatment phase and at follow up. While the intervention did not negatively affect outcomes, a large degree of overlap and variability prevent the suggestion of any desirable effect of the intervention on Cameron’s targeted conversational behaviors.
**Multilevel Model**

Careful visual analysis of the graphical display paired with the conceptual model informed the HLM specified for the analyses. Estimated autocorrelations for each outcome ranged from .003 (story elicitation) to .23 (story strengthening) and were non-significant (\(p\)-values \(\geq .10\)). We report parameter estimates of individual and average between-case treatment effects by outcome in Tables 12–14. The Intercept parameters correspond to baseline levels. Phase parameter estimates can be interpreted as treatment effects. For oral narrative language and story elicitation outcomes, the Phase \(\times\) Time interaction refer to the degree of slope change associated with the introduction of the Tell Me More intervention. We centered the time variable based on the first treatment phase session, so the Time parameter estimates for the story strengthening outcome report the trend in baseline expected to continue into the treatment phase.

**Oral Narrative Language.** For the oral narrative language outcome, individual baseline levels varied across participants, with Talia reporting the lowest level (14.84) while all other participants had baselines higher than the overall group mean, \(\beta = 26.68\) (\(p = .02\)). While individual baseline differences were not statistically different from the group mean, the results indicated that the baseline variable was a significant predictor in the two-level model (see Table 12). All participants had a positive and large degree of change in level in response to the intervention, with an average increase of 16.16 in the level of oral narrative quality and individual effects ranging from 12.12 (Talia) to 22.51 (Cameron). However, neither individual nor group treatment effects were statistically significant (\(p = .0634\) for group; \(p\)-value range of .64–.99 for participants). While the average change in slope at the onset of the treatment phase was positive and small, \(\beta = .46\), it was also non-significant (\(p = .34\)). Individual participant slopes did not vary significantly from the group mean.
Table 12. Individual and Group EB Estimates of *Tell Me More* Effect on Caregiver Oral Narrative Language

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Caregiver(s)</th>
<th>Intercept</th>
<th>Phase</th>
<th>Phase*Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral narrative</td>
<td>Talia</td>
<td>14.84</td>
<td>12.12</td>
<td>0.46</td>
</tr>
<tr>
<td>language</td>
<td>Holly &amp; Alexander</td>
<td>29.81</td>
<td>16.34</td>
<td>0.46</td>
</tr>
<tr>
<td>Freya</td>
<td></td>
<td>27.63</td>
<td>13.65</td>
<td>0.46</td>
</tr>
<tr>
<td>Cameron</td>
<td></td>
<td>34.42</td>
<td>22.51</td>
<td>0.46</td>
</tr>
<tr>
<td>Average group effect (SE)</td>
<td></td>
<td>26.68* (6.42)</td>
<td>16.15 (7.04)</td>
<td>.46 (.48)</td>
</tr>
</tbody>
</table>

*Note. EB = empirical bayes; SE = standard error; *p <.05; non-significant between case variance estimates: Intercept = 96.39, Phase = 48.45, Phase*Time interaction = 0.00, autocorrelation (1) = .15

**Story Elicitation Skills.** Table 13 provides information about individual and group *Tell Me More* effects on caregivers’ story elicitation skills. There was a small degree of variability in caregiver baseline rates for story elicitation skills, ranging from 0.30–0.38, but participant baseline averages did not vary significantly from the group mean (0.34, *p* = .004). The overall immediate effect of the second module on story elicitation skills was positive and small, but non-significant, β = .22, *p* = .20. Individual caregiver treatment effects reflected the expected variability in observations across participants, ranging from .45 (Talia) to 0.08 (Freya). Finally, the overall degree of change in trend during the treatment phase was very small and non-significant, β = .003, *p* = .81. Only Freya reported a positive change from the group mean, although it was non-significant and near-zero, b = 0.02, *p* = .61.

Table 13. Individual and Group EB Estimates of *Tell Me More* Effect on Caregiver Story Elicitation Skills

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Caregiver(s)</th>
<th>Intercept</th>
<th>Phase</th>
<th>Phase*Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story elicitation skills</td>
<td>Talia</td>
<td>0.36</td>
<td>0.45</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Holly &amp; Alexander</td>
<td>0.30</td>
<td>0.10</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Freya</td>
<td>0.32</td>
<td>0.08</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Cameron</td>
<td>0.38</td>
<td>0.25</td>
<td>0.003</td>
</tr>
<tr>
<td>Average group effect (SE)</td>
<td></td>
<td>0.34* (0.04)</td>
<td>.22 (.14)</td>
<td>.003 (.01)</td>
</tr>
</tbody>
</table>

*Note. EB = empirical bayes; SE = standard error; *p <.05; Between case variance estimates non-significant: Intercept = .003, Phase = .04, Phase*Time interaction = 0.0002, autocorrelation (1) = -.058
**Story Strengthening Skills.** The last training module targeted the improvement of skills intended to support or strengthen children’s conversational language. Table 14 includes estimates of baseline levels and trends, as well as individual treatment effects. Like the previous two skills, the intercept parameter was a significant contributor to the specified model, \( \beta = 1.34, p = .001 \). Individual participant baselines did not vary significantly from the group mean, with \( p \)-values ranging from .13 to .81. The average treatment effect across participants was positive and small but non-significant, \( \beta = .29, p = .15 \). Individual treatment effects on story strengthening skill rates varied across participants, \( bs = .21–.37 \), although individual estimates of shift in level were non-significant for all participants. Furthermore, there was no significant change in slope from baseline to treatment within or across participants.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Caregiver(s)</th>
<th>Intercept</th>
<th>Phase</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story strengthening skills</td>
<td>Talia</td>
<td>1.04</td>
<td>0.27</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Holly &amp; Alexander</td>
<td>1.69</td>
<td>0.37</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Freya</td>
<td>1.38</td>
<td>0.29</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Cameron</td>
<td>1.24</td>
<td>0.21</td>
<td>0.02</td>
</tr>
<tr>
<td>Average group effect (SE)</td>
<td>1.34*(.19)</td>
<td>0.29 (18)</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* EB = empirical bayes; SE = standard error; \(*p < .05\); non-significant between case variance estimates: Intercept = 96.39, Phase = 48.45, Time = 0.00, autocorrelation (1) = .23

**Feasibility and Acceptability**

**Participation Logs.** Based on REDCap participation logs, we were able to document intervention completion as a measure of feasibility. Four out of five families completed the study. The four caregivers accessed and completed all three training modules as intended, suggesting the intervention to be feasible for them. Furthermore, for these four participants, there was no missing survey or questionnaire data. One participant voluntarily withdrew from the study after they reported difficulties due to changes in their work schedule that limited their
ability to record conversations with their child. Because they only completed Module 1, we could not examine intervention effects. Therefore, they were not included in the study. For this participant, the intervention was insufficiently feasible.

**End-of-Module Feedback Surveys.** All participants completed acceptability surveys at the end of each module. End-of-module survey responses are not included in a table because all caregivers answered “yes” to each item across all three end-of-module feedback surveys. In summary, caregivers reported that each module was easy to access and easy to complete in one sitting. They also reported that the examples and videos in each module were appropriate and adequate. Caregivers also responded “yes” to the item, “The module content was easy to relate to (relevant to me and my child).” However, two caregivers contacted the primary author to report problems navigating the first training module, to include a non-functional exit button on the final screen (it did not close out the window). Within the end-of-module feedback surveys, all caregivers shared about technological issues they encountered while engaged in the online training. Commonly reported issues were related to the training audio, such as “doubled voice” and low volume during video models. The same caregivers reportedly needed to refresh their browsers or force a hard restart when accessing the training modules using their smartphone. Freya, who consistently used her computer to watch the videos, reported no problems.

**End-of-Study Questionnaire.** After the study ended, each caregiver completed a final measure of intervention acceptability. Overall and ingredient-specific results are reported below.

**Tell Me More Overall.** As shown in Table 15, all caregivers were satisfied with the online delivery of *Tell Me More* ($M = 4.25$, range of scores: 3–5). In the section on the overall acceptability of *Tell Me More*, caregivers agreed that the intervention improved their knowledge about storytelling. Specifically, caregivers agreed (4) or strongly agreed (5) that the intervention
helped them to gain a better understanding of how children learn to tell stories. There was also
strong agreement across participants that the intervention taught them how to tell a good story
(oral narrative language), how to strengthen their child’s stories ($M = 4.5$, range: 4–5), and to a
slightly lesser degree, how to get their child to tell a good story ($M = 4.25$, range: 4–5). However,
when asked to consider if the knowledge and skills gained would improve their child’s language,
caregivers were less confident in the training’s efficacy. On average, caregivers rated the degree
to which their skills and knowledge would transfer to their child’s oral narrative quality as 3.5
(range: 3–4) and improve overall conversational skills ($M = 4.25$, range: 3–5).

**Training Modules.** With respect to the training modules specifically, the caregivers
found them easy to access ($M = 4.5$, range: 4–5) and strongly agreed that the training modules
were easy to complete in one sitting, despite reporting technological issues within the end-of-
module surveys. All but one caregiver strongly agreed (5) that the length of time to complete
each module was appropriate, with Cameron providing a neutral (3) response. However, there
was more variability in caregivers’ opinion about whether three modules were enough to learn
the skills. While Talia and Freya agreed that the training module dose was adequate, two families
(Holly and Alexander, Cameron) were neutral ($M = 3.75$, range: 3–5).

**Family Activities.** Three of the four families participating in the study completed all six
sets of family activities. One family (Holly and Alexander) stopped using the family activities
after set 1, reporting a lack of confidence that their child would cooperate with the story retelling
and story generation portions. Instead of using the optional family activities, they practiced the
taught skills in their “regular conversations.” No difficulties were reported by caregivers who
used the family activities. There was strong agreement that the family activities were easy to
access and completed quickly with their children. On average, the caregivers reported agreement
that the family activities helped them learn how to apply the storytelling, elicitation, and strengthening skills effectively ($M = 4$, range 3–5). Freya and Cameron provided unsolicited feedback about the family activities, reporting them to be visually appealing and enjoyable. Freya requested additional family activities after the study had ended.

**Discussion**

The purpose of this study was to examine the preliminary effects, feasibility, and acceptability of the *Tell Me More* intervention on caregivers’ ability to use specific strategies in conversation with their young children. We collected and transcribed caregiver-child language samples to measure intervention effects on three caregiver behaviors: oral narrative language, story elicitation skills, and story strengthening skills. In addition to measuring intervention effects using a multiple baseline across behaviors design, we documented intervention feasibility as it relates to completion of intervention components. Finally, caregivers provided feedback during and at the end of the study using surveys and questionnaires, from which we gleaned information about the acceptability of the *Tell Me More* intervention.

**Intervention Effects on Caregiver Behaviors**

Over the course of the study, all caregivers improved oral narrative language quality (linguistic complexity and structure), story elicitation, and story strengthening skills, although to varying degrees within and across participants. The effect size estimates for all outcomes were positive, although small. However, visual analyses of participant MBD graphs and non-significant HLM results do not suggest a causal relation between the *Tell Me More* intervention and changes in caregiver verbal behavior. We cannot confidently presume that the intervention was the primary reason for improvements in caregivers’ oral narrative language, story elicitation,
and story strengthening skills. Instead, the lack of significant improvements in any or all taught skills highlights the need to re-examine and modify intervention ingredients.

While narrative conversations with caregivers may be the primary context in which young children acquire knowledge of structure and content to generate their own independent personal narratives (Kelley, 2018), the baseline data results across participants suggest the need for caregivers to receive knowledge and training in skills necessary to act as instructional oral narrative language models. In other words, they did not already have these skills at baseline. Furthermore, caregivers varied in their quantity and quality of oral narrative language skills throughout the study. At least one caregiver was aware of this lack of knowledge; in a text message sent to the first author during week 4, one parent stated, “I’ve been a bad storyteller this whole time.” There is emerging evidence that such self-reflective behavior can predict continued improvement in new skills (Lo & Wong, 2022; Riva Crugnola et al., 2018). In addition, improvements in caregiver story elicitation were the least noticeable compared to oral narrative language and story strengthening skills. Although visual analysis of Talia’s second baseline graph (Figure 5) suggests the presence of a treatment effect on story elicitation skills, across-participant comparison of changes in level and trend weakened our confidence of an overall treatment effect.

Caregivers did not appear to enter the study with knowledge of how to elicit a conversation from their child; they used more close-ended (yes/no) questions similar to an interrogation. As a result, children were less responsive to their parents’ attempts. Based on behaviorist motivation theory (Michael, 1993), we speculate that caregiver story elicitation skills were reinforced by children’s responsiveness and a lack of reinforcement led to diminished rates of caregiver story elicitation skills over time. This behavioral sequence may reveal why the
teaching of language skills is difficult for parents and primary caregivers, as opposed to professionals who understand response cycles and reinforcement (Sundberg, 1993). Additional tips and strategies may need to be added to the second module to prevent parents from giving up on further attempts when initial conversation starting attempts are ineffective.

Research on maternal support during joint storytelling with young children suggests that mothers’ interaction style promotes reciprocity and overall length of conversation (Kelley, 2018). However, there is little evidence available to determine which behaviors at the utterance level are most likely to increase engagement and reinforce children’s narrative language. Therefore, assuming that caregivers would be able to elicit language from their child, we taught six strategies meant to strengthen the child’s contributions. However, we encountered two problems with this assumption. First, we assumed this sample of caregivers would have some individual skills within their repertoire based on the popularity of responsive parenting in online parenting groups. Second, in the ordering of the modules, we made an assumption a priori that the second module would result in increased opportunities to shape up children’s oral language using the RIVER skills. Our hypotheses were incorrect in both cases. Caregiver demonstration of RIVER skills in baseline was variable and rates of repeats and elaborations were lower than validating and relating statements. It is also possible that because parents did not improve story elicitation skills, there was no increased opportunity to practice the RIVER strategies in conversation with their child. Regardless, there is a need for further exploration as to why the training modules did not demonstrate a causal relationship with the targeted caregiver behaviors.

**Authenticity vs. Control.** One possible reason for non-significant results is the general nature of studying parent-child conversation in naturalistic settings. Caregiver-child discourse is a beautiful but complex phenomenon; it is extremely informative and grounded in a family’s
culture. However, in the context of research, it is unpredictable, highly variable, and easily influenced by the mood and subtle environmental factors. While we acted on research suggesting naturalistic observations to be the best and most authentic measures of parental linguistic input (Adamson et al., 2021), we did not have control of the contingencies governing the relationship between caregivers’ conversational topics nor length of the conversation. We did not place restrictions on the time, setting, or topic of conversation, so the duration and magnitude of conversational turns per sample was variable.

Some caregivers had difficulty generating conversational topics, resulting in less child engagement and shorter interactions. Other parents took measures to increase the likelihood of child engagement. For example, Talia (Family 1) requested permission to hide her smartphone when recording video conversations, because her child was distracted by it. Cameron also had difficulty recording video at times and shared audio recordings of conversations with her son instead. While we could not report information about the observation setting, audio was usable in the transcription process. In addition, there were several times that other persons in the environment interrupted the ongoing conversation to ask an unrelated question or receive help for an unrelated problem. We expected such interruptions to occur as a consequence of naturalistic observations. We did not, however, predict the variety of environments in which caregivers would videotape conversations. In this study, observations occurred during car rides home from daycare, on walks through the neighborhood, during bath time, while sitting on the living room floor or at the dining table, jumping on a backyard trampoline, and as a part of the child’s bedtime routine. This finding was extremely encouraging and supports the utility and versatility of storytelling. However, as a result, we suspect that had some influence on both duration and conversational quality or productiveness.
**Intervention Dosage.** Intervention dosage may not have been enough to result in the transfer of knowledge from training module into caregiver practice in natural conversational opportunities with their child. By holding the duration of each training module constant (between 8–10 minutes) and using only three modules in total, we limited the magnitude and specificity of knowledge and model demonstrations presented to caregivers. This was done to ensure each family had approximately equal intervention dosage (equal opportunity/access to a training module within a two-week period). However, it is possible that caregivers may need more time and examples (as requested by one participant in the end-of-study survey). Unfortunately, we did not gather more specific information about how many times the caregivers watched each module to completion, the average duration of time spent in each module, or patterns in days and times of access. This information may be relevant in the study of module potency. Because digital implementation of the intervention allows for easy collection of such data when using background data collection applications like Firebase, the next iteration of studies should include usage and engagement metrics. Such information will allow us to further explore module-specific engagement, module ingredients, and caregivers’ processes of behavior change.

**Sampling Bias.** Sampling bias may be another reason there was a small but non-significant treatment effect. The extent to which parents’ educational attainment affected their baseline scores and responsiveness to intervention is unknown without a comparison group of parents and caregivers without advanced degrees. It is likely that future studies with different populations may not replicate these outcomes. For example, Alexander used words like *fermented* and *secure* during baseline conversations. Although caregiver participants were not representative of the general population or those who would most likely need a dyadic oral language intervention, it appeared that even these caregivers needed instruction. For meaningful
behavior change to occur, a more potent intervention will be needed. The characteristics of the learners (i.e., caregivers in this case) will likely intersect with the rate of learning in the *Tell Me More* intervention. Unfortunately, the current study does not yield actionable guidance about this, only that researchers should be cautious and careful about sampling and examine learner characteristics vis a vis behavior change.

**Tell Me More Feasibility and Acceptability**

When interventions are delivered in-person, outside of the home, and require coordination of schedules between parent and interventionist, their reach and impact can be extremely limited (Breitenstein et al., 2016; Greenwood et al., 2020). For this study, online delivery of the *Tell Me More* intervention was both feasible and acceptable for the small sample of parents living in different geographic areas in the United States. Feedback from caregivers indicated that the asynchronous *Tell Me More* training made it easier for them to complete the program because they could do so at any time. Although their acceptability was generally high, it is worth noting that caregivers had less confidence that their new knowledge and skills would translate into improved oral language of their children. An average score of 3.5 on this acceptability item aligns with the mediocre results of the intervention’s effect on caregiver behaviors. Caregivers were unsure about the potency of the intervention for helping their children’s language, despite finding the intervention acceptable.

The notion that caregivers found *Tell Me More* to be acceptable is an important finding. All participants had full-time employment in addition to their caregiver responsibilities. Records of module completion and end-of-module surveys also indicated high degrees of parent participation. Five out of six (counting Holly and Alexander separately and the caregiver who withdrew) enrolled participants completed all three modules. A completion rate of 83% is high,
given that other studies of digitally delivered parenting programs report an average rate of completion of 78.3% (range: 41.7–99.2%; Breitenstein et al., 2014). Parents were enthusiastic about participating and several shared funny stories about attempts to record conversations on video. One parent, Freya, requested more family activities after the study had ended because she and her son enjoyed doing them together. While one parent did not have any trouble accessing and completing the training using the browser on their personal laptop, four individual reports of modules “freezing” and caregivers having to force close the application reveal areas in need of improvement in later design iterations (e.g., “Running on my phone would cause audio issues…nothing a restart couldn't fix.”). In general, the positive feasibility and acceptability results suggest that the Tell Me More intervention is worthy of continued investigation and iterative development. Had the caregivers not liked the intervention, it would be difficult to justify further revisions of Tell Me More to improve its effect on caregiver and child behaviors. As it stands now, this study yielded extensive information about where the intervention is weak and what elements show promise that can eventually lead to an efficacious, feasible, and acceptable intervention.

Limitations

There are several methodological limitations to this study related directly related to research studying parent-child language intervention effects. First, caution must be taken in assumptions derived from the supplemental statistical analyses. For example, the level-2 model did not find variability in baseline levels across participants for some behaviors, reporting covariance parameter estimates of near-zero. Although they are a part of the solution and reported in this paper, it may be that estimates of baseline level variance were not obtained because of the small sample size. Future research may also consider exploring the bounce, or
magnitude of individual deviations from their own unique trajectories, occurring from session to session. Through investigation of within-participant variation, future designs will need to consider use of a component analysis to explore the potential influence of environmental variables on participants’ response to intervention. Variables like observation setting, context (e.g., the ongoing activity in which the conversation is embedded), and presence of others in the environment may be correlated with language outcome effects. In other studies, measures of household noise and activity have also been used as a moderator when estimating the effects of interventions delivered in a child’s home environment (Wilhoit et al., 2021).

Although we intended to recruit a more diverse sample of caregivers by using social media platforms like Instagram, Twitter, and Facebook, we did not anticipate the degree to which our own networks may have influence on selection bias. For example, tertiary sharing of the study recruitment flyer on Facebook landed it in a Facebook group called “PhD Mamas.” While the first author was not initially aware of this, it became very apparent during the informed consent process leading the first author to follow up with each caregiver to ask where they learned about the study. Three participants were members of that Facebook group. It is likely that self-selection bias played a role in the degree to which parents were motivated to participate and not solely the result of an engaging and accessible online intervention.

**Contributions and Future Research**

One strength of this study is that from inception, we designed *Tell Me More* to be an online asynchronous intervention. While some may see online interventions as inequitable for low SES or rural populations, two of four families completing the study lived in areas designated as rural. Neither had broadband internet and both were able to access and complete the training modules and family activities, one reported no connectivity issues. Furthermore, we designed the
intervention so that it can be accessed on a smartphone, tablet, laptop, or desktop computer. Four out of five families accessed the training using their phones. This finding is consistent with research conducted in rural U.S. communities (Collins et al., 2019) and rural refugee camps in Europe (UNHCR Innovation Service, 2016). Online delivery of the training was an essential and defining feature of *Tell Me More* and will prime the intervention for scaling after changes are made to increase its efficacy.

Considering the lack of large and significant treatment effects, online training can be made more efficacious when automated tailored messaging and coaching is provided in a timely manner (Pellachia et al., 2020). Therefore, this is a logical addition to *Tell Me More*. Previous research has shown that use of text messaging increases family participation and engagement (Justice et al., 2020; Muench & Baumel, 2017). Since parents in this study preferred to interact with the first author using text messaging, it follows that use of text messaging in the provision of coaching may further improve the intervention’s efficacy.

In their systematic review of technology in parenting interventions, Correlajo and Domenech Rodriguez (2018) found that most programs were validated with White families that would require significant content changes to increase scalability to other demographic populations. Therefore, we intentionally included a pedagogical agent with brown skin and hair, while not specifying the agent’s race or ethnicity. In addition, we made sure to program for diversity in exemplars shown in the modules. For example, photographs depicted caregivers of various races, ethnicities, genders, ages (e.g., grandparents), religions (e.g., Muslim). When a training discussed the role of family, we included images of two-father households, as well as intergenerational models. Although our study included families living in different geographic regions in the U.S., the sample was otherwise homogenous. Therefore, we do not yet know if the
consideration of diverse cultural representation had an impact on the acceptability of the intervention. We have an ethical and fiscal responsibility to attend to the culture, language, and beliefs of our end users and the scientific evidence underpinning our intervention design. Future research with a heterogenous sample of caregivers is needed to answer this question.

This study combined research in adult learning theory, parent implemented programs, and technology-based interventions with child development. The results of this study indicate that caregivers are highly motivated to receive training in methods that support their child’s language development and caregivers find online asynchronous training feasible and acceptable. Beginning with the end in mind, online asynchronous caregiver interventions eliminate known access barriers like transportation and scheduling, while also reinforcing caregivers’ motivation and interest in learning more about their child’s language development.

References


163


### Table 11

*Means (M), Standard Deviations (SD), and Range by Variable and Participant*

<table>
<thead>
<tr>
<th>Caregiver(s)</th>
<th>Duration</th>
<th></th>
<th></th>
<th>Oral Narrative</th>
<th>Story Elicitation</th>
<th>Story Strengthening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B (M, SD)</td>
<td>Tx</td>
<td>B (M, SD)</td>
<td>Tx</td>
<td>M (SD)</td>
<td>Tx</td>
</tr>
<tr>
<td>Talia</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92 (43)</td>
<td>144 (72)</td>
<td>19.4 (7.8)</td>
<td>17.5 (11.2)</td>
<td>11 (1.0)</td>
<td>28 (12.2)</td>
</tr>
<tr>
<td></td>
<td>61–168</td>
<td>62–288</td>
<td>12–30</td>
<td>6–48</td>
<td>10–12</td>
<td>7–43</td>
</tr>
<tr>
<td>Holly &amp; Alexander</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>226 (98)</td>
<td>261 (83)</td>
<td>19.5 (6.0)</td>
<td>21.6 (7.7)</td>
<td>31.3 (16.7)</td>
<td>47.5 (22.8)</td>
</tr>
<tr>
<td></td>
<td>92–350</td>
<td>60–462</td>
<td>11–26</td>
<td>4–40</td>
<td>18–63</td>
<td>11–102</td>
</tr>
<tr>
<td>Freya</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>193 (75)</td>
<td>213 (50)</td>
<td>27 (10.2)</td>
<td>26.4 (5.2)</td>
<td>31.2 (13.2)</td>
<td>43.2 (13.4)</td>
</tr>
<tr>
<td>Cameron</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>327 (230)</td>
<td>527 (316)</td>
<td>47.7 (38.0)</td>
<td>59 (39.8)</td>
<td>31.2 (23.5)</td>
<td>66.1 (22.3)</td>
</tr>
<tr>
<td>Across participants</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>215 (151)</td>
<td>293 (221)</td>
<td>28.8 (22.7)</td>
<td>31.7 (26.6)</td>
<td>26.8 (17.4)</td>
<td>47.1 (22.5)</td>
</tr>
</tbody>
</table>

*Note.* B = baseline phase, Tx = treatment phase; Duration is rounded to the nearest second (i.e., .5 rounded to 1.0); Turns and Oral Narrative outcomes are reported as a count; Story Elicitation and Story Strengthening are reported as a rate per conversational turn.
Table 15

Caregiver Responses to End-of-Study Questionnaire: Overall, Training Modules, and Family Activities

<table>
<thead>
<tr>
<th>Survey item</th>
<th>Mean Rating</th>
<th>Caregiver feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tell Me More overall</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three modules was enough to learn these skills.</td>
<td>3.75</td>
<td>• No problems</td>
</tr>
<tr>
<td>I am satisfied with the online delivery of the training and family activities.</td>
<td>4.25</td>
<td>• [Module 3] was my favorite of all!</td>
</tr>
<tr>
<td>Tell Me More helped me gain a better understanding of how children learn to tell stories.</td>
<td>4</td>
<td>• I think my biggest obstacle is filming when my child is talking.</td>
</tr>
<tr>
<td>This understanding has allowed me to teach my child more effectively.</td>
<td>4</td>
<td>• Otherwise, in natural settings, I am finding these tactics helpful!</td>
</tr>
<tr>
<td>The skills and knowledge I learned will help my child tell better stories.</td>
<td>3.5</td>
<td>• Help to explain how to begin a conversation or what topics to approach. I struggled coming up with topics at times.</td>
</tr>
<tr>
<td>The skills and knowledge I learned will help my child improve their conversation skills.</td>
<td>4.25</td>
<td>• I'm not clear as to whether this is about using made up stories to teach things or if it is about taking about real events. It showed a lot of book reading but also talked about resolving emotions about real problems.</td>
</tr>
<tr>
<td>Tell Me More taught me how to tell a good story.</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Tell Me More taught me how to get my child to tell their own story.</td>
<td>4.25</td>
<td></td>
</tr>
<tr>
<td>Tell Me More taught me how to strengthen my child's stories.</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td><strong>Training modules</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The modules were easy to access.</td>
<td>4.5</td>
<td>• They were very well done. I actually don't have any improvements to suggest.</td>
</tr>
<tr>
<td>The length of time to complete each module is appropriate.</td>
<td>4.5</td>
<td>• The audio was sometimes difficult to hear. Provide more clear examples (voices in example I hard to hear)</td>
</tr>
<tr>
<td>The modules were easy to complete in one sitting.</td>
<td>5</td>
<td>• Adding a few more examples.</td>
</tr>
<tr>
<td>The modules had good examples and videos.</td>
<td>5</td>
<td>• Running on my phone would cause audio issues...nothing a restart couldn't fix.</td>
</tr>
<tr>
<td>The module content was easy to relate to (relevant to me and my child).</td>
<td>5</td>
<td>• I had to reload a few times when the example videos were about to be shown.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• I watched on my phone and a call came in. Chrome stopped the video but was hard to restart it.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Weird doubled voice when I clicked on the benefits of storytelling buttons. Also I had to reset it once because it froze.</td>
</tr>
<tr>
<td><strong>Family activities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The family activities were easy to access (log-in, start).</td>
<td>5</td>
<td>• It might be good to put a little note to parents to say that they can encourage/prompt the child more than the actual questions on the activities. My son struggled with some of them but with scaffolding (which I began to offer towards the end) did incredibly well.</td>
</tr>
<tr>
<td>The family activities were quick to complete with my child.</td>
<td>5</td>
<td>• I didn't really use them, so I'm not sure. I didn't think my child would cooperate with retelling the first story so I just tried to use the skills in my regular conversations. Maybe make the first stories silly?</td>
</tr>
<tr>
<td>The family activities were relevant to me and my child.</td>
<td>4.67</td>
<td></td>
</tr>
<tr>
<td>The family activities helped me learn how to apply the storytelling, elicitation, and strengthening skills effectively.</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Note. *Holly and Alexander did not elect to use the family activities; Likert items measured on a scale of 1 (strongly disagree) to 5 (strongly agree)
Figure 5. Multiple Baseline Design Across Behaviors with Replications Across Participants
CHAPTER FIVE:
CONCLUSION

This dissertation included three different intervention studies that tell a story about the utility and effectiveness of oral narrative language interventions implemented in natural environments by teachers, speech-language pathologists (SLPs), and parents. A familiar practice in many households, storytelling is a powerful teaching tool (Curenton, 2006; Dunst et al., 2012; Spencer & Petersen, 2020). Oral narrative language development is also related to children’s development of reading (Dickinson & McCabe, 2001; Snow et al., 2007), writing (Kirby et al., 2021; Spencer & Petersen, 2018), and social skills (Bliss & McCabe, 2012; Brinton & Fujiki, 2019; Hart et al., 2004). The results of these studies may have implications for the use of oral language interventions to improve early childhood language nutrition (Zauche et al., 2017) and support students’ development of reading, writing, and other academic outcomes.

While there is growing evidence of oral narrative intervention effects on language and literacy outcomes for children in preschool through early elementary grades, most findings are the result of methods employing researchers as interventionists and data collectors (Greenwood et al., 2020). There is insufficient evidence of intervention effectiveness under “real world” conditions, when implemented by persons natural to the environment and outcomes measured in naturalistic contexts (Anderson et al., 2021; Coburn et al., 2013). Thus, these three studies
prioritized the involvement of intended end users when examining the effects of oral narrative interventions delivered to preschoolers, kindergarteners, and first graders.

The first dissertation study examined the effects of a brief teacher-implemented oral narrative intervention on six kindergarteners’ text generation skills. A multiple baseline across groups design and action-research framework was used to study the extent to which the kindergarteners’ narrative writing after only six small group lessons. The study demonstrated that teacher-implemented oral narrative intervention resulted in immediate and lasting effects on student writing when the intervention was delivered by the classroom teacher during their regular literacy center rotations. The results are early evidence of a functional relation between oral narrative intervention effects on writing outcomes (Spencer & Petersen, 2018; Petersen et al., 2022).

The second dissertation study resulted from a research-practice partnership (RPP; Alonzo et al., 2022; Coburn et al., 2013; Coburn & Penuel, 2016) that explored the effects of an oral language program implemented within the context of a Multitiered System of Language Supports (MTSLS) by first grade teachers and school-based SLPs. Specifically, the study investigated the extent to which a new multitiered oral narrative intervention, *Story Champs Curriculum*, improved the language and literacy skills of first graders identified at risk for or having language disabilities. Using a randomized waitlist-controlled trial design, we measured students’ narrative and expository retell, vocabulary inferencing, narrative writing, and listening and passage comprehension at pretest, posttest, and follow-up. Analysis of the results found large and statistically significant improvements in proximal measures of narrative retell and vocabulary at posttest. Effects on intermediate measures of narrative writing were large and significant at posttest, extending support for the findings of the first study. Positive treatment effects were also
found on students’ expository oral retell skills at follow-up. This experiment is one of two (Petersen et al., 2022) that show that oral narrative intervention delivered in an MTSLS framework is both feasible and effective for improving oral language and writing skills of first graders’ at risk for or having language disabilities.

There is very little literature on the use of parents and caregivers as early childhood oral narrative language interventionists (Pico et al., 2021). Thus, the third dissertation study sought to address this knowledge gap by designing and testing the preliminary effectiveness, feasibility, and acceptability of an online asynchronous parent training intended to improve preschoolers’ oral narrative language. The caregivers improved in all measured outcomes, although the effects were small and non-significant. One possible explanation for the lack of significant intervention effect is the nature of the research conducted by parents in natural contexts, as well as collecting parent-child conversation samples to measure adult language outcomes. However, results of the study indicated that *Tell Me More* intervention was feasible and acceptable and will inform changes in the iterative design of the oral narrative intervention.

Results of the research presented in this dissertation demonstrate that oral narrative language intervention is effective for improving young students’ language and literacy skills when implemented by teachers and SLPs. Results of the first study revealed how use of an oral narrative language intervention alone can improve developing writers’ text generation skills, in absence of any explicit writing instruction. This knowledge adds to the understanding of how narratives serve as a bridge between oral and written language (Kirby et al., 2021; Spencer & Petersen, 2018). Extending these results, the second study also showed that oral narrative intervention delivered in a MTSLS model by school teams had a large, positive, and significant effect on the narrative writing skills of first graders at risk for or having language disabilities.
Further, the second study also contributed to literature demonstrating the effectiveness of oral narrative intervention on students’ vocabulary inferencing and narrative and expository retell. The latter is one of the first signs of narrative intervention resulting in the improvement of expository language.

The last dissertation study lends support for the design and use of accessible and scalable parenting programs to support early childhood language nutrition. While neither large nor significant, the effects of *Tell Me More* on improving caregivers’ oral language skills in conversation with their children will result in the improvement of its primary ingredients: training modules and family activities. The iterative design process holds the promise of creating a feasible, acceptable, and effective tool to empower parents and caregivers to become early childhood oral narrative language facilitators. Therefore, our future studies of *Tell Me More* will include a more diverse sample of caregivers and examine the effects of caregiver language improvements on child oral language outcomes. Such information will lead to improved operationalization of primary intervention ingredients and identify variables that can be adapted to improve its reach to traditionally marginalized communities.

The current findings add to our understanding of oral narrative intervention effects on children’s language development and academic achievement. Studies 1 and 2 examined the extent to which oral narrative intervention is effective for improving K-1 students’ language and literacy skills, particularly when programs are implemented by natural end users in various contexts. The results of the first two studies showed that interventions implemented by teachers and, in the case of study 2, SLPs, are indeed effective change agents, despite differences in fidelity of implementation. There is a chance that we would have found a larger effect if we had more control over the contingencies governing implementation and measurement outcomes.
However, the effects on student language and literacy skills, at times large and statistically significant, are more exciting and encouraging because school teams and not researchers implemented the intervention and collected the outcome measures under real world conditions.

Study 3 contributed knowledge to an emerging field of online asynchronous parenting programs focused on improving language nutrition in early childhood (Bellon-Harn et al., 2020; Greenwood et al., 2020). The study methods and measures also provide additional exemplars for the study of parent-child dialogue in natural settings and contexts. By reporting the feasibility and acceptability of the online asynchronous format of Tell Me More, the results may also be useful to those studying methods for designing scalable interventions. All caregivers were satisfied with the online training program and found the asynchronous nature of delivery a strength since they could access the modules and activities anywhere at any time. In general, the study adds to a growing literature base supporting the design and use of online parent training.

Results of these studies also have possible implications for bridging the gap between research and practice. Because the interventions were implemented by parents across a variety of settings or by school professionals in natural instructional contexts like MTSLS, the data offer a more complete account of implementation facilitators and barriers. Information gleaned from having teachers and SLPs as interventionists and school teams as data collectors led to results that may be more interpretable and meaningful to practitioners. With knowledge that the intervention was implemented and studied in real world conditions, professionals can draw conclusions about how oral narrative interventions may be effective for improving their own students’ listening comprehension, vocabulary, reading, writing, and expository language. RPPs are useful in the study of intervention effectiveness and may yield findings with greater relevance to the intended audiences.
We hope to see further research building off the results of these three studies. Oral narrative interventions can be effective for a variety of children, but evidence of oral narrative intervention effects when delivered within an MTSLS framework is still emerging. Large scale randomized controlled trials will be necessary to provide additional evidence of effectiveness, utility, and scalability. To improve the speed with which effective interventions transfer into the hands of those who teach, we encourage our colleagues to involve teachers, SLPs, parents, and caregivers in the design of educational research. In addition, universities, school districts, and other community leaders will need to improve funding to support this meaningful and impactful approach to research.

Finally, while the findings presented in the third dissertation study advance research in online parenting programs, changes are needed to the intervention to examine its effectiveness on parent language skills prior to examining their impact on child-level outcomes. Modifications were expected because this study was the first to examine the Tell Me More intervention and use the novel language sampling procedures. As intervention science demands iteration and design-based methods, the current study fulfilled its purpose in that we gleaned the critical information to guide revisions and future directions. Nonetheless, there is much to be done in the area of parent-implemented oral language interventions. While we encourage other researchers to join this line of work, our findings suggest more attention needs to be given to the procedures for data collection and measurement of language. Parents and children make up a complex system of interaction and there was little empirical guidance available for how to measure dyadic productions within natural routines. We contend that when ample literature on the use of naturalistic language sampling procedures is available, researchers will more confidently
measure this critical and complex phenomenon, which will ultimately hasten the development of high impact interventions.

Overall, the three studies presented in this dissertation provide further evidence of the utility of storytelling and support for its use as a mechanism for improving children’s oral language skills. Further, the cultural and linguistic flexibility of narrative language, as well as a lack of need for physical materials, may enhance the utility and scalability of oral narrative interventions. Only when effective oral narrative language interventions are both feasible and acceptable will they reach the children who will benefit from them.

References


APPENDICES
Appendix A: Narrative Language Measures (NLM) Flow Chart (Language Dynamics Group, 2020)
Appendix B: CUBED Narrative Language Measures (NLM) Listening (Language Dynamics Group, 2019)

NLM Scoring

The NLM subtest is designed for real-time scoring. This means that while the student is retelling a story, the examiner can score the story simultaneously. Each NLM protocol sheet has a model and an accompanying score sheet that reflects the content of the selected story. For students who speak very softly or quickly, the NLM may need to be scored from an audio recording or from a transcript derived from an audio recording.

NLM Listening protocol sample

Examiner script

Story for examiner to model. Story Grammar icons are included to help real-time scoring.

LISTENING RETELL:
A student retells a story, examiner scores retell in real-time. There are benchmark, moderate risk, and high risk cut points for this total score.

COMPREHENSION QUESTIONS:
Story Questions: A supplemental section designed as a secondary measure of comprehension, administered usually when a student does not produce a complete episode (problem, attempt, consequence) or if a student does not meet benchmark on the retell score.

Vocabulary Questions: A supplemental section designed to measure a student's knowledge of low frequency words (also called tier-2 words), and/or to measure a student's ability to infer the meaning of words from context.

PERSONAL GENERATION:
A supplemental section designed to measure a student's ability to generate their own personal narrative based off the theme of the model story. To score, use the Narrative Language Measures: (NLM) Flow Chart in Appendix B.
Appendix C: Expository Language Measures (ELM) Flow Chart (Language Dynamics Group, 2020)
Appendix D: *Tell Me More* Study Activities and Timeline
### Appendix E: Scoring Rubric—Caregiver Skills to Strengthen Children’s Stories

<table>
<thead>
<tr>
<th>Code</th>
<th>Variable Name</th>
<th>Operational Definition (observable &amp; measurable)</th>
<th>Variable Type</th>
<th>Scoring Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Repeat</td>
<td>A repeat is when the caregiver echoes the child’s utterance with point-to-point/word-for-word correspondence to reinforce their grammatically correct utterance and/or use of more complex descriptive language. <em>Example: Child says, “We went to the store,” the caregiver responds, “Oh, you went to the store.”</em></td>
<td>Continuous ((0 – \infty))</td>
<td>Count per video</td>
</tr>
<tr>
<td>I</td>
<td>Improve</td>
<td>An improvement is the caregiver’s repeat of the child’s grammatically incorrect or incomplete utterance for the purpose of correcting grammatical errors, while maintaining the original meaning. It does not require the child to echo or correct their original utterance. <em>Example: Child says, “He go dare,” the caregiver waits for the child to finish and replies, “He went over there.”</em></td>
<td>Continuous ((0 – \infty))</td>
<td>Count per video</td>
</tr>
<tr>
<td>V</td>
<td>Validate</td>
<td>A caregiver’s response to the child’s utterance that explicitly acknowledges and affirms they are listening to the child. A validation can be a verbal statement, gesture, and/or facial expression. <em>Example: nodding head or saying “mm hmm” with surprised facial expression</em></td>
<td>Continuous ((0 – \infty))</td>
<td>Count per video</td>
</tr>
<tr>
<td>E</td>
<td>Elaborate</td>
<td>Any statement made by the caregiver that extends the child’s storyline by elaborating upon their utterance with new information, advanced vocabulary, and complex language. Elaborations serve as establishing operations to motivate the child or prompt the child to extend their story. Cannot be a question. <em>Example: Child says, “He go.” Caregiver replies, “Yes, he did move away. He lives with his papa in Florida now.”</em></td>
<td>Continuous ((0 – \infty))</td>
<td>Count per video</td>
</tr>
<tr>
<td>R</td>
<td>Relate</td>
<td>Statements that connect or relate content from the child’s story to their own experiences, events, persons—actual or fictitious (for the purpose of providing a relational statement). <em>Examples: “Did you know that when I was a little kid, I was also afraid of dogs?” or “Yes, I am scared of snakes, too.”</em></td>
<td>Continuous ((0 – \infty))</td>
<td>Count per video</td>
</tr>
</tbody>
</table>

| **RIVER** | Ability to Strengthen Children’s Stories | Composite; Sum of all sub-skill totals above | Continuous \((0 – \infty)\) | Total count/video |
Appendix F: Caregiver Language Sample Scoring Record

**TMM1A LANGUAGE SAMPLE SCORING**

<table>
<thead>
<tr>
<th>Adult Language Sample</th>
<th>Participant ID: ____</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcript File Name: ____</td>
<td>Video Duration: ____</td>
</tr>
<tr>
<td>Transcriber’s Initials: ____</td>
<td>Scorer’s Initials: ____</td>
</tr>
<tr>
<td>Sample Date: ____</td>
<td>Scoring Date: ____</td>
</tr>
</tbody>
</table>

**Language Complexity**

1. Count of relative pronouns used: ____
2. Count of verb/noun modifiers used: ____
3. Social-emotional vocabulary words (and number of times used): ____
4. Count of advanced vocabulary used (no social-emotional words): ____
5. Count of temporal ties used: ____
6. Count of causal ties used: ____
7. Dialogue: ____

**Narrative Structure**

<table>
<thead>
<tr>
<th>TOTAL SCORE: ____</th>
</tr>
</thead>
</table>
| 1. Character: ____
2. Setting: ____
3. Problem: ____
4. Plan/Attempt: ____
5. Consequence: ____
6. Emotion: ____
7. Ending: ____
8. Episode Complexity: ____

**TOSS: Story Elicitation Skills**

<table>
<thead>
<tr>
<th>TOTAL SCORE: ____</th>
</tr>
</thead>
</table>
| 1. Count of all “Tell Me About…” phrases used: ____; Can you tell me: ____; Other “tell me”: ____
| 2. Count of all open-ended questions asked: ____
| 3. Count of all personal story shares: ____

**RIVER: Story Strengthening Skills**

<table>
<thead>
<tr>
<th>TOTAL SCORE: ____</th>
</tr>
</thead>
</table>
| 1. Count of all repeats of child language: ____
2. Count of all improvements: ____
3. Count of all validations: ____
4. Count of all elaborations: ____
5. Count of all relating statements: ____

**Caregiver’s Conversational Turns:** ____
**Appendix G: Scoring Rubric—Caregiver Story Elicitation Skills**

<table>
<thead>
<tr>
<th>Code</th>
<th>Variable Name</th>
<th>Operational Definition (observable &amp; measurable)</th>
<th>Variable Type</th>
<th>Scoring Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>Tell me about…</td>
<td>The caregiver’s use of the phrase “tell be about…” as a directive to prompt the child to tell a story; Child’s can respond with a story or refusal (“no”)</td>
<td>Continuous</td>
<td>Count per transcribed sample</td>
</tr>
<tr>
<td></td>
<td>phrase</td>
<td>Non-example: “Can you tell me about?” (yes/no question)</td>
<td>(0 – ∞)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any question asked by the caregiver that results in the child’s response contributing more information to the construction of the story</td>
<td>Continuous</td>
<td>Count per transcribed sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Example: “What happened next?” or “How did you feel when…?”. Non-example: “Did you go to the doctor when you were sick?”</td>
<td>(0 – ∞)</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>Open-ended question</td>
<td>Any instance of the caregiver’s sharing of a story (personal or fictional generation) as an establishing operation for the child’s narrative thereafter.</td>
<td>Continuous</td>
<td>Count per transcribed sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Example: “I will tell you about a time when I was little, and my brother broke my toy.”</td>
<td>(0 – ∞)</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>Personal story share</td>
<td></td>
<td>Continuous</td>
<td>Count per transcribed sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0 – ∞)</td>
<td></td>
</tr>
<tr>
<td>TOSS</td>
<td>Ability to elicit</td>
<td>Composite; Sum of all sub-skill totals above</td>
<td>Continuous</td>
<td>Total count per transcribed sample</td>
</tr>
<tr>
<td></td>
<td>children’s stories</td>
<td></td>
<td>(0 – ∞)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix H: End-of-Study Questionnaire

Instructions: Please share your opinions and experiences with us. Circle one number for each item to indicate how much you agree (5) or disagree (1) with the statement.

1. The *Tell Me More* intervention helped me gain a better understanding of how children learn to tell stories.
2. This understanding has allowed me to teach my child more effectively.
3. The skills and knowledge taught I learned will help my child tell better stories.
4. The skills and knowledge I learned will help my child improve their conversation skills.
5. I found the following skills taught by *Tell Me More* to be very useful:
   (a) How to tell a good story
   (b) How to get my child to tell their own story
   (c) How to strengthen my child’s stories
6. I think the number of training modules and length of time to complete them is appropriate.
7. The family activities helped me learn how to apply the storytelling, elicitation, and strengthening skills effectively.
8. I am satisfied with the online delivery of the training and family activities.
9. The online training modules:
   (a) were easy to access (log-in, start)
   (b) complete in one sitting
   (c) were relevant to me and my child
10. The family activities (Boom Cards):
    (a) were easy to access (log-in, start)
    (b) quick to complete with my child
    (c) were relevant to me and my child

For items 11–12, please provide as much detail as possible to help us improve *Tell Me More*.

11. The training modules can be improved by: _________
12. The family activities can be improved by: _________
Appendix I: Individual Participant Multiple Baseline Design Across Behaviors Graphs

Figure 11. Treatment Effects for Family 1 (Talia)
Figure 12. Treatment Effects for Family 2 (Holly and Alexander)
Figure 13. Treatment Effects for Family 3 (Freya)
Figure I4. Treatment Effects for Family 4 (Cameron)
Appendix J: IRB Approval Letter

August 19, 2021

Megan Kirby
1400 E. Powhatan Ave
Tampa, FL 33604

Dear Megan Kirby:

On 8/18/2021, the IRB reviewed and approved the following protocol:

<table>
<thead>
<tr>
<th>Application Type</th>
<th>Initial Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRB ID</td>
<td>STUDY002996</td>
</tr>
</tbody>
</table>

Review Type: Expedited

Title: Development and Feasibility of an Online Self-paced Oral Narrative Language Training for Caregivers of Children Ages 3- to 5-Years-Old

Approved Protocol and Consent(s)/Assent(s):
- TMM Study Protocol
- HRP-502b(3) SB Combined Consent and Parental Permission

Approved study documents can be found under the ‘Documents’ tab in the main study workspace. Use the stamped consent found under the ‘Last Finalized’ column under the ‘Documents’ tab.

This research involving children as participants was approved under 45 CFR 46.404: Research not involving greater than minimal risk to children is presented.

Requirements for Assent and/or Permission by Parents or Guardians: 45 CFR 46.408 Permission of one parent is sufficient.

Assent is waived because it is not appropriate due to the age, maturity, and/or psychological state of the child.
Within 30 days of the anniversary date of study approval, confirm your research is ongoing by clicking Confirm Ongoing Research in BullsIRB, or if your research is complete, submit a study closure request in BullsIRB by clicking Create Modification/CR.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

Jennifer Walker
IRB Research Compliance Administrator
Appendix K: IRB Consent and Permission to Participate

Consent to Participate in Research and Parental Permission for my Child to Participate in Research

Information for parents to consider before agreeing to participate and allowing your child to participate in this research study

Title: Development and Feasibility of an Online Self-paced Oral Narrative Language Training for Caregivers of Children Ages 3- to 5-Years-Old
Study # 002996

Overview: The following information is being presented to help you and your child decide whether you would like to be a part of a research study. The sections in this Overview provide the basic information about the study. More detailed information may be provided in the remainder of the document.

Study Staff: This study is being led by Megan Kirby who is a doctoral candidate at The University of South Florida. This person is called the Principal Investigator. Megan Kirby is being guided in this research by Dr. Trina D. Spencer. Other approved research staff may act on behalf of the Principal Investigator.

Study Details: This study is being conducted online. This research is being sponsored by the Department of Child and Family Studies at University of South Florida. The purpose of the study is to test the initial effectiveness, feasibility, and acceptability of a new online caregiver training designed to improve caregiver-child storytelling conversations. The free training (three modules and 6 practice activities) will be available online and accessible at a time convenient for your family. Each training module is about 6-8 minutes long, but you can start and stop the training video multiple times if you need to. The family activities are optional and take about 5-6 minutes on average to complete with your child. At the start of the study, you will be asked to complete a questionnaire to provide basic information about you and your child, as well as a survey to gather information about your storytelling knowledge, attitudes, and practices. After completing the surveys, you will be asked to record a video of you and your child having a conversation about a past event or experience. For example, you may record a conversation about a time that you and your child went to the beach or went to another family member’s home. You will be asked to record three videos per week (recorded on different days of the week) for the duration of the study (8 weeks). The videos can be as short or as long as your conversation with your child. At the end of the study, you will be asked to share your opinions about the training modules and activities so that the research team can improve the intervention. Four weeks after the study has ended, the research team will ask you to record three more videos of storytelling conversations with your child and share those videos with the team.

Participants: You are being asked to take part, and to allow your child to take part, because you are a primary caregiver over 18 years of age, you have a child between 3-5 years old, you and your child primarily speak English in your home, you have a personal device to access the training online (e.g., smartphone, laptop or desktop computer, or tablet) and you are able to record and share videos with the research team.

Voluntary Participation: You and your child’s participation is voluntary. You and your child do not have to participate and may stop your participation at any time. There will be no penalties or loss of benefits or opportunities if you and your child do not participate or decide to stop once you start. There are no known alternatives to participating in the study.
Benefits, Compensation, and Risk: While access to the online training is of no cost to you (i.e., free), you may incur data usage costs as determined by your internet and/or phone service provider. We do not know if you or your child will receive any benefit from participation. You will be compensated up to $100 for your participation, based on the number of tasks you complete throughout the study (i.e., videos, questionnaires). This research is considered minimal risk. Minimal risk means that study risks are the same as the risks you face in daily life.

Confidentiality: Even if we publish the findings from this study, we will keep you and your child’s study information private and confidential. Anyone with the authority to look at you and your child’s records must keep them confidential.

Why are you & your child being asked to take part?

The purpose of this study is to develop an effective early childhood language intervention that can be delivered by parents to their children without the need for in-person or face-to-face training from professionals. Primarily, your participation in this study will help us evaluate how well the intervention works to improve storytelling and parent-child conversations about past events. In addition, your experience and feedback will allow us to make necessary changes to the intervention to improve its acceptance and effectiveness for other families who may use it in the future.

Study Procedures:

If you and your child take part in this study, you will be asked to:

- Week 1: Meet with the Primary Investigator, complete surveys to provide the research team with more information about you and your child, and record three videos of you and your child telling a story about a past event. Note: the videos must be recorded on three different days.
- Week 2: Record three conversations with your child about a past event on three different days.
- Weeks 3-4: Use the hyperlink provided to you by the research team to access the first online training module. Complete the training module and use the provided online activities to practice storytelling with your child.
  - Week 3: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 1A-C.
  - Week 4: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 2A-C. Access to Module 1 will terminate at the end of Week 4.
- Weeks 5-6: Use the hyperlink provided to you by the research team to access the second online training module. Complete the training module and use the provided online activities to practice storytelling with your child.
  - Week 5: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 3A-C.
Week 6: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 4A-C. Access to Module 2 will terminate at the end of Week 6.

- Weeks 7-8: Use the hyperlink provided to you by the research team to access the third and last online training module. Complete the training module and use the provided online activities to practice storytelling with your child.
  - Week 7: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 5A-C.
  - Week 8: Record three conversations with your child about a past event on three different days. Practice new skills with your child using family activities 6A-C. Access to Module 2 will terminate at the end of Week 8.

- Follow-up: Complete end-of-study surveys and meet with the researcher(s) as necessary to share your opinions about and experience with the training and related family activities. Record three final videos of storytelling conversations one month after the study ends.

**Total Number of Participants**

About 10 caregiver-child dyads (10 adults and 10 children) will take part in this study at USF.

**Alternatives / Voluntary Participation / Withdrawal**

You and your child do not have to participate in this research study.

**Benefits**

We are unsure if you or your child will receive any benefits by taking part in this research study.

**Risks or Discomfort**

This research is considered to be minimal risk. That means that the risks associated with this study are the same as what you face every day. There are no known additional risks to those who take part in this study.

**Compensation**

We will compensate you and your child with a $100 Amazon gift card if you complete all study-related tasks: submitting three videos per week for 9 weeks, completing questionnaires and surveys, and having a final meeting with the researcher(s) to discuss your experience. If you or your child withdraw for any reason from the study before completion, you will be paid $5 for each video recording that you submit.

**Cost**

While access to the online training is of no cost to you (i.e., free), you may incur data usage costs as determined by your internet and/or phone service provider.

**Privacy and Confidentiality**

We will do our best to keep you and your child’s records private and confidential. We cannot guarantee...
absolute confidentiality. You or your child's personal information may be disclosed if required by law. Certain people may need to see you or your child's study records. These individuals include:

- The research team, including the Principal Investigator, co-Investigator, and all research assistants assigned to the study team by the Primary Investigator.
- Certain government and university people who need to know more about the study. For example, individuals who provide oversight on this study may need to look at your records. This is done to make sure that we are doing the study in the right way. They also need to make sure that we are protecting your rights and your safety.
- The USF Institutional Review Board (IRB) and its related staff who have oversight responsibilities for this study, and staff in USF Research Integrity and Compliance.

We will remove all identifying information from data collection forms, such as surveys and conversation transcription. After such removal, data resulting from the study may be used for future research studies without additional informed consent from the subject or the Legally Authorized Representative. Only the research team will have access to the videos and transcriptions for up to five years after collection, after which time the videos will be permanently deleted from encrypted storage on the USF server.

Note: If you or your child refer to each other by name in the video recording, we will use your study participant identification number in place of such when transcribing your conversations. Although we cannot edit the videos to remove names, only authorized study team members will have access to watch/listen to the conversations as saved to a password-protected secured study folder saved on the University Box Drive. Unless you give explicit permission to the team to use some or all of your video in a revision to the training (e.g., as a model for other families who may use the intervention in the future), no one else will be able to watch or listen to the videos you provide. Permission or denial to use your videos for any other reason than specified above is presented at the end of this form.

If completing the study online, it is possible, although unlikely, that unauthorized individuals could gain access to your responses. Confidentiality will be maintained to the degree permitted by the technology used. No guarantees can be made regarding the interception of information sent via the Internet. However, your participation in this study involves risks similar to a person's everyday use of the Internet.

We may publish what we learn from this study. If we do, we will not include you or your child's name. We will not publish anything that would let people know who you are.

**What if new information becomes available about the study?**

During the study, we may find more information that could be important to you or your child. This includes information that, once learned, might cause you to change your mind about you or your child's participation in this study. We will notify you as soon as possible if such information becomes available.

**You can get the answers to your questions, concerns, or complaints.**

If you have any questions, concerns or complaints about this study, please call or text Megan Sullivan Kirby, Primary Investigator, at [Number]. If you have questions about your rights, complaints, or
issues as a person taking part in this study, call the USF IRB at (913) 974-5638 or contact by email at RSCH-IRB@usf.edu.

Consent to Participate and Parental Permission for My Child to Participate in this Research Study

I freely give my consent to take part and to let my child take part in this study. I understand that by signing this form I am agreeing to take part and to let my child take part in research. I have received a copy of this form to take with me.

Signature of Adult Participant/Parent of Child Taking Part in Study

Date

Printed Name of Adult Participant Taking Part in Study

Printed Name of the Child Taking Part in Study

Statement of Person Obtaining Informed Consent

I have carefully explained to the person taking part in the study what he or she can expect from their participation. I confirm that this research subject speaks the language that was used to explain this research and is receiving an informed consent form in their primary language. This research subject has provided legally effective informed consent.

Signature of Person Obtaining Informed Consent

Date

Printed Name of Person Obtaining Informed Consent
Additional Consent: Use of Research Recordings for Educational Purposes or Presentation Purposes

During your participation in the study, you will voluntarily provide the researchers with videos of you and your child engaged in conversation. We may wish to present segments of video from this study at professional meetings. We may want to use video segments of your participation in the research for education and training of future researchers/practitioners. Videos will only be shared by the PI or co-PI via a password protected file which cannot be copied or downloaded. Your face and voice will be used and you potentially could be recognizable by a viewer of the video recording.

In addition to consenting to participate in the research study, you have the option of consenting to these uses of recordings. Please check whether you do/don’t consent to use of video segments as follows:

- [ ] I agree that segments of the recordings made of my participation in this research may be used for conference presentations.
- [ ] I do not want segments of the recordings made of my participation in this research to be used for conference presentations.
- [ ] I agree that segments of the recordings made of my participation in this research may be used for education and training of future researchers/practitioners.
- [ ] I do not want segments of the recordings made of my participation in this research to be used for education and training of future researchers/practitioners.